

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2023

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, 30.06.2023



Pre Úrad splnomocnenca vlády SR pre rozvoj občianskej spoločnosti
spracoval **Lukáš Jankovič**

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

HVD – „high value datasets“ – datasety s vysokou hodnotou podľa smernice Európskeho parlamentu a Rady EÚ o otvorených dátach a opakovanom použití informácií verejného sektora

1	Úvod	5
2	Kvantitatívna analýza	7
2.1	Počet zverejnených datasetov	7
2.2	Počet zverejnených datasetov podľa OVM	9
2.3	Tagy (kľúčové slová)	11
2.4	Licencie datasetov	13
2.5	Formáty datasetov	14
2.6	Aktualizácia datasetov	16
2.7	Zhrnutie	18
3	Kvalitatívna analýza	19
3.1	Hodnotenie datasetov podľa hviezdíčiek	19
3.2	Najčastejšie problémy znižujúce použiteľnosť datasetov	21
3.3	Rebríčky	22
	Open data maturity index	22
	OECD OURdata Index - vládne otvorené dáta	24
	Open data watch	24
	Index digitálnej ekonomiky a spoločnosti (DESI)	25
	Ostatné rebríčky	25
3.4	Zhrnutie	25
4	Používanie portálu	27
4.1	Počet registrovaných používateľov portálu	27
4.2	Návštevníci podľa krajín	27
4.3	Počet zobrazení podľa rokov	28
4.4	Najpopulárnejšie datasety	29
4.5	Počet aplikácií	31
4.6	Zhrnutie	31
5	Štandardy a iniciatívy v oblasti otvorených dát	33
5.1	Publikačné minimum štátnej správy	33
5.2	Datasety s vysokou pridanou hodnotou	34
5.3	Prieskum dopytu verejnosti po najžiadanejších datasetoch a API	34



Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2023

5.4	Národný projekt Otvorené údaje 2.0	35
5.5	Školenia	36
5.6	Akčný plán Iniciatívy pre otvorené vládnutie na roky 2022-2024	37
5.7	Občianske združenie Alvaria	38
5.8	Opencode.gov.sk – metodický portál pre zverejňovanie zdrojových kódov ISVS	39
5.9	Zhrnutie	39
6	Záver	41
7	Odporúčania	42
7.1	Opakujúce sa odporúčania z poslednej analýzy	42



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. 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, 2022. 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.

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 2023

Obrázok 1: <https://data.gov.sk>

Správcom portálu data.gov.sk je *Národná agentúra pre sieťové a elektronické služby* (NASSES), 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 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, 24.5.2023.

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

V súvislosti s projektom nového centrálného portálu⁴ otvorených dát bol v čase vzniku tejto analýzy vybudovaný backend a SPARQL Endpoint⁵ nového portálu. Nový portál je postavený na novovzniknutom štandarde DCAT-AP-SK-2⁶, ktorý sa odlišuje od pôvodného štandardu. Zmenila sa definícia datasetu a distribúcie, následkom čoho sa zvýšil počet datasetov a distribúcií. **Z dôvodu spätnej kontinuity uvádzame v tejto kapitole oba počty - podľa súčasného portálu data.gov.sk (označené ako „SP“) a podľa SPARQL Endpointu nového portálu (označené ako „NP“).**

2.1 Počet zverejnených datasetov

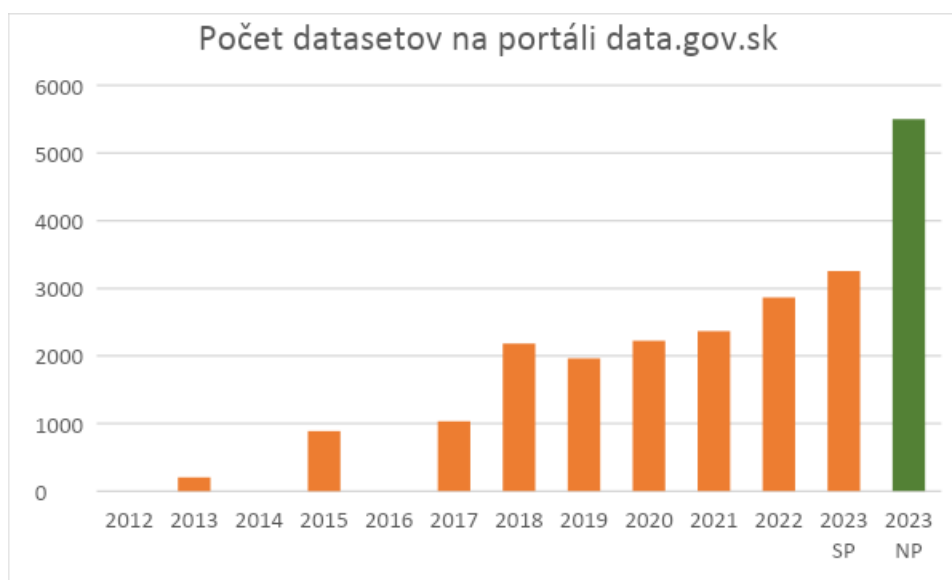
Aktuálny počet datasetov k 24.5.2023 bol zistený z portálu data.gov.sk a z backendu nového portálu. Počty datasetov za minulé roky sú prebraté z analýzy za rok 2022. **Počet datasetov postupne rastie, v roku 2023 prekonal hranicu 3000 datasetov.**

⁴ vid' kapitolu ...

⁵ <https://opendata.mirri.tech/>

⁶ <https://datova-kancelaria.github.io/dcat-ap-sk-2.0/>

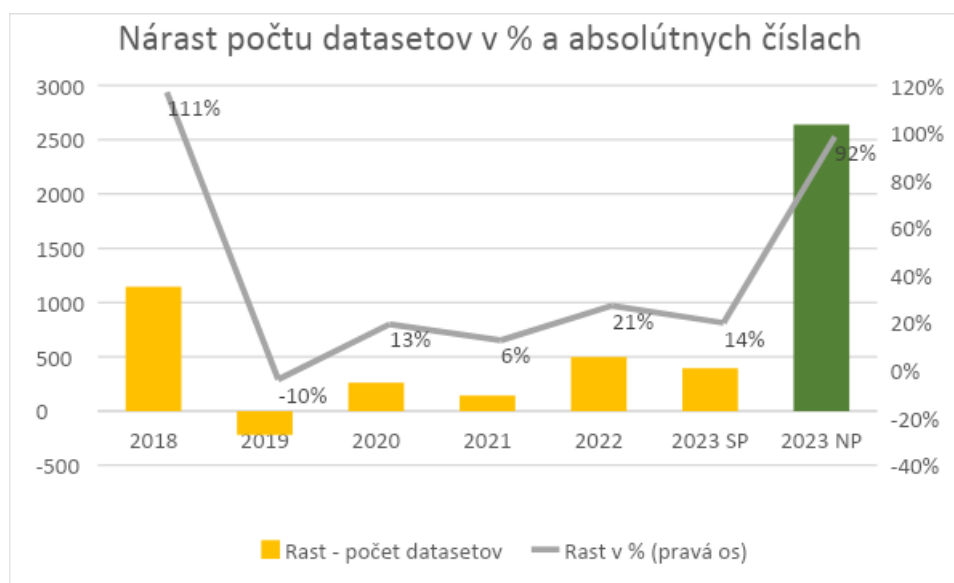
Graf 1



SP – súčasný portál, NP – SPARQL Endpoint nového portálu

Pokiaľ sa pozrieme na percentuálny nárast v porovnaní s minulým rokom, zaznamenali sme **o niečo menší rast ako v minulom roku 2022 – 14%**. V minulom roku bol rast počtu datasetov 500, tento rok to bolo o viac ako 100 menej – 395. Podľa nového portálu bol počet datasetov viac ako 5500, toto navýšenie z veľkej časti vzniklo zmenou definície datasetu a distribúcie a nie zverejnením nových datasetov.

Graf 2



SP – súčasný portál, NP – SPARQL Endpoint nového portálu

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 pozerať 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.sk⁷, 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

⁷https://www.slovensko.sk/sk/agendy/agenda/_ministerstva-a-ostatne-ustredne1

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2023

spoločnosť, ktorá je správcom slovenskej internetovej domény sk. Datasetsy na portáli zverejňujú aj územné samosprávy.⁸

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

OVM alebo názov organizácie	2013	2015	2017	2018	2019	2020	2021	2022	2023 SP	2023 NP	Ras t SP
Generálna prokuratúra SR				8	8	8	8	8	8	8	0
Kancelária Národnej rady SR					0	10	10	10	10	168	0
Ministerstvo dopravy a výstavby SR	8	10	10	32	66	60	77	83	87	558	4
Ministerstvo financií SR ⁹	11	22	22	20	13	14	15	24	24	24	0
Ministerstvo hospodárstva SR	7	8	9	17	17	34	37	37	39	43	2
Ministerstvo investícií, regionálneho rozvoja a informatizácie Slovenskej republiky	0	23	12	36	39	36	58	100	144	144	44
Ministerstvo kultúry Slovenskej republiky	10	11	11	35	25	26	26	26	31	32	5
Ministerstvo obrany SR	3	4	4	19	8	8	8	9	9	41	0
Ministerstvo pôdohospodárstva a rozvoja vidieka SR	65	68	67	209	457	469	475	491	468	1137	-5
Ministerstvo práce, sociálnych vecí a rodiny SR	5	5	5	5	13	13	13	13	16	19	3
Ministerstvo spravodlivosti SR	5	8	8	17	16	17	21	26	27	43	1
Ministerstvo školstva, vedy, výskumu a športu Slovenskej republiky	5	24	24	34	26	26	23	24	85	85	61
Ministerstvo vnútra SR	24	40	41	253	61	62	62	62	62	418	0
Ministerstvo zahraničných vecí a európskych záležitostí SR	14	14	16	25	25	11	12	12	1	29	-11
Ministerstvo zdravotníctva SR	9	18	95	184	262	381	408	435	461	1128	26
Ministerstvo životného prostredia SR	9	9	22	41	25	42	52	82	90	100	8
Najvyšší kontrolný úrad SR					2	10	10	32	46	46	14
Národný bezpečnostný úrad				4	23	23	23	23	23	24	0
Protimonopolný úrad SR				7	7	10	10	50	50	106	56
SK-NIC								1	1	1	0
Správa štátnych hmotných rezerv Slovenskej republiky				5	7	10	14	15	15	17	0
Štatistický úrad SR	9	586	606	1101	665	745	764	1023	1139	1141	116
Úrad geodézie, kartografie a katastra SR	2	2	2	2	2	5	19	19	26	27	7
Úrad jadrového dozoru Slovenskej republiky				6	7	7	7	12	13	13	1
Úrad pre normalizáciu, metrológiu a skúšobníctvo Slovenskej republiky				3	7	7	7	10	13	32	3

⁸ Ide o mestá Prešov, Trnava, Vranov nad Topľou, Nitriansky samosprávny kraj, Prešovský samosprávny kraj, obce Horné Strháre, Kolačkov

⁹ Ministerstvo financií SR v apríli 2023 sprístupnilo verejnosti vlastný portál otvorených dát (<https://opendata.mfsr.sk/>), ktorý ku dňu 25.7.2023 obsahuje 116 datasetov. Ich poskytovateľmi sú úrad MF SR a jednotlivé podriadené organizácie. Datasetsy na tomto portáli nezapočítavame do datasetov v tabuľke, keďže zatiaľ nie je harvestovaný na centrálny portál data.gov.sk ani do SPARQL Endpointu nového portálu. Práce na zabezpečení harvestovania prebiehajú.

