



DATALAB

NEWSLETTER

DECEMBER 2020

ZÁKON O ÚDAJOCH JE TESNE PRED
MEDZIREZORTNÝM PRIPOMIENKOVÝM KONANÍM

DÁTOVÁ KANCELÁRIA SA CHCE SÚSTREDIŤ
AJ NA ROZVOJ CENTRÁLNEHO DÁTOVÉHO MODELU

DÁTOVÁ INTEROPERABILITA
JE PILIEROM EGOVERNMENTU



Operačný program
**Efektívna
verejná správa**



Európska únia
Európsky sociálny fond

Tento projekt je podporený z Európskeho sociálneho fondu



**Milan
Andrejkovič**
Dátová kancelária

Milí čitatelia,
ku koncu sa blíži rok, na aký si takmer nikto z nás nepamätá. Napriek všetkému, čo nás postretlo, sa pracovný život nezastavil, pokračovali sme v rozbehnutých aktivitách a dokonca sme spustili nové projekty. Náš newsletter mal síce v dôsledku pandémie aj zmien na našom úrade (ktorý sa transformoval na ministerstvo) niekoľkomesačnú prestávku, ale dúfame, že jeho obnovenie uvítate a ponúknuté témy vás zaujmú.

Z našej súčasnej činnosti je kľúčové, že čoskoro predložíme do opätovného medzirezortného pripomienkového konania návrh zákona o údajoch, ktorý vytvára legislatívny rámec pre systematické riadenie údajov verejnej správy. Nemenej prioritná je aj kontinuálna práca na riadení dátovej kvality, ktorá sa už prejavuje na „čistení“ viacerých zdrojových registrov. Naším cieľom je pokračovať v nastúpenom

trende, aby sme údaje čo najviac priblížili k cieľovej nulovej chybovosti.

Naša práca prináša konkrétne výsledky aj v čase pandémie, ako sa o tom môžete presvedčiť v článku o prínose Centrálnej správy referenčných údajov počas pandemickej situácie. Vďaka jej službám si napríklad úrady práce, sociálnych vecí a rodiny dokážu bezpapierovo overovať údaje o státisícoch zamestnávateľov a živnostníkov. Stále sa v spolupráci s gestormi venujeme aj vyhlasovaniu referenčných údajov a publikovaniu základných číselníkov. Sfinalizovali sme návrh publikačného minima štátnej správy a popri ňom sa chceme sústrediť aj na rozvoj centrálneho dátového modelu.

Vo všetkých spomenutých oblastiach nás čaká ešte veľa práce, a tak spoločne dúfajme, že rok 2021 bude vo všetkých smeroch pokojnejší a úspešnejší ako tento.

TÉMA

Meranie biznis pravidiel
na Registri právnických
osôb, podnikateľov
a orgánov verejnej moci

3

DÁTOVÁ LEGISLATÍVA

Zákon o údajoch je tesne
pred medzirezortným
pripomienkovým konaním

6

REFERENČNÉ ÚDAJE A ČÍSELNÍKY

Dátová kancelária sa chce
sústrediť aj na rozvoj
centrálneho dátového modelu

10

OTVORENÉ ÚDAJE

Podporujeme zverejňovanie
najžiadanejších datasetov

13

STOP BYROKRACII

Aké sú prínosy IS CSRÚ
počas pandemickej situácie

5

DÁTOVÁ INTEROPERABILITA

Dátová interoperabilita
je pilierom eGovernmentu

8

REFERENČNÉ ÚDAJE

V čom spočíva význam
referencovania údajov

11

PREDSTAVUJEME

Peter Barlík

14

MERANIE BIZNIS PRAVIDIEL NA REGISTRI PRÁVNICKÝCH OSÔB, PODNIKATEĽOV A ORGÁNOV VEREJNEJ MOCI

Štatistický úrad zrealizoval na údajoch registra právnických osôb zatiaľ tri merania – v marci, júli a septembri tohto roku. Na začiatku celého procesu v januári 2020 Štatistický úrad stanovil 67 biznis pravidiel, ktoré sa postupne podľa potreby dopĺňali. Pomocou nich vieme percentuálne stanoviť správnosť, respektíve chybovosť vo vybranej množine údajov.