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2023

Úrad pre verejné obstarávanie	7	7	46	16	79	90	103	127	139	140	12
Úrad priemyselného vlastníctva Slovenskej republiky				19	8	9	10	15	18	123	3
Úrad vlády SR	7	11	11	13	15	18	22	22	22	40	0
Všeobecná zdravotná poisťovňa								1	1	1	0
Nezaradené	4	1	1	3	11				15	21	6
územná samospráva	0	18	22	67	68	71	71	67	102	116	35
Spolu	204	889	1034	2181	1962	2232	2375	2869	3259	5825	494

SP – súčasný portál, NP – SPARQL Endpoint nového portálu

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.sk¹⁰, 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 95 (o 1 viac ako v roku 2022).

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**, **Úrad pre verejné obstarávanie** a **Protimonopolný úrad**. Tento zoznam sa oproti minulému roku zmenil len v tom, že do neho Pribudol Protimonopolný úrad, ktorý 50 datasetom zverejňovaným v minulom roku pridal 56 datasetov.

Najväčším skokanom za predchádzajúci rok je **Štatistický úrad SR**, ktorý pridal 116 datasetov, 61 datasetov pridalo **Ministerstvo školstva, vedy a výskumu SR**, 56 datasetov **Protimonopolný úrad SR**, 44 datasetov pridalo **Ministerstvo investícií, regionálneho rozvoja a informatizácie SR**. Väčší počet datasetov ako 20 pridali spoločne už len organizácie miestnej samosprávy (35 datasetov).

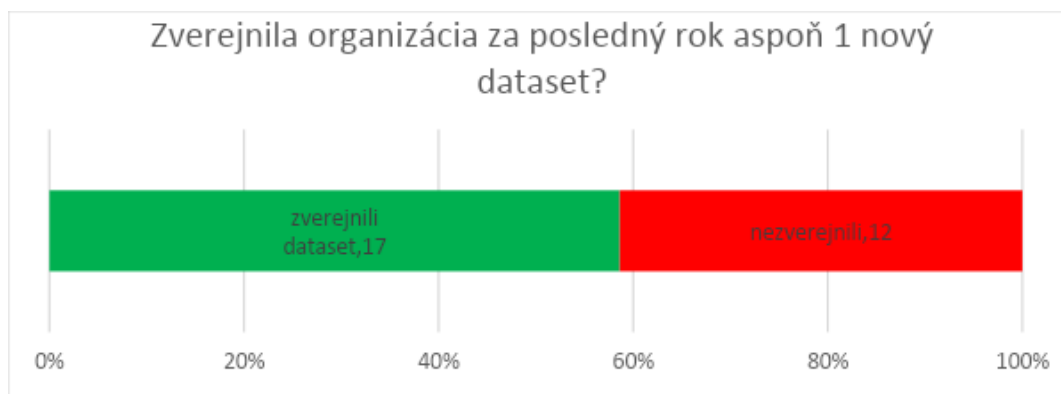
Počty datasetov pri organizáciách v drvivej väčšine rástli alebo stagnovali, pokles zaznamenalo Ministerstvo zahraničných vecí a európskych záležitostí SR, ktorý je však dôsledkom vzniku samostatného portálu, ktorého datasety na centrálnom portáli nie sú katalogizované. Mierny pokles o 5 datasetov zaznamenalo Ministerstvo pôdohospodárstva a rozvoja vidieka SR.

Väčšina organizácií však zverejnila aspoň 1 nový dataset. **Počet organizácií, ktoré pridali aspoň 1 dataset však medziročne klesol.** Môže to byť ovplyvnené finálnou fázou budovania nového centrálného portálu otvorených dát a opatrnosťou organizácií, môže

¹⁰<https://www.slovensko.sk/sk/institucie>

však ísť aj o pokles záujmu o zverejňovanie otvorených dát. Toto bude možné zhodnotiť až v budúcoročnej analýze.

Graf 3



Zdroj???

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 Datasety umožňuje filtrovať datasety podľa tagu (kľúčového slova) pre ľahšie vyhľadávanie¹¹, 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 zistených podľa tagu sa líši od celkového počtu datasetov, zisteného podľa organizácií, pretože daný dataset môže mať priradených aj viac tagov.

Najväčší podiel majú datasety s tagom **štatistika, zdravotníctvo a voľby**, čo sú tagy asociované najmä s datasetmi Štatistického úradu SR a Ministerstva zdravotníctva SR (s podriadenými organizáciami). V populárnych tagoch sa tiež vyskytuje **cesta, register adries, publikačné minimum, eurofondy či geografická reprezentácia**.

Celkový počet tagov je až 12 053, v tagoch sa vyskytujú duplicity či preklepy. Je otáznne, 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

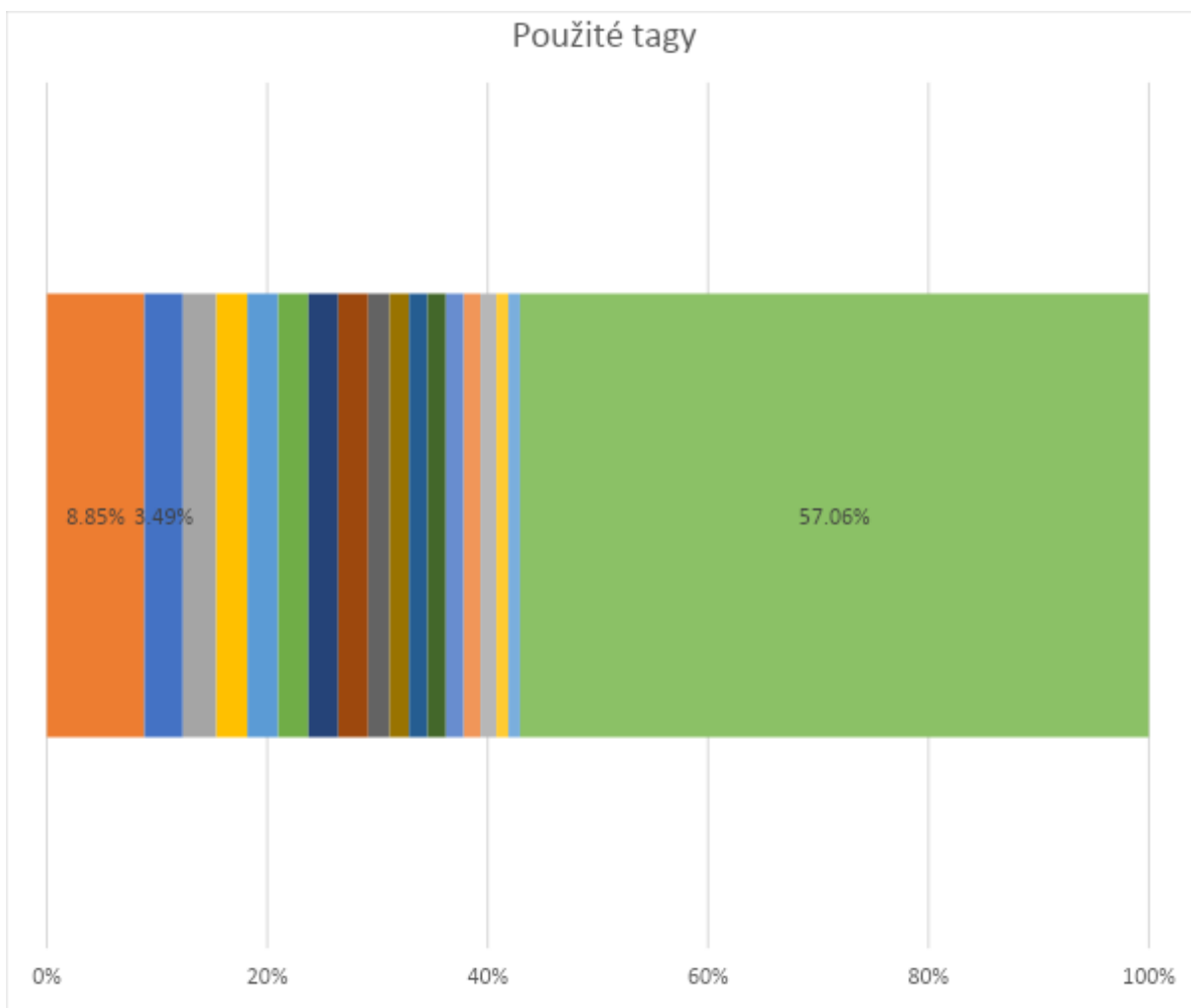
¹¹https://data.gov.sk/dataset?_tags_limit=0



Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2023

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

Graf 4



Zdroj dát: SPARQL Endpoint nového portálu

V porovnaní s analýzou za minulý rok vzrástol celkový počet tagov z 6 001 na viac ako 12 053, čo bolo ovplyvnené najmä preklopením údajov do backendu nového centrálného portálu a zvýšením počtu datasetov. V súvislosti so zvýšením počtu datasetov (viď. kap. 2.1.) prišlo k zvýšeniu podielu niektorých tagov (napr. cesta, doprava, mdvsr, register adries, atď.).

2.4 Licencie datasetov

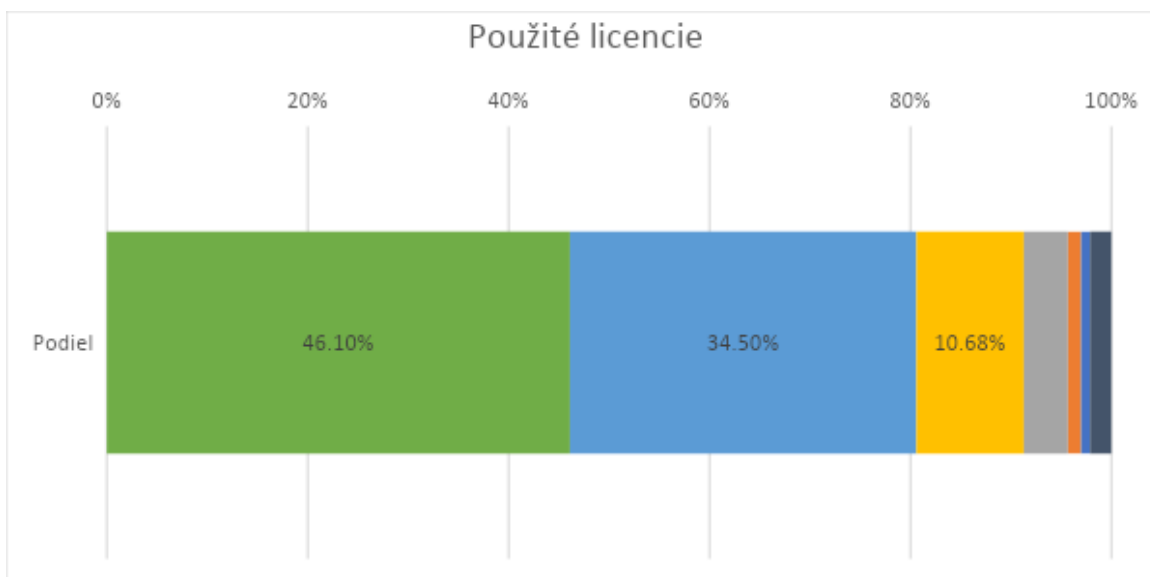
Licencia znamená právny režim použitia údajov. Aby používateľ vedel, ako môže údaje používať a čo musí dodržať, je nutné aby boli pri datasetoch uvedené licencie. Licencia by mala umožňovať komerčné aj nekomerčné použitie údajov, čo spĺňajú napríklad licencie Creative Commons CC0 alebo CC-BY.



Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2023

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



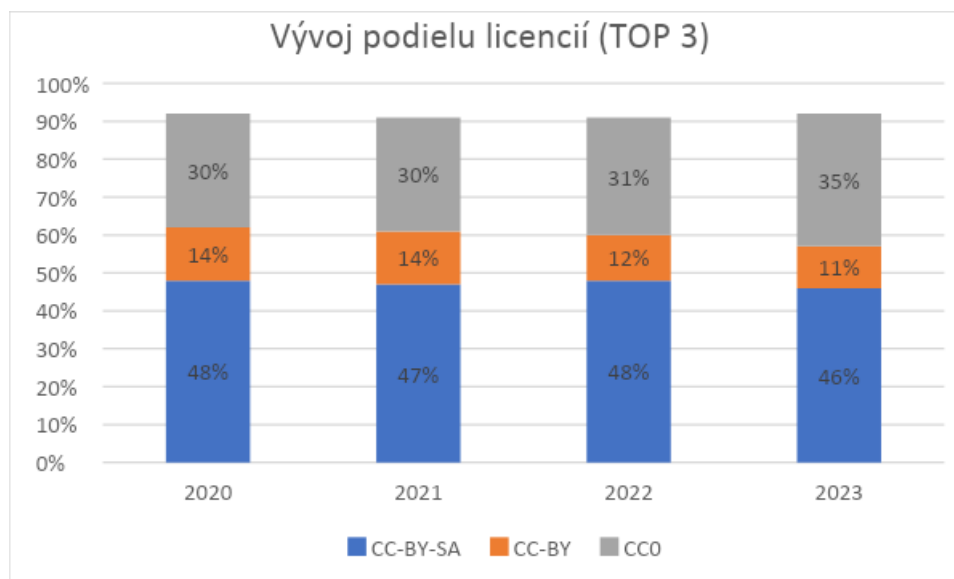
Zdroj dát: Portál data.gov.sk

Najpoužívanejšia 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 4,41 %. Pri 0,10 % 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 štandardmi¹². Odporúčané licencie sú CC0, CC-BY. Veľký problém je vysoký podiel licencií CC SA, ktoré tvoria až necelú polovicu datasetov 46%.

¹²Vyhláška 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



Zdroj dát: portál data.gov.sk

Podiel licencií sa mení pomaly. Rastie podiel licencií CC0, znižuje sa podiel CC-BY a CC BY-SA. **Tento vývoj nie je pozitívny, žiadúce by bolo, aby sa licencie rýchlejšie preklápali smerom k odporúčaným licenciám CC-0 a CC-BY.**

Kvôli momentálnej nedostupnosti dát z SPARQL Endpointu nového portálu používame v tabuľke vyššie dáta získané zo súčasného portálu.

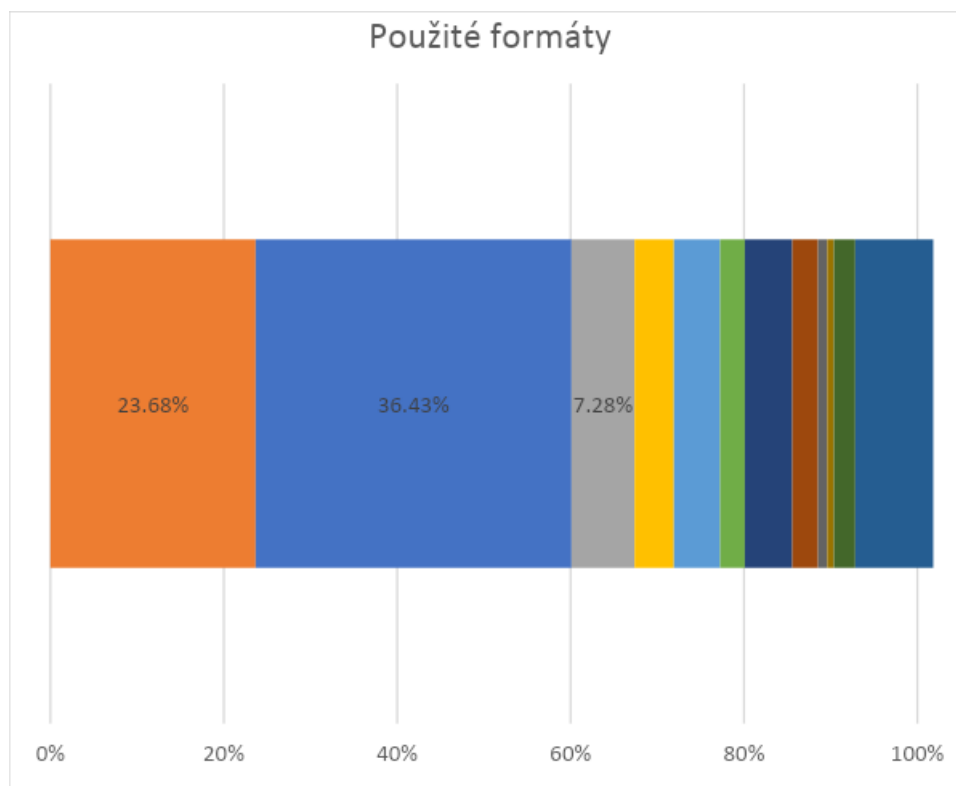
2.5 Formáty datasetov

Formát je spôsob uloženia dát (štruktúra) do súboru. Formát z väčšej časti 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/distribúcií).

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



Zdroj dát: API portálu data.gov.sk

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. Za posledný rok nevidujeme zásadnú zmenu v tomto nelichotivom stave.** Okrem týchto najpoužívanejších formátov sa používajú aj mnohé iné formáty, ktoré sú strojovo spracovateľné, napr. geojson, kml, shapefile a ďalšie.

Poriadok do označenia formátov zavádza backend nového portálu a štandard DCAT-AP-SK-2.0, ktorý používa predefinovaný číselník, zamedzí sa tak duplicitám a preklepom. Kvôli momentálnej nedostupnosti dát z SPARQL Endpointu nového portálu používame v tabuľke vyššie dáta získané z API súčasného portálu.

Tabuľka 2: Predefinované formáty na novom portáli

http://www.iana.org/assignments/media-types/application/msword
http://www.iana.org/assignments/media-types/text/html
http://www.iana.org/assignments/media-types/application/pdf?type=pdf1x
http://www.iana.org/assignments/media-types/image/tiff
http://www.iana.org/assignments/media-types/text/rtf
http://www.iana.org/assignments/media-types/application/vnd.ms-excel
http://www.iana.org/assignments/media-types/application/xml



http://www.iana.org/assignments/media-types/text/csv
http://www.iana.org/assignments/media-types/application/vnd.openxmlformats-officedocument.spreadsheetml.sheet
http://www.iana.org/assignments/media-types/application/geo+json
http://www.iana.org/assignments/media-types/application/vnd.openxmlformats-officedocument.wordprocessingml.document.main+xml
http://www.iana.org/assignments/media-types/application/xhtml+xml
http://www.iana.org/assignments/media-types/application/json
http://www.iana.org/assignments/media-types/text/plain
http://www.iana.org/assignments/media-types/application/zip

Porovnanie po rokoch neuvádzame, keďže v minulých rokoch boli použité dáta získané z portálu data.gov.sk (filter na podstránke Datasets vľavo), ktoré však vyhodnocujú počet odlišne ako API. Údaje by teda neboli porovnateľné.

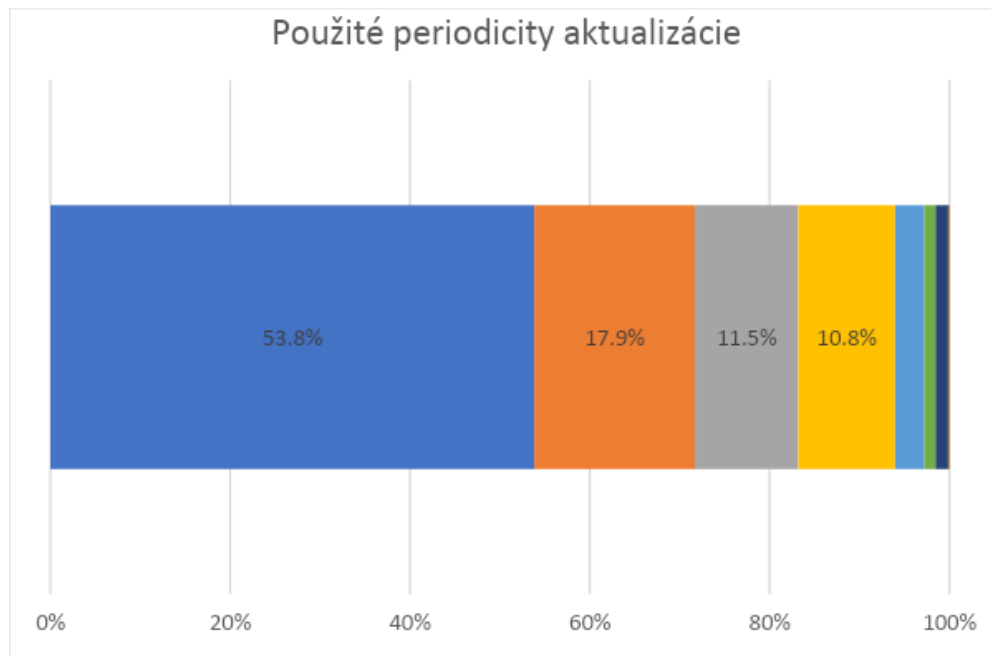
2.6 Aktualizácia datasetov

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 periodičita dostatočná pre konkrétne použitie, ktoré má používateľ v pláne.

Na portáli má každá distribúcia (súbor, odkaz) vlastnú hodnotu aktualizácie, ktorú sme získali z API portálu.

Viac ako 70 % distribúcií má hodnotu iné a nepravidelne, čo sú hodnoty, ktoré by sa mali využívať, ak nie je možné vybrať konkrétnejšiu periodicitu aktualizácie. Z konkrétnejších periodícít je najviac zastúpené ročne a mesačne (okolo 11 %), denne je aktualizovaných 3,3 % distribúcií, každého pol roka a štvrtročne je to okolo 1 %, týždenne len 0,1 % dátových zdrojov.