Andrej Fukas



Radoslav Brída

(Štatistický úrad SR)

Proaktívne aktivity sú dopredu plánované pravidelné merania a opatrenia na zvolených evidenciách a registroch, ktoré nemajú za cieľ len odstrániť prípadné nedostatky, ale cez formalizovanie biznis pravidiel aj bližšie špecifikovať objekty evidencie. Sú teda aj akýmsi prostriedkom na lepšie spoznanie pravidiel a vlastností daného registra. Za príklad dobrej praxe proaktívnych aktivít v oblasti dátovej kvality považujeme prístup Štatistického úradu pri správe Registra právnických osôb.

Register právnických osôb, podnikateľov a orgánov verejnej moci (RPO) je informačný systém verejnej správy zriadený zákonom 272/2015 Z. z. o registri právnických osôb, podnikateľov a orgánov verejnej moci.

Údaje do RPO zasielajú povinné osoby (napr. Ministerstvo spravodlivosti SR, Ministerstvo vnútra SR) zo svojich evidencií, t.j. zdrojových registrov, ktoré spravujú. Aktuálne sa v RPO nachádza viac ako 884-tisíc platných záznamov z 83 zdrojových registrov a 849 obcí (SHR).

Štatistický úrad zrealizoval na údajoch registra zatiaľ tri merania – v marci, júli a septembri tohto roku. Predchádzalo tomu stanovovanie biznis pravidiel, ktoré sú formalizovanými požiadavkami

pre hodnoty údajov v jednotlivých atribútoch v registri, ale zároveň aj merateľnými ukazovateľmi. Pomocou nich vieme percentuálne stanoviť správnosť, respektíve chybovosť vo vybranej množine údajov. Na začiatku celého procesu v januári 2020 Štatistický úrad stanovil 67 biznis pravidiel, ktoré sa postupne podľa potreby dopĺňali (96 biznis pravidiel v septembri). Všetky sú súčasťou publikovaného [centrálneho zoznamu biznis pravidiel](#) dostupného na webe Dátovej kancelárie.

V rámci tohto príspevku by sme chceli spomenúť len tie biznis pravidlá, ktorých meranie prinieslo zistenia výraznejších a zároveň najzávažnejších nedostatkov. Samotné hodnotiace správy si môžete stiahnuť aj z webovej stránky Štatistického úradu (dostupné na webe www.statistics.sk). Všetky hodnotiace správy z merania kvality údajov aj iných registrov sú dostupné na [webe Dátovej kancelárie](#).

Tabuľka č. 1: Počet záznamov nespĺňajúcich biznis pravidlo BP_SU_RPO_002 – IČO právnickej alebo fyzickej osoby je referenčný identifikátor a jeho hodnota je 8-znakový reťazec zložený z číslíc.

Zdrojový register	Počet záznamov nespĺňajúcich pravidlo		
	Marec 2020	Júl 2020	September 2020
Obchodný register	1 596	1 593	1586 !
Register spoločenstiev vlastníkov bytov a nebytových priestorov	1 179	4	0 ✓
Zoznam advokátov	952	941	660 !
Register záujmových združení právnických osôb	144	0	0 ✓
Živnostenský register	42	42	40 !
Register organizácií s medzinárodným prvkom	42	0	0 ✓

Pri vyhodnotení merania biznis pravidiel 001 až 006 sa prejavili dlhodobo známe problémy s duplicitnými alebo nesprávnymi IČO (jedno IČO pre dva rôzne subjekty, dve rôzne IČO pre ten istý subjekt) a chyby s duplicitnými záznamami pre to isté IČO v jednom registri. Podiel nesprávnych záznamov voči celkovému počtu záznamov v RPO nie je veľký (menší ako 0,4%), ale tieto prípady (v marci 2020 v absolútnych číslach viac ako 4000) spôsobujú používateľom ťažkosti pri používaní IČO ako unikátneho identifikátora. Opravy je potrebné vykonávať priamo v zdrojových registroch v spolupráci so Štatistickým úradom. Viaceré zdrojové registre však majú svoje špecifiká, najmä procesného a legislatívneho charakteru.

Positívny efekt riadenia dátovej kvality sa prejavil v Registri spoločenstiev vlastníkov bytov a nebytových priestorov, Registri záujmových združení právnických osôb a Registri organizácií s medzinárodným

Aktivity v oblasti dátovej kvality môžeme rozdeliť na proaktívne a reaktívne. Reaktívne aktivity predstavujú reakcie na odstránenie konkrétnych problémov spôsobených dátovou nekvalitou.

Podnetom býva upozornenie od používateľov elektronických služieb štátu alebo realizácia integrácie rôznych informačných systémov verejnej správy, keď je potrebné porovnať údaje vo viacerých evidenciách. V takomto prípade sa k zisteniu nedostatkov v dátovej kvalite dostaneme „náhodou“.