Graf 8

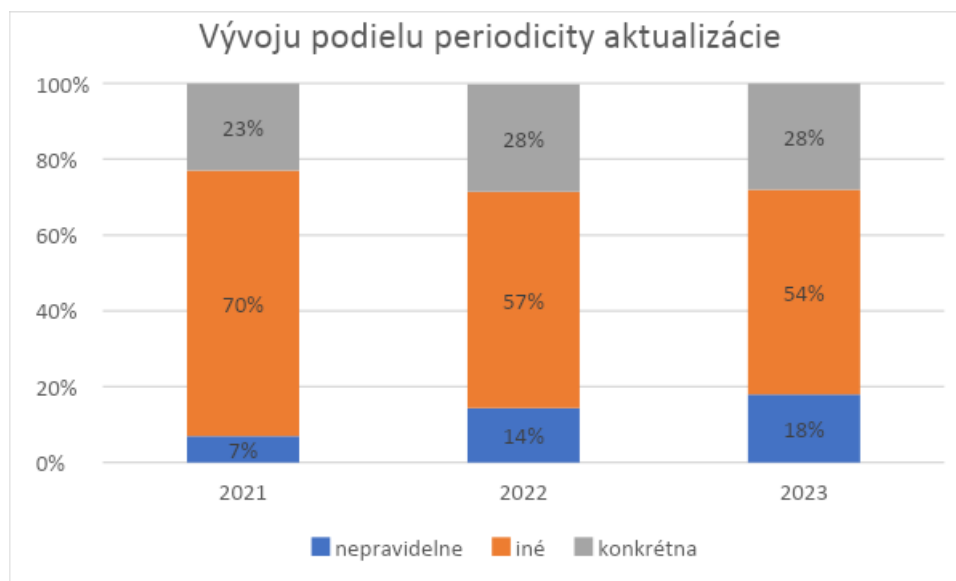


Zdroj dát: API portálu data.gov.sk

Vývoj podielu periodicity aktualizácie distribúcií je mierne negatívny. **Za posledné 2 roky klesol podiel hodnoty iné na úkor hodnoty nepravidelne.** Za posledný rok sa však podiel konkrétnych hodnôt nezmenil. Stále ostáva väčšina distribúcií, ktorá má nepravidelnú aktualizáciu alebo hodnotu iné. Tieto distribúcie by bolo treba analyzovať a určiť, či je perióda aktualizácie určená správne, respektíve či sú podľa vyplnených metadát aj pravidelne aktualizované, čo je veľmi dôležité pre ich použiteľnosť.

Kvôli momentálnej nedostupnosti dát o aktualizácii distribúcií zo SPARQL Endpointu nového portálu používame v tabuľke vyššie dáta získané z API súčasného portálu.

Graf 9



2.7 Zhrnutie

Za posledný rok **pribudlo** na portáli približne **400 datasetov**, čo je mierne menej v porovnaní s minulým rokom. Pozitívne je tiež to, že **väčšina OVM publikujúcich na portáli zverejnila aspoň 1 nový dataset**. Vplyvom inej definície datasetu a distribúcie počet datasetov prekonal hranicu 5 500 (predtým viac ako 3 000).

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, Ministerstvo vnútra, Ministerstvo dopravy a Štatistický úrad.

Zloženie licencií sa mení len pomaly, **stále ostáva vysoký (nadpolovičný !) podiel licencií, ktoré nie sú dostatočne otvorené**. Počet distribúcií zverejnených v zle strojovo spracovateľnom formáte zostáva niekoľko stoviek (viac ako 700), s pribúdajúcimi distribúciami však klesá ich podiel. Väčšina dátových zdrojov má nastavených ako periódu aktualizácie nepravidelne alebo iné, čo môže pre používateľov znamenať problém s predvídateľnosťou aktualizácie.



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 hviezdíčkami¹³:

- * 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 vyhláske č. 78/2020 Z. z. (Vyhláska o štandardoch pre informačné technológie verejnej správy)¹⁴, § 38 až § 40.

¹³<https://5stardata.info/cs/>

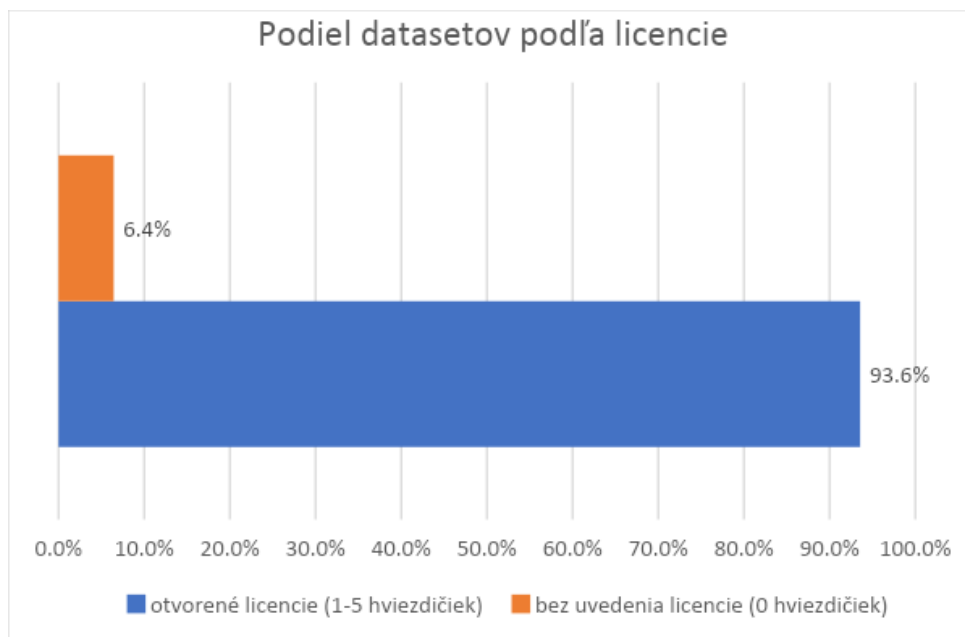
¹⁴<https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2020/78/20200501>

Obrázok 2: Hodnotenie otvorených dát hviezdčkami

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

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

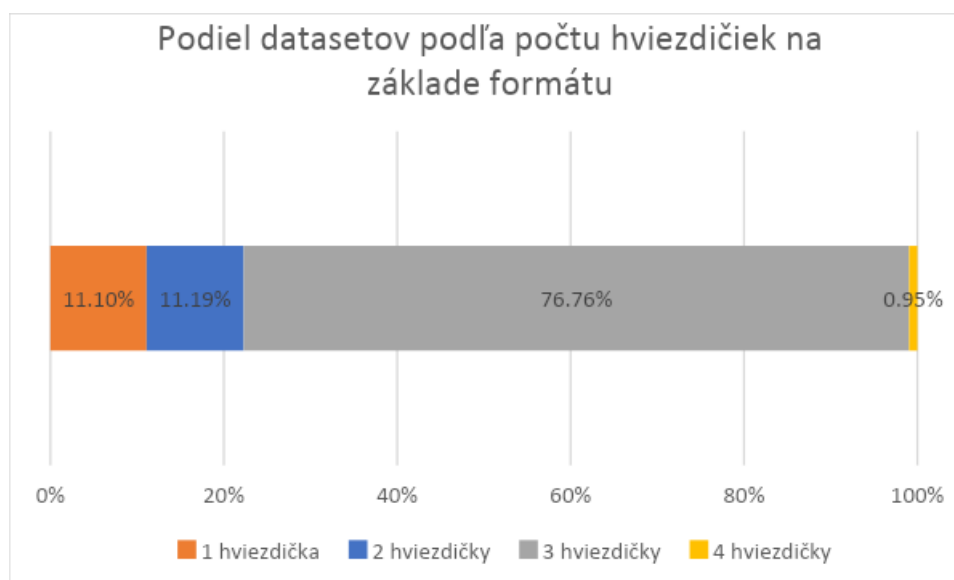
Graf 10



Oproti minulému roku nebol zaznamenaný nárast ani pokles podielu datasetov bez licencie. **Datasety bez licencie však stále predstavujú problém, keďže takéto publikovanie datasetov nie je v súlade s princípmi otvorených dát.**

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 hviezdíček prepojené dáta.

Graf 11



Najväčší podiel majú 3 hviezdíčky, viac ako 3/4, čo sú plnohodnotné otvorené dáta. 2 hviezdíčky, ktoré síce nie sú plnohodnotné otvorené dáta, ale sú ešte stále dobre strojovo spracovateľné sa vyskytujú pri takmer 12 % datasetov. **Približne 11 % datasetov dosahuje len 1 hviezdíčku, čo je tiež podiel datasetov, ktoré je treba nahradiť datasetmi s vyšším počtom hviezdíčiek.**

Formát RDF, ktorý poskytuje ontológiu pre metadáta ku datasetu, je na portáli data.gov.sk len pri 24 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 hviezdíčiek je použitie ontológie pri definícii údajov (metadátach), použitie referencovateľného identifikátora (permanentnej URI/URL) pre dataset a použitie referencovateľných identifikátorov v dátach. Zverejnené 3-hviezdíčkové datasety na portáli data.gov.sk majú spravidla permanentnú URI adresu, také možno považovať za 4-hviezdíčkové.

Úroveň 5 hviezdíček spĺňajú 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. 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

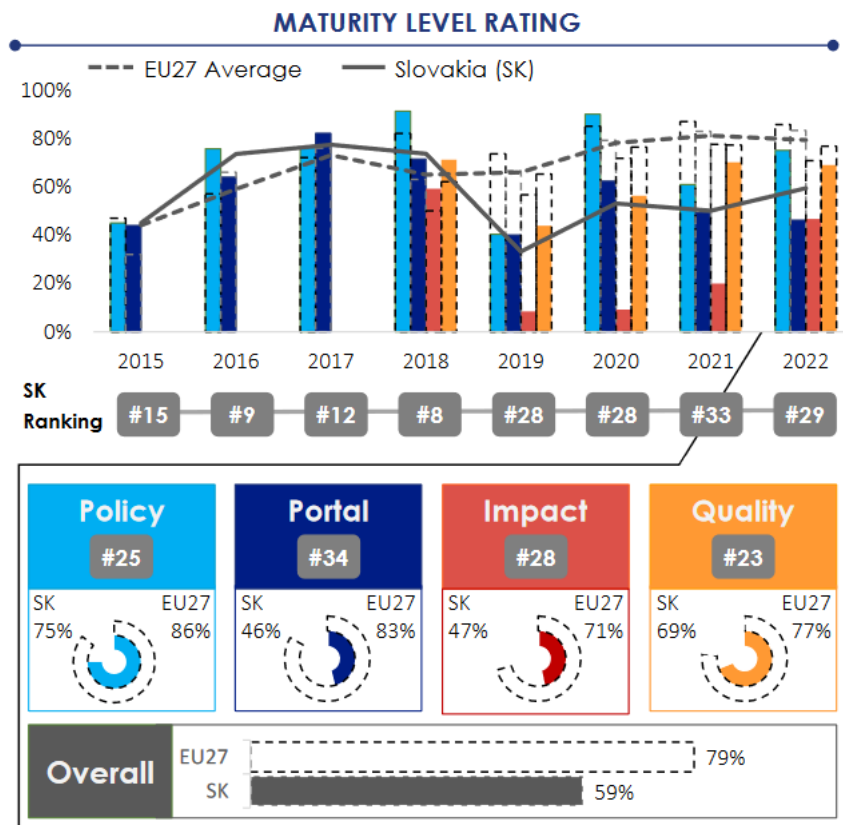
Open data maturity index¹⁵

Open data maturity index je rebríček každoročne vyhodnocovaný Európskou komisiou. Slovensko sa umiestnilo v poslednom sledovanom období **2022** na **29. priečke**, **4. najhoršia z krajín EÚ** v najslabšej skupine sledovaných európskych štátov (začiatoníci) zo štyroch skupín (lídri trendu, rýchli nasledovníci trendu, nasledovníci, začiatoníci). Metodika je založená na posúdení 4 dimenzií otvorených dát: politika, portál, dopad, kvalita.