Tabuľka č. 2: Počet záznamov nespĺňajúcich biznis pravidlo BP_SU_RPO_012 – Subjekt má uvedenú ako právnu formu platnú položku z číselníka právnych foriem a zároveň nemá uvedenú právnu formu „995 – Nešpecifikovaná právna forma“.

Zdrojový register	Počet záznamov nespĺňajúcich pravidlo			!
	Marec 2020	Júl 2020	September 2020	
Obchodný register	10 733	10 636	10 581	!

Tabuľka č. 3: Počet záznamov a podiel záznamov nespĺňajúcich biznis pravidlo BP_SU_RPO_028 – Ak je štatutárny orgán fyzická osoba, tak má uvedené IFO.

	Marec 2020	Júl 2020	September 2020	!
Počet záznamov nespĺňajúcich pravidlo	244 846	226 110	222 282	!
Podiel záznamov nespĺňajúcich pravidlo – z celkového počtu záznamov o štatutároch	26,43%	23,98%	23,13%	!

prvkom v úplnom odstránení nedostatkov a vyčistení aj zdrojového registra aj RPO. Práve spomenuté registre, respektíve subjekty v nich zapísané, boli zároveň aj predmetom riešenia povinnej aktivácie elektronických schránok pre právnické osoby nezapísané v Obchodnom registri. Dátová kancelária koordinovala od 1.6.2020 spoluprácu medzi správcami zdrojových registrov, Štatistickým úradom a Národnou agentúrou pre sieťové a elektronické služby (NASSES) najmä v oblasti čistenia údajov. Vznikli tu jasne definované požiadavky na dátovú kvalitu, najmä zabezpečenie rovnakých údajov v troch rôznych evidenciách a sotožnenie fyzických osôb pre bezproblémové vytvorenie prístupu.

Ďalším závažným problémom, ktorý sa nachádza výlučne v údajoch z Obchodného registra, je neuvedenie právnej formy. Je to jeden z hlavných údajov

o subjekte a tieto chyby by mali byť opravené prioritne. Deje sa to však pomalým tempom, čo vidieť aj v tabuľke č. 2.

Častým problémom, s ktorým sa na Štatistický úrad v poslednej dobe obracia NASSES ako správca elektronických schránok zriadených na Ústrednom portáli verejnej správy (ÚPVS), je situácia, keď štatutár nemá prístup do elektronickej schránky subjektu. Prístup osôb do schránok je založený na identifikácii pomocou rodného čísla alebo IFO (technický identifikátor fyzickej osoby z Registra fyzických osôb) uvedených v RPO. Ak osoba pri danom subjekte v RPO rodné číslo nemá korektne zapísané (biznis pravidlá 015, 027, 029) alebo dokonca pri subjekte nie je uvedená vôbec, tak štatutár nemá prístup do schránky, je z celej situácie zmätený a frustrovaný a má dojem, že elektronizácia štátnej správy nefunguje.



Zdrojové registre by údaje o štatutároch mali referencovať na Register fyzických osôb. Okrem vrátenia IFO tento proces zabezpečí jednoznačnú a správnu identifikáciu osoby. Táto požiadavka nie je splnená pri 23 percentách záznamov (biznis pravidlo 028). V tabuľke č. 3 vidieť pozitívny vývoj pri naplňaní tohto údaja.

Štatistický úrad priebežne žiada povinné osoby ako správcov zdrojových registrov o nápravu ziste-

ných nedostatkov. Niektoré nedostatky sú spôsobené nesprávnymi procesmi na strane zdrojových registrov, často spojenými s nevyhovujúcou legislatívou. Dátová kancelária bude požadovať od správcov zdrojových registrov, aby spustili procesy riadenia dátovej kvality práve v najproblematickejších oblastiach. Súčasťou týchto procesov musí byť aj prebratie zodpovedajúcich biznis pravidiel z RPO, ich meranie, identifikácia príčin nedostatkov a nápravné opatrenia.

STOP BYROKRACII



**Andrej
Hajdúch**

Aké sú prínosy IS CSRÚ počas pandemickej situácie

Vďaka službám Informačného systému Centrálnej správy referenčných údajov si úrady práce, sociálnych vecí a rodiny v októbri a novembri overili automatizovane údaje o viac ako 150-tisíc zamestnávateľoch a živnostníkoch (SZČO).