¹⁵ <https://data.europa.eu/en/impact-studies/open-data-maturity>

Obrázok 1

State-of-Play on open data - 2022

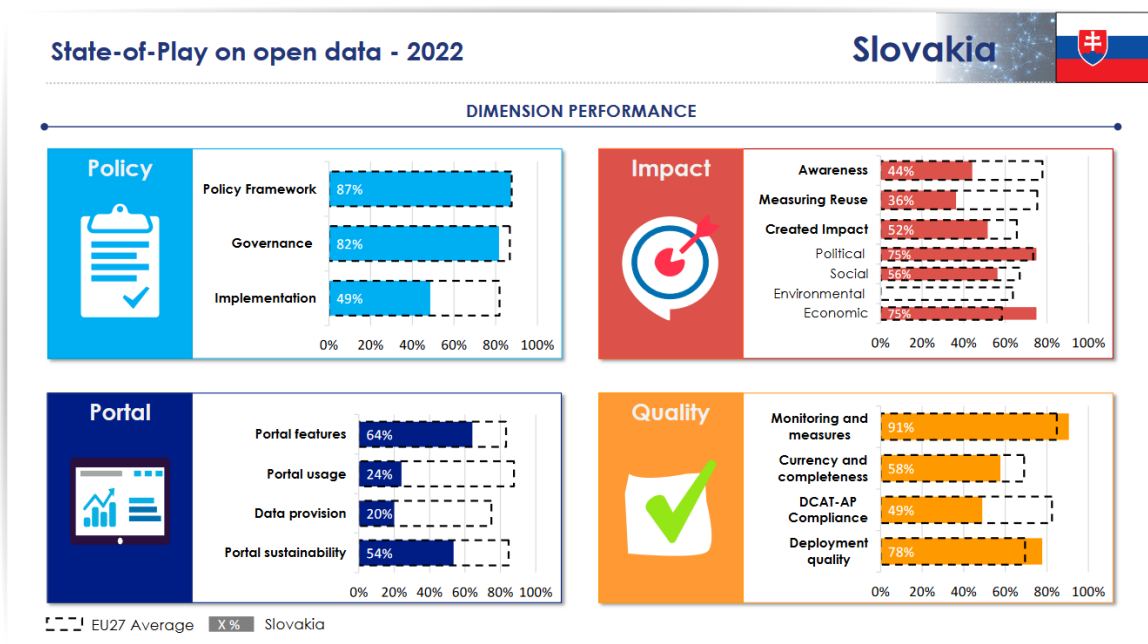


Zdroj: https://data.europa.eu/sites/default/files/country-factsheet_slovakia_2022.pdf

Slovensko sa oproti minulému roku prepadlo v hodnotení portálu a kvality, polepšilo si pri dopade a politikách.

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

Obrázok 2



Zdroj: https://data.europa.eu/sites/default/files/country-factsheet_slovakia_2022.pdf

Dobrá správa je, že oproti minulému roku si Slovensko polepšilo o 3 priečky a že toto hodnotenie sa môže v najbližších rokoch ešte viac zlepšiť, najmä vplyvom zavedenia nového centrálného portálu otvorených dát, publikačného minima a zvýšenej pozornosti venovanej dátam s vysokou pridanou hodnotou a sledovaniu vplyvu (dopadu) otvorených dát.

OECD OURdata Index - vládne otvorené dáta ¹⁶

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**.¹⁷

¹⁶ <https://www.oecd.org/gov/digital-government/open-government-data.htm>

¹⁷ <https://www.oecd.org/gov/digital-government/ourdata-index-slovak-republic.pdf>

Open data watch

Slovensko sa umiestnilo v poslednom sledovanom období **2020/21¹⁸ 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 2020¹⁹ je na základe dát Štatistického úradu SR²⁰. **Najhoršie 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) 2022²¹ patrí Slovensku **23. miesto v rebríčku z 27 krajín EÚ, o jedno miesto horšie ako pred rokom. Slovensko²² je v hodnotení DESI pod priemerom EÚ, 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 štyroch dimenzií: konektivita, ľudské zdroje (digitálne zručnosti), integrácia digitálnej technológie, verejné digitálne služby. Najlepšie hodnotenie má Dánsko, Fínsko a Holandsko.

Najhoršie je Slovensko hodnotené pri digitálnych verejných službách kam patrí aj hodnotenie otvorených dát. V tejto oblasti zaostáva vo všetkých hodnotených indikátoroch (používatelia e-governmentu, predvyplnené formuláre, digitálne služby pre občanov a firmy, otvorené dáta).

V máji 2021 vláda schválila stratégiu a akčný plán na zlepšenie postavenia Slovenska v indexe DESI do roku 2025²³. 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“,

¹⁸<https://odin.opendatawatch.com/>

¹⁹ <https://odin.opendatawatch.com/Report/Reports>

²⁰ <http://www.statistics.sk>

²¹ <https://ec.europa.eu/newsroom/dae/redirection/document/88764>

²² <https://ec.europa.eu/newsroom/dae/redirection/document/88712>

²³

<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>



- 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 barometer²⁴ a Global open data index²⁵. 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 s viacerými hviezdičkami 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).** Približne 11% distribúcií je zverejnených v nevyhovujúcej kvalite 1 hviezdičky.

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 posunulo v rebríčku **Open data maturity index na 29. priečku (4. najhoršie z krajín EÚ).** Mnohé zmeny, ktoré by mali Slovensku posunúť vyššie v rebríčku už začali. **Podarilo sa výrazne pokročiť v projekte nového centrálného portálu, monitorovania dopadu a zverejňovania HVD.** Riadiť by sme sa mali aj odporúčaniami OECD²⁶, ktoré hoci sú 4 roky staré, sú stále relevantné.

²⁴ <https://opendatabarometer.org>

²⁵ <https://index.okfn.org/>

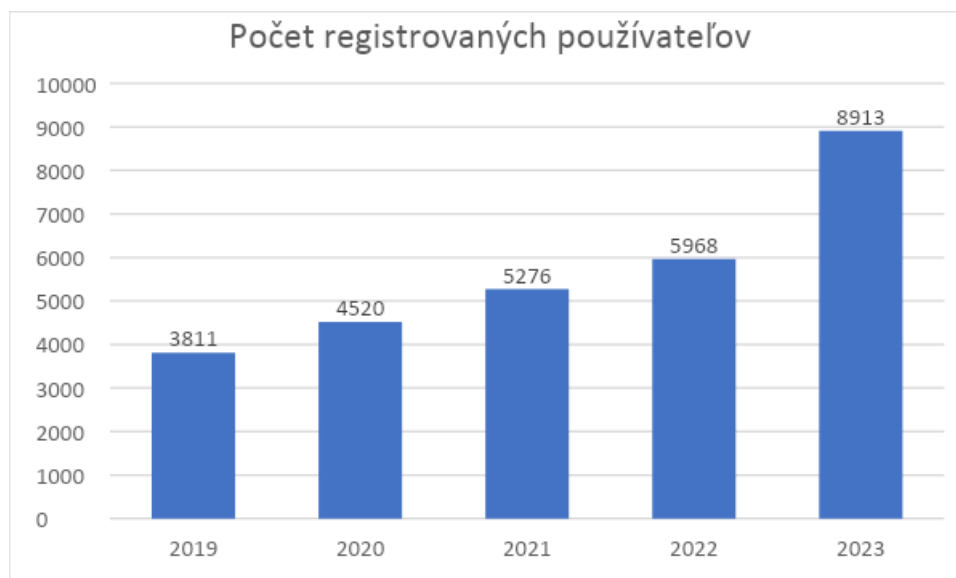
²⁶ <https://www.oecd.org/gov/digital-government/open-government-data.htm>

4 Používanie portálu

4.1 Počet registrovaných používateľov portálu

Na portáli data.gov.sk v záložke Štatistiky sa uvádza počet registrovaných používateľov portálu. K 24.5.2023 bol tento počet **8 913**, čo je o **2 945 (49 %) viac ako pred rokom**. Počet registrovaných používateľov v poslednom roku **výrazne vzrástol**.

Graf 12



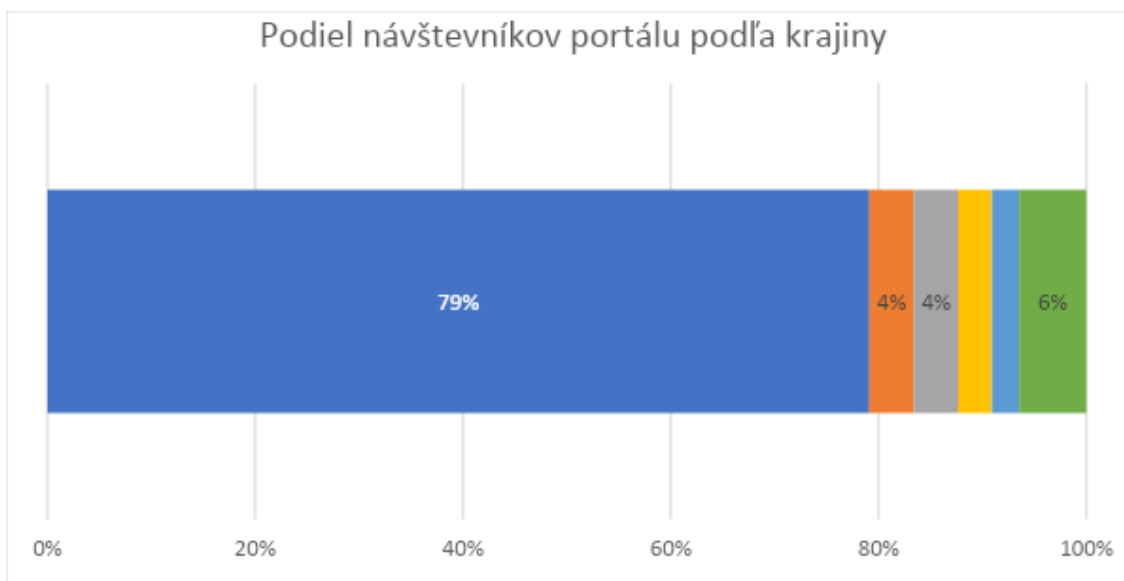
Zdroj: Data.gov.sk a minulé analýzy

4.2 Návštevníci podľa krajín

Počet návštevníkov na portáli v období od 1.1.2023 do 30.4.2023 bol 114 110. Z týchto návštevníkov sme vylúčili návštevníkov zo Singapuru a Mexika, kde predpokladáme že sa nejednalo o skutočných ľudí (priemerné trvanie návštevy bolo 0,01 s.). Brali sme do úvahy zvyšok, **16 626 návštevníkov**.

Najviac návštevníkov pochádzalo zo Slovenska – takmer 80 %. Druhou krajinou s najvyšším počtom návštevníkov bolo Francúzsko s podielom takmer 4 %. Na treťom mieste bolo Česko s podielom taktiež 4 %. Ďalšie najčastejšie krajiny sú Nemecko a USA (obe okolo 3%). Oproti minulému roku sa z popredných miest rebríčka vytratilo Španielsko a Hong-Kong, čo je pravdepodobne spôsobené tým, že išlo o prístupy strojov.

Graf 13



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 26.6.2023. 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 bol rekordným rokom, portál zaznamenal viac ako 730 tis. zobrazení, čo je nárast oproti roku 2021 o viac ako 75 %.** 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. Rok 2023, ktorý je zaznamenaný do približne polovice vyzerá byť zatiaľ slabší ako rok 2022, ale lepší ako rok 2021.

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 14



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 stiahnutí²⁷. Tento údaj nemáme bohužiaľ od prevádzkovateľa portálu (NASES) k dispozícii, máme však k dispozícii **počet zobrazení za prvé 4 mesiace roku 2023**. 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, počet prístupov UPVS za deň, zoznam vyplatených faktúr Pôdohospodárskej platobnej agentúry. Populárne sú aj datasety Národnej rady Slovenskej Republiky – programy schôdzí, register adries, informácie o epidemiologickej situácii v okrese Trebišov alebo aj informácie o čerpaní OPII.

Tabuľka 3: Počet zobrazení datasetov na portáli (TOP 20)

P.	Názov datasetu	Vlastník	Počet zobrazení
1	Register adries - Zmenové dávky registra adries	Ministerstvo vnútra SR	20 565

²⁷ 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.

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2023

2	Vestník Úradu priemyselného vlastníctva SR	Úrad priemyselného vlastníctva Slovenskej republiky	1 761
3	Programy schôdzí 8. volebného obdobia	Kancelária Národnej rady SR	1 697
4	Zoznam vyplatných faktúr PPA	Pôdohospodárska platobná agentúra	1 604
5	Programy schôdzí 7. volebného obdobia	Kancelária Národnej rady SR	1 047
6	Počet prístupov na ÚPVS za deň	Národná agentúra pre sieťové a elektronické služby – NASES	993
7	Počet prístupov na ÚPVS	Národná agentúra pre sieťové a elektronické služby – NASES	854
8	Register Adries - Register vchodov (orientačných čísiel)	Ministerstvo vnútra SR	791
9	Register Adries - Register ulíc	Ministerstvo vnútra SR	757
10	Register Adries - Register obcí	Ministerstvo vnútra SR	745
11	Register adries - Zmenové dávky	Ministerstvo vnútra SR	649
12	Mesačná informácia o epidemiologickej situácii v okrese Trebišov	Regionálny úrad verejného zdravotníctva Trebišov	624
13	Zoznam schválených ŽoNFP OPII	Ministerstvo dopravy a výstavby SR	619
14	Sporové konania	Úrad priemyselného vlastníctva Slovenskej republiky	569
15	Počet prístupov na www.antimon.gov.sk	Protimonopolný úrad SR, Bratislava	534
16	Register Adries - Register budov (súpisných čísiel)	Ministerstvo vnútra SR	527
17	Zoznam faktúr PMÚ SR	Protimonopolný úrad SR, Bratislava	520
18	Vývoj certifikovaného čerpania v rámci OPII	Ministerstvo dopravy a výstavby SR	516
19	Vyhodnotenia analýz pôdnych vzoriek pre Agrochemické skúšanie pôd	Ústredný kontrolný a skúšobný ústav poľnohospodársky v Bratislave	483
20	Zoznam objednávok PMÚ SR	Protimonopolný úrad SR, Bratislava	467

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 a Úrad priemyselného vlastníctva SR. V porovnaní s minulým rokom zo zoznamu vypadlo Ministerstvo zdravotníctva SR.

Oproti minulému roku najviac vzrástla popularita top datasetov Úradu vlády SR (posun o 6 priečok) a Ministerstva obrany SR (posun o 4 priečky).

Tabuľka zachytáva prvé 4 mesiace roku 2023. OVM, ktoré 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 viac zaujímať alebo ako viac spropagovať svoje otvorené dáta.

Tabuľka 4: Najpopulárnejšie datasety OVM

P.	Vlastník	Najpopulárnejší dataset	Počet zobrazení
----	----------	-------------------------	-----------------

Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2023

1	Ministerstvo vnútra SR	Register adries - Zmenové dávky registra adries	20565
2	Úrad priemyselného vlastníctva Slovenskej republiky	Vestník Úradu priemyselného vlastníctva SR	1761
3	Kancelária Národnej rady SR	Programy schôdzí 8. volebného obdobia	1697
4	Ministerstvo pôdohospodárstva a rozvoja vidieka SR	Zoznam vyplatených faktúr PPA	1604
5	Ministerstvo investícií, regionálneho rozvoja a informatizácie Slovenskej republiky	Počet prístupov na ÚPVS za deň	993
6	Ministerstvo dopravy a výstavby SR	Zoznam schválených ŽoNFP OPII	619
7	Protimonopolný úrad SR	Počet prístupov na www.antimon.gov.sk	534
8	Úrad pre verejné obstarávanie	Vestník verejného obstarávania február 2023	344
9	Ministerstvo zdravotníctva SR	Zdravotnícka ročenka Slovenskej republiky	244
10	Ministerstvo obrany SR	Faktúry MO SR	194
11	Štatistický úrad SR	Priemerné ceny pohonných látok v SR (mesačné) [sp0202ms]	131
12	Ministerstvo školstva, vedy, výskumu a športu Slovenskej republiky	Centrálny register študentov	90
13	Ministerstvo spravodlivosti SR	Zoznam pojednávani na okresných a krajských súdoch v civilnej agende	86
14	Úrad vlády SR	Rozpočtová rezerva predsedu vlády SR	83
15	Ministerstvo životného prostredia SR	CITES - výročné správy	82
16	Všeobecná zdravotná poisťovňa	VSŽP - zoznamy dlžníkov	81
17	Úrad geodézie, kartografie a katastra SR	Zoznam Datasetov ÚGKK SR	80
18	Ministerstvo financií SR	041 Dotácie poskytnuté z kapitoly VPS	56
19	Ministerstvo kultúry Slovenskej republiky	Register nehnuteľných národných kultúrnych pamiatok - Bratislavský samosprávny kraj	55
20	Ministerstvo práce, sociálnych vecí a rodiny SR	Centrálny register poskytovateľov sociálnych služieb	54
21	Úrad pre normalizáciu, metrológiu a skúšobníctvo Slovenskej republiky	Zoznam objednávok ÚNMS SR	50
22	Úrad jadrového dozoru Slovenskej republiky	Zoznam zahraničných cestovných nákladov (2023-01-13)	28
23	Ministerstvo hospodárstva SR	Zoznam MH SR využívaného otvoreného softvéru	21
24	Správa štátnych hmotných rezerv Slovenskej republiky	Zoznam datasetov SŠHR SR	19
25	Najvyšší kontrolný úrad SR	NKÚ SR Úradné hodiny (28.03.2022)	16
26	Ministerstvo zahraničných vecí a európskych záležitostí SR	Systém pomoci pri cestách do zahraničia	14
27	Národný bezpečnostný úrad	Zoznam datasetov NBÚ	14
28	Generálna prokuratúra SR	Zoznam odsúdených právnických osôb	9

Zdroj: NASES

4.5 Počet aplikácií

Portál umožňuje registráciu aplikácií²⁸ 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

²⁸ <https://data.gov.sk/apps>

zaregistrovaných aplikácií. V minulom roku bolo aplikácií 11, **v tomto roku sa ich počet nezmenil**. Aplikácií nad otvorenými dátami existuje nepochybne viac, zoznamy je možné nájsť na stránkach neziskových organizácií²⁹. Taktiež samosprávy vytvárajú vlastné aplikácie³⁰.

4.6 Zhrnutie

Z hľadiska návštevnosti v posledných 2 rokoch portál zaznamenal výrazné zvýšenie záujmu. Významne sa zvýšil počet používateľov aj zobrazení portálu. Počet zaregistrovaných aplikácií sa za posledný rok nezmenil, ostal na čísle 11.

Ako najpopulárnejšie datasety za prvé 4 mesiace roku 2023 figurujú Register adries, Vestník Úradu priemyselného vlastníctva SR a Programy schôdzí 8. volebného obdobia (NR SR).

²⁹ <https://utopia.sk/wiki/display/opendata/SK+a+CZ+opendata+aplikacie+a+projekty>,
<https://www.alvaria.sk/open-data-aplikacie/>

³⁰ <https://geopresovregion.sk/home/>, <https://opendata.kosice.sk/pages/vizualizacie>

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 zverejňovať³¹. 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.

Publikačné minimum je vytvorené, zverejnené³². 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. Viaceré datasety budú zisťované automatizovane, z registrov a informačných systémov štátu, čo hodnotíme pozitívne.

Úroveň zverejňovania datasetov PM je rôzna, napríklad pri úradných hodinách je vyššia, kým napríklad pri zahraničných pracovných cestách je nízka – dáta zverejňuje len 5 OVM. **Priemerne sa však pohybuje okolo hranice 8 OVM, čo je približne 42 % OVM. Oproti minulému roku sa zvýšil počet OVM, ktoré zverejňujú publikačné minimum.** Veľkou otázkou, ktorou by sa malo zaoberať MIRRI SR, zostáva kvalita týchto datasetov, ich použiteľnosť a aké aplikácie alebo vizualizácie či využitia by mohli resp. budú nad nimi vznikať.

Tabuľka 5: Počet OVM publikujúcich datasety publikačného minima štátnej správy

Dataset	Počet OVM
Úradné hodiny	11
Kontaktné miesta	11
Dotácie (schémy, obmedzenia, výzvy, žiadosti)	10
Organizačná štruktúra	10
Objednávky	9
Faktúry	8
Podriadené organizácie	8

³¹ OVM môže a má zverejňovať aj datasety mimo Publikačného minima, minimum slúži ako základná báza

³² <https://metais.vicemier.gov.sk/publicspace?pageId=67152325>

Strategické dokumenty	7
Prijaté žiadosti podľa Infozákona	7
Pracovné skupiny	6
Najžiadanejšie otvorené údaje OVM	6
Zahraničné služobné cesty	5

Zdroj: Dátová kancelária MIRRI

5.2 Datasety s vysokou pridanou hodnotou

Smernica EP č. 2019/1024³³ 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.