Ministerstvo práce, sociálnych vecí a rodiny vypláca prostredníctvom jednotlivých úradov práce, sociálnych vecí

a rodiny v rámci pandemickej situácie tzv. „Prvú pomoc zamestnávateľom, zamestnancom a živnostníkom“. Súčasťou pomoci je napr. preplatenie 80 percent celkovej ceny práce zamestnávateľom v maximálnej výške 1100 eur, ktorá zahŕňa preplatenie nielen hrubej mzdy, ale aj odvodov za zamestnanca. Pomoc je určená zamestnávateľom, SZČO, ktorí museli svoje prevádzky zatvoriť alebo obmedziť svoju činnosť na základe rozhodnutia Úradu verejného zdravotníctva alebo im poklesli tržby o najmenej 20 percent.

Úrady práce, sociálnych vecí a rodiny pomoc vyplácajú na

základe uzatvorených dohôd so žiadateľmi, súvisiacich žiadostí a následných vyplnených výkazov. Žiadosti a výkazy sa podávajú výlučne elektronicky prostredníctvom portálu www.slovensko.sk.

Zamestnávatelia a SZČO, ktorí sú zamestnávateľmi, prikladajú k podávaným žiadostiam o i. vyhlásenia, že majú splnené daňové a odvodové povinnosti na sociálnom a zdravotnom poistení. Pre pandemickej situáciu mali zamestnávatelia možnosť dohodnúť sa s daňovými úradmi a Sociálnou poisťovňou na odklade plnenia povinností prostredníctvom splátok/splátkových kalendárov dlžných súm.

Aby zamestnávatelia a SZČO nemuseli následne dokladovať splnenie daňových a odvodových povinností prostredníctvom predkladania mesačných listinných potvrdení daňových úradov, Sociálnej poisťovne a zdravotných poisťovní, využili úrady práce dostupné služby IS Centrálnej správy referenčných údajov (CSRÚ). Vďaka službám CSRÚ využili úrady práce možnosť overiť splnenie podmienok zamestnávateľov automatizovane za viac ako 150-tisíc subjektov a prostredníctvom jednotného volania služieb IS CSRÚ bolo možné získať údaje o splnení záväzkov súčasne za daňové úrady a aj odvodové po-

vinnosti voči Sociálnej poisťovni a trom zdravotným poisťovňam. Nebolo teda potrebné, aby zamestnávatelia predložili úradom práce viac ako 500-tisíc listinných potvrdení a aby ich úrady práce museli manuálne spracúvať. Keďže pandemickej pomoc bude pokračovať pravdepodobne ešte viacero mesiacov, pozitívne efekty z využitia služieb CSRÚ sa budú nabaľovať, nehovoriac o prínosoch súvisiacich s odbúraním nutnosti osobného kontaktu úradníkov s dotknutými subjektmi poberajúcimi pandemickej pomoc.

Dátová kancelária aktuálne intenzívne pracuje na rozšírení dostupných „služieb“ CSRÚ aj o ďal-

šie dodatočné elektronické výpisy a potvrdenia týkajúce sa dýchacích a nemocenských dávok, ktorých časť bude sprístupnená už od 1.1.2021. Zároveň pripravuje ďalšiu vlnu legislatívnych zmien smerujúcich k odstráneniu byrokracie súvisiacich so súborom takmer 20 okruhov listinných dokladov. Prostredníctvom CSRÚ a jej portálu OverSi budú mať oprávnené orgány verejnej moci možnosť získať údaje napr. vo vzťahu k rodným, sobášnym a úmrtným listom, ako aj informácie, či je subjekt v konkurze, úpadku, likvidácii a reštrukturalizácii, prípadne či neporušil zákaz nelegálneho zamestnávania osôb atď.

DÁTOVÁ LEGISLATÍVA



**Michaela
Galia**

Zákon o údajoch je tesne pred medzirezortným pripomienkovým konaním

V najbližších dňoch bude opätovne do legislatívneho procesu v rámci medzirezortného pripomienkového konania predložený návrh zákona o údajoch.