MIRRI funguje v tomto projekte ako koordinačný a metodický orgán. Pre každý typ HVD datasetu bola vypracovaná tzv. **kartu datasetu**, ktorá slúži ako pomôcka pre poskytovateľov datasetu ale aj pre účely vypracovania správy pre Európsku Komisiu o stave dostupnosti jednotlivých datasetov a podmienkach ich poskytovania. Povinnosť poskytnutia správy o opatreniach bude vyžadovaná do 2 rokov od účinnosti Nariadenia Komisie.

HVD karta obsahuje atribúty, ktoré sa týkajú napr. technických možností používania API rozhrania, zoznamu jednotlivých kľúčových atribútoch, licenčnú politiku, atď. Prebehlo aj prvotné otestovanie kariet Úradom geodézie, kartografie a katastra s jednotlivými gestormi datasetov. MIRRI úzko spolupracuje s Ministerstvom životného prostredia SR ako gestorom priestorových údajov a štandardu pre európsku infraštruktúru priestorových informácií (INSPIRE).

Poskytovanie metadát o HVD datasetov bude riešené prostredníctvom centrálného portálu data.gov.sk, z ktorého budú harvestované metadáta o HVD datasetoch európskym portálom otvorených údajov. Informácie o HVD datasetoch sú zverejňované na metodickom portáli otvorených údajov³⁴. MIRRI tiež aktívne spolupracuje s ostatnými členskými krajinami EU pri implementácii tohto nariadenia.

Pri tvorbe nového portálu sa počítalo aj s príznakom „HVD“ pomocou ktorého bude možné tieto datasety medzi ostatnými filtrovať. Taktiež Metodika hodnotenia dopadu otvorených dát na vybrané oblasti obsahuje indikátory a metriky týkajúce sa využívania datasetov HVD.

³³https://datalab.digital/wp-content/uploads/Smernica2019_1024OPENDATA.pdf

³⁴ <https://wiki.vicpremier.gov.sk/display/opendata/7.+HVD+-+Datasety+s+vysokou+hodnotou>

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.³⁵ 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 %. Od minulého roku MIRRI neaktualizovalo toto vyhodnotenie.

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.

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 2015³⁶, napriek tomu tieto dáta neboli doteraz zverejnené vo formáte otvorených dát. Ministerstvá a ostatné ÚOŠS majú stále k dispozícii výsledky prieskumov dopytu po najžiadanejších datasetoch, na základe ktorých môžu prioritizovať zverejňovanie datasetov v ich pôsobnosti.³⁷

Akčný plán Iniciatívy pre otvorené vládnutie na roky 2022-2024 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 31.10.2023.

5.4 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

³⁵<https://datalab.digital/prieskum-najziadanejsich-datasetov/>

³⁶

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

³⁷ <https://datalab.digital/prieskum-najziadanejsich-datasetov/>

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 vyhľadávanie - sémantické.

Vytvoria sa nové služby - služba na publikáciu meta údajov do stránky, textové sémantické vyhľadávanie, 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 nahratie zdrojových kódov.

Projekt sa skladá z viacerých komponentov, pričom k času písania tejto analýzy boli hotové komponenty NKOD a Komerčná časť vládneho cloudu. Webový portál pre NKOD (frontend) a Gov GIT bol v štádiu riešenia.



Obrázok 3

1. Aktuálny stav OD2.0

Komponenty OD2.0

VO3	Webový portál pre NKOD autentifikácia, úložisko pre súbory, OpenAPI NKOD, služba pre ukladanie súboru, frontend ID-SK, štatistiky (základné, kvalita metadát, publikačné minimum), OpenData komunita: sledovanie datasetov, komentovanie, podnety, registrácie aplikácií		VO4	Nový
VO1	Národný katalóg otvorených dát (NKOD) Dátový model, Databáza, Dátový procesing, harvestovanie externých katalógov a meranie kvality metadát, SPARQL prístupový bod		Gov GIT	PoC Transformačné a validačné API pre datasety
VO2	Komerčná časť vládneho cloudu			Dohodári /interne

Zdroj: PS Otvorené dáta a dátoví kurátori 15.3.2023

SPARQL Endpoint³⁸ bol spustený, prebehlo k nemu aj školenie. Endpoint obsahuje predefinované dotazy týkajúce sa počtu datasetov, formátov, licencií, stiahnuteľnosti datasetov, atď. Bohužiaľ nie všetky dotazy sú momentálne funkčné (stav 28.6.2023). Problém je takisto neexistencia niektorých dôležitých metadát v backende nového portálu ako je napríklad dátum vytvorenia a poslednej modifikácie datasetu.

Aktuálny stav projektu je možné sledovať aj na platforme SD³⁹.

5.5 Školenia

Zoznam zrealizovaných školení Dátovou kanceláriou MIRRI za posledný rok:

- Úvod do otvorených dát,
- Dotazovanie metadát otvorených údajov cez SPARQL Endpoint,
- Školenie administrátorov NKOD,
- Dátová kvalita (séria školení),
- Tvorba publikačného minima štátnej správy.

Všetky školenia sú zverejnené na Youtube Dátovej kancelárii⁴⁰. MIRRI plánuje realizovať v najbližšom čase školenia zamerané na čistenie dát, analýzu a vizualizáciu dát.

³⁸ <https://opendata.mirri.tech/>

³⁹ <https://platforma.slovensko.digital/t/dennik-np-opendata-2-0/7933/129>

⁴⁰ <https://www.youtube.com/@datovakancelaria7258/videos>



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

Akčný plán (AP) na roky 2022-2024 bol schválený uznesením vlády č. 457/2022 z 13.7.2022. AP obsahuje jeden záväzok (3 úlohy) v oblasti otvorených dát. Minulé akčné plány obsahovali viacero (aj niekoľko desiatok) úloh spojených so zverejňovaním otvorených údajov, tento počet sa však časom zmenšoval.

AP obsahuje tieto 3 úlohy vyplývajúce zo záväzku týkajúceho sa otvorených dát:

B.2. 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 30. júna 2023

B.3. 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. októbra 2023

B.4. *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 31. marca 2024

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

V čase písania tejto analýzy bola vo verejnom pripomienkovom konaní Metodika hodnotenia dopadu otvorených dát na vybrané oblasti spoločnosti (úloha B.2. AP).⁴¹

Metodika⁴² sa zameriava najmä **na hodnotenie prístupu, použitia a dopadu otvorených dát štátnej správy**. Metodika vychádza z analýz Európskeho dátového portálu, hodnotenia Open Data Maturity Index a českej metodiky. Participatívny proces pri tvorbe metodiky pozostával z interaktívnych workshopov⁴³ organizovaných od 12/2022 do 6/2023, kde boli preberané témy ako je prístup k dátam, využitie dát, jednotlivé oblasti dopadu otvorených dát (vládna, sociálna, environmentálna, ekonomická). Workshopov sa zúčastnili relevantní hostia (firmy, neziskové organizácie, experti z ČR). Podnety účastníkov workshopov (najmä dátových kurátorov) boli zaznamenávané vo forme Jamboardu. MIRRI plánuje v 2. polroku 2023 spustiť zber údajov zo

⁴¹

https://www.minv.sk/?ros_vsetky-spravy&sprava=na-pripomienkovanie-metodika-hodnotenia-dopadu-otvorených-dat

⁴² <https://wiki.vicemier.gov.sk/pages/viewpage.action?pageId=101829221>

⁴³ <https://wiki.vicemier.gov.sk/pages/viewpage.action?pageId=101829238>

zdrojov, ktoré definuje metodika (napr. prieskumy dátových kurátorov, používateľov portálu, dopadu). MIRRI momentálne postupuje v súlade s časovým plánom pri plnení úloh AP.

Nultý návrh **Akčného plánu Iniciatívy pre otvorené vládnutie na roky 2024 – 2026** bol v čase písania tejto analýzy sprístupnený na verejné pripomienkovanie. Návrh obsahuje 1 záväzok týkajúci sa otvorených dát - Zvyšovanie dopadu otvorených údajov podporou aktivít k ich využívaniu a opätovnému používaniu relevantnými aktérmi. Cieľom záväzku je podporiť využívanie a znovupoužívanie otvorených údajov štátnej (a verejnej správy), ako aj zvýšiť dopyt po zverejnení údajov, ktoré zatiaľ zverejnené nie sú, a podporiť vzťah medzi (potenciálnymi) užívateľmi otvorených dát a ich majiteľmi. Obsah záväzku tvoria dve komplementárne aktivity - **pravidelné meranie dopadu otvorených údajov na vybrané aspekty spoločnosti** (politická, sociálna, environmentálna, ekonomická, príp. iné) podľa pripravenej metodiky a aktivity (napr. hackatony, ideathony, workshopy) zamerané na **budovanie partnerstva medzi majiteľmi otvorených dát a ich užívateľmi smerom k lepšiemu využívaniu už zverejnených údajov**, ale aj k identifikácii doposiaľ nezverejnených údajov.

Pozn.: Autor Metodiky hodnotenia dopadu otvorených dát na vybrané oblasti spoločnosti je aj autorom tejto analýzy.

5.7 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 Sieti⁴⁴**, aby spolu odhaľovali potenciál dát, identifikovali spoločenské výzvy a riešenia, premenili idey na konkrétne výsledky a spolu iniciovali pozitívnu spoločenskú zmenu. V rámci siete aj verejne na webe a sociálnych sieťach je pravidelne distribuovaný newsletter s novinkami v oblasti otvorených dát.

V mesiacoch december 2022 až marec 2023 realizovalo OZ projekt **Spoznaj otvorené dáta** určený pre pedagógov a študentov na 5 stredných školách.⁴⁵ Do projektu bolo celkovo zapojených viac ako 200 študentov a 7 pedagógov, zrealizovalo sa 12 workshopov. Workshopy realizované fyzicky boli zamerané na vzdelávanie študentov v pochopení prínosu otvorených dát a otvoreného vládnutiu a vizualizáciu otvorených dát v nástroji Microsoft PowerBI. V rámci projektu bol vytvorený bezplatný e-learning⁴⁶ a videoprednášky.

Z projektu vyplynulo, že vzdelávanie v oblasti vizualizácie a spracovania dát je na slovenských stredných školách žiadané (najmä gymnáziá a technicky zamerané školy),

⁴⁴ <https://www.alvaria.sk/open-data-siet/>

⁴⁵ <https://www.alvaria.sk/realizovali-sme-projekt-spoznaj-otvorene-data-urceny-pre-stredne-skoly/>

⁴⁶ <https://www.alvaria.sk/kurz/spoznaj-otvorene-data-kurz-pre-stredne-skoly/>

študenti dokázali pochopiť prínos otvorených dát a vedeli s nimi pracovať. 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.

Okrem OZ Alvaria sa z občianskej spoločnosti témou otvorených údajov zaoberajú aj neziskové organizácie Open Data⁴⁷ a Slovenská spoločnosť pre otvorené informačné technológie (SOIT)⁴⁸. 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.

5.8 Opencode.gov.sk – metodický portál pre zverejňovanie zdrojových kódov ISVS

Na riadiacom výbore sa schválil OPZ GOVGit a katalóg požiadaviek. Postupuje sa podľa vzoru Spojeného kráľovstva. V blízkej dobe sa očakáva vyhlásenie verejného obstarávania. Malo by byť nasadené cloudové riešenie GitHub alebo GitLab alebo obdobné riešenie. Tvorí sa **metodika zverejňovania a testovania bezpečnosti zdrojových kódov**. Metodika bude zverejnená na pripomienkovanie. Do legislatívy sa plánuje zaviesť povinnosť pre dodávateľov OVM otestovať svoj zdrojový kód pomocou metodiky a zverejniť ho.

5.9 Zhrnutie

Dôležitý záväzok AP Iniciatívy pre otvorené vládnutie **Mapovanie vplyvu otvorených dát na vybrané aspekty spoločnosti** sa nachádza v schválenom Akčnom pláne Iniciatívy pre otvorené vládnutie na roky 2022-2024. V súvislosti s týmto záväzkom bolo zorganizovaných 6 workshopov v období november 2022 – jún 2024. Každý workshop bol zameraný na špecifickú oblasť dopadu. V čase písania tejto analýzy bola vo verejnom pripomienkovom konaní **Metodika hodnotenia dopadu otvorených dát na vybrané oblasti spoločnosti**.

Mierne sa zlepšilo zverejňovania datasetov publikačného minima štátnej správy – priemerne ich zverejňuje 8 OVM z 19.

Zverejňovanie datasetov a API je naďalej veľkým problémom, **zverejnených bolo len 22,8 % dopytovaných dát** (dáta za minulý rok, nie sú dostupné aktualizované dáta).

Projekt Otvorené dáta 2.0 je v štádiu realizácie. Bol vybudovaný backend a SPARQL Endpoint, pracuje sa na webovom portáli (frontende), ktorý by mal byť spustený ku koncu roka 2023.

⁴⁷ <https://opendata.sk/>

⁴⁸ <https://www.soit.sk/>

MIRRI sa tiež venovalo **organizovaniu školení**, boli zorganizované 4 školenia a 1 séria školení. V oblasti zverejňovania datasetov s vysokou hodnotou bola pre každý typ HVD datasetu vypracovaná tzv. **kartu datasetu**, ktorá slúži ako pomôcka pre poskytovateľov datasetu ale aj pre účely vypracovania správy pre Európsku Komisiu o stave dostupnosti jednotlivých datasetov a podmienkach ich poskytovania. Prebehlo aj **prvotné otestovanie** kariet Úradom geodézie, kartografie a katastra s jednotlivými gestormi datasetov. Informácie o HVD datasetoch sú zverejňované na metodickom portáli otvorených údajov

Bol realizovaný projekt **OZ Alvaria Spoznaj otvorené dáta – určený pre stredné školy**, projekt bol zameraný na pochopenie významu a fungovania otvorených dát, základné možnosti vizualizácie dát bez znalosti programovania. V rámci projektu bol vytvorený bezplatný e-learning pre stredné školy.

6 Záver

Posledný rok priniesol viacero pozitívnych zmien v oblasti otvorených dát. Výrazne pokročila tvorba nového centrálného portálu otvorených dát (projekt OD 2.0). Bol prijatý štandard DCAT-AP-SK-2.0⁴⁹, vybudovaný backend a SPARQL Endpoint nového portálu, čaká sa na spustenie frontendu portálu ku koncu roka 2023. Nový portál by mal priniesť poriadok v katalogizácii otvorených dát na centrálnom portáli a mal by umožniť harvestovanie z lokálnych portálov, vrátane samosprávnych. MIRRI sa začalo výraznejšie venovať dopadu otvorených dát, datasetom s vysokou hodnotou (HVD) a tiež školeniam určeným najmä pre dátových kurátorov.

Počet datasetov sa mierne zvýšil, výrazne sa zvýšila návštevnosť a počet používateľov portálu. Nepodarilo sa však zverejniť významnejší počet žiadaných datasetov a API ani významne zvýšiť využívanie otvorených dát pri životaschopných projektoch verejnej správy, neziskových organizácií či firiem. Podpora otvorených dát v strategických dokumentoch štátu ostáva na rovnakej úrovni ako tomu bolo minulý rok. Rok 2023 je tiež rokom ukončovania projektov z eurofondov, týka sa to projektu nového portálu ako aj projektov na manažment údajov inštitúcií verejného sektora. Pre budúci rozvoj otvorených dát bude kľúčové okrem úspešného dokončenia nového portálu aj aké nástroje participácie a spolupráce bude nový portál ponúkať a či sa na ňom podarí harvestovať väčšiu časť datasetov lokálnych portálov vrátane samosprávnych.

Štátna správa je pomalá pri inováciách a odstraňovaní nedostatkov v doterajších systémoch. Táto analýza sa vypracováva už niekoľko rokov, každý rok sú identifikované nedostatky a odporúčania na zlepšenie. Ústredné orgány štátnej správy nie sú schopné viacero rokov za sebou odstrániť nedostatky už zverejnených a novo zverejňovaných datasetov, ktoré sú v analýze vymenované. Pozitívne je, že sa nasadzuje nový portál otvorených dát, ktorý viaceré nedostatky odstráni. Nový portál vzniká na základe Akčného plánu pre zlepšenie postavenia Slovenska v indexe DESI do roku 2025 (z roku 2021). Od roku 2020 vieme o výraznom prepade v hodnotení Slovenska v rebríčku Open Data Maturity (z dát za rok 2019), pričom očakávame zlepšenie až v roku 2024 (z dát za rok 2023), najneskôr v roku 2025 (z dát za rok 2024). To je horizont 5 rokov. Ak by sa priebežne odstraňovali jednoduchšie nedostatky, zlepšenie by mohlo byť viditeľné v horizonte do 2 rokov.

⁴⁹ <https://datova-kancelaria.github.io/dcat-ap-sk-2.0/>

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 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.
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 zobrazíť zoznam ústredných orgánov štátnej správy⁵⁰ a k nim priradené podriadené organizácie⁵¹. 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.
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 navedol povinné osoby vybrať si zo zoznamu vhodných licencií. Ku zverejneným datasetom bez licencie treba doplniť licenciu. Ponuka

⁵⁰https://www.slovensko.sk/sk/agendy/agenda/_ministerstva-a-ostatne-ustredne1

⁵¹<https://www.slovensko.sk/sk/institucie>



Analýza zverejňovania datasetov ústredných orgánov štátnej správy na portáli otvorených dát za rok 2023

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 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 hviezdčkami⁵². Kvalita poskytovaného datasetu je v tomto význame zadefinovaná vo vyhláške ÚPVII č. 78/2020 Z. z. (Vyhláška o štandardoch pre informačné technológie verejnej správy)⁵³, § 38 až § 40. Na zverejňovanie otvorených dát sa vzťahuje tiež Národná koncepcia informatizácie verejnej správy⁵⁴, Stratégia a akčný plán prístupnosti a používania otvorených údajov verejnej správy⁵⁵. 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 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.
10. Verejnoscť 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 podaní⁵⁶ verejnoscťou.

⁵²<https://5stardata.info/cs/>

⁵³<https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2020/78/>

⁵⁴<https://www.mirri.gov.sk/sekcie/informatizacia/narodna-koncepcia-informatizacie-verejnej-spravy-nikvs/index.html>

⁵⁵https://www.slovensko.sk/_img/CMS4/strategia_otvorene_udaje.pdf

⁵⁶<https://data.gov.sk/published-services>

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 portáloch⁵⁷.
14. 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.
15. Publikovať manuál a/alebo videonávod ako napojiť existujúce portály otvorených dát automatizovane na centrálny portál a v prípade záujmu OVM k tomuto usporiadať workshop.
16. 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 OVM dostatočne dlhé prechodné obdobie, v ktorom by nevyhovujúce datasety boli nahradené vyhovujúcimi, aby sa tak žiadne údaje nestratili a neznižoval sa počet datasetov.
17. 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.
18. Umožniť používateľovi tlačidlom a formulárom bez nutnosti elektronického podpisu nahlásiť chybu pri datasete. Oznamy by sa mali odosielať dátovým kurátorom.
19. Umožniť používateľovi podať návrh na zverejnenie údajov na portáli otvorených dát aj formou e-mailu alebo vo formulári bez nutnosti elektronického podpisu.
20. Umožniť používateľom filtrovať datasety aj podľa periódy aktualizácie.

Odporúčania pre dátovú kanceláriu MIRRI

17. Po vzore ČR⁵⁸ monitorovať „časté zlé praktiky“ a zverejňovať krátke návody, popisujúce konkrétny nedostatok a kroky potrebné pre nápravu. Pozn.: MIRRI má vlastné návody⁵⁹, ktoré však nie sú natoľko prehľadné v porovnaní s ČR a neobsahujú komplexný zoznam najčastejších chýb.

⁵⁷ <https://utopia.sk/wiki/display/opendata/SK+a+CZ+opendata+aplikacie+a+projekty>

⁵⁸ <https://opendata.gov.cz/špatná-praxe:start>

⁵⁹ <https://wiki.vicempremier.gov.sk/pages/viewpage.action?pageId=82027683>

18. Zabezpečiť prioritné zverejnenie datasetov a API, ktoré vzišli z prieskumov dopytu verejnosti po najžiadanejších datasetoch a API.
19. Realizovať ďalší prieskum dopytu verejnosti po najžiadanejších datasetoch a API po splnení predchádzajúcej úlohy minimálne na 75 %.
20. Pravidelne vyhodnocovať úroveň zverejnenia najžiadanejších datasetov a API.
21. Pri realizácii prieskumu dopytu verejnosti po najžiadanejších datasetoch umožniť OVM a dátovým kurátorom kontaktovať respondentov, ktorí dobrovoľne na seba uvedú kontakt, v prípade nejasností.
22. Zverejniť návrhy na zverejnenie údajov na portáli otvorených dát v anonymizovanej podobe ako dataset v pravidelnej periodicite najmenej raz ročne.
23. Znižovať podiel datasetov bez uvedenia licencie (hodnota „iná“), datasetov s nedostatočne otvorenou licenciou (napr. CC-SA, CC-NC), okrem odôvodnených prípadov.
24. Vyhodnocovať dopyt verejnosti po datasetoch a API územnej samosprávy.

Odporúčanie pre povinné osoby – poskytovateľov datasetov

18. Pred publikovaním datasetov odporúčame *preštudovanie Metodickéj príručky pre povinné osoby – ako zverejňovať otvorené dáta (2015)*⁶⁰, čím by sa dalo vyhnúť nedostatkom opísaným v tejto analýze.

⁶⁰https://www.slovensko.sk/_img/CMS4/Navody/Metodicka_prirucka_pre_povinne_osoby.pdf