Predkladaný zákon vytvára legislatívny a strategický rámec pre riadenie údajov verejnej správy s cieľom dynamického zvyšovania kvality údajov evidovaných v jednotlivých registroch a eviden ciách orgánov verejnej moci a tiež zvyšovania interoperability informačných systémov verejnej správy v správe jednotlivých orgánov verejnej moci. Tieto ciele sú v návrhu zákona zadané ako verejný záujem v oblasti údajov, pričom zákon obsahuje vo viacerých ustanoveniach opatrenia na zabezpečenie realizácie takto definovaného verejného záujmu. Ide najmä o zankotvenie silnej pozície dátového



kurátora ako subjektu s právomocami umožňujúcimi účinne pôsobiť na zvyšovanie kvality údajov. Kvalite údajov napomáha aj ich referencovanie a inštitút referenčných údajov, ktorý je aktuálne obsiahnutý v zákone č. 305/2013 Z. z. (zákon o e-Governmente) v znení neskorších predpisov. Fundamentálnym znakom údajov po jeho vyhlásení za referenčný údaj je, že jeho hodnota sa považuje za úplnú a správnu, preto procesu vyhlásenia takéhoto údajov musí nevyhnutne predchádzať zmeranie a „vyčistenie“ jeho kvality. V záujme vytvorenia „veľkého“ a efektívneho zákonníka údajov bola predmetná právna úprava do uvedeného návrhu zákona prevzatá spolu s nadväzujúcou právnou úpravou základných číselníkov zo zákona č. 95/2019 Z. z. o informačných technológiách vo verejnej správe. Návrh zákona zabezpečuje podporu dátovej integrácie na modul procesnej integrácie a integrácie údajov v zmysle zákona o e-Governmente, čím podporuje reálnu realizáciu konceptu „stop byrokracií“ – znižovania administratívnej záťaže občana, ale záro-

veň aj znižovania pracovnej záťaže zamestnancov verejnej správy, keď predpokladá sprostredkovanie údajov „na pozadí“ informačných systémov automatizovaným spôsobom, a to pri zachovaní práv fyzických a právnických osôb na adekvátnu ochranu ich osobných údajov.

Koncept analytických jednotiek obsiahnutý v návrhu zákona umožňuje vytvoriť zo Slovenskej republiky moderný digitálny štát, ktorý sa pri tvorbe stratégií a plánovaní ďalšieho rozvoja môže opierať o kvalitné analýzy realizované na základe legálne získaných a kvalitných reálnych dát, nielen o tzv. kvalifikovaný odhad, a využívať ich.

Návrh zákona obsahuje koncept Mojich údajov ako nový právny inštitút v slovenskom právnom poriadku, ktorý má medzinárodnú oporu a nadväzuje na právnú úpravu ochrany osobných údajov



obsiahnutú najmä v nariadení Európskeho parlamentu a Rady (EÚ) 2016/679 o ochrane fyzických osôb pri spracúvaní osobných údajov a o voľnom pohybe takýchto údajov. Fyzické a právnické osoby majú právo byť promptne informované o svojich údajoch obsiahnutých v jednotlivých registroch/evidenciách orgánov verejnej moci, o zmenách v takýchto údajoch i o skutočnosti, kto, kedy a na základe akého právneho dôvodu k ich údajom pristupoval. Uplatňovanie uvedenej triedy práv fyzických a právnických osôb sa má v zmysle návrhu zákona diať automatizovaným spôsobom pomocou informačných systémov a moderných digitálnych technológií. Právna úprava obsiahnutá v návrhu zákona nadväzuje na medzinárodnoprávne záväzky Slovenskej republiky, napr. v oblasti

naplňania konceptu „otvorených údajov“ – ide o všestrannú podporu sprístupňovania čo najväčšieho množstva údajov čo najširšiemu okruhu subjektov v strojovo spracovateľnej forme bez akýchkoľvek obmedzení, pokiaľ takýmto sprístupneniu nebráni v konkrétnych prípadoch špecifická ochrana poskytovaná určitým kategóriám údajov, napr. osobné údaje, citlivé údaje, bankové tajomstvo atď.; v oblasti naplňania indikovaných ukazovateľov európskeho hodnotenia pod názvom eGov Benchmark, ktoré mapuje približovanie služieb verejnej správy fyzickým i právnickým osobám vďaka elektronizácii a digitalizácii, a to nielen prostredníctvom webových služieb, ale aj v podobe mobilných aplikácií.

Prijatím zákona o údajoch a jeho včasným uvedením do aplikáčnej praxe sa vytvorí právny nástroj, ktorým bude možné promptne reagovať na prudký európsky i medzinárodný rozvoj v oblasti údajov a digitalizácie, napr. v kontexte dokumentu „Európska dátová stratégia“.

DÁTOVÁ INTEROPERABILITA



**Miroslav
Líška**

Dátová interoperabilita je pilierom eGovernmentu

Jedným z kľúčových problémov pri práci s údajmi ISVS je ich obrovské množstvo, rôznorodosť reprezentácie a uloženie v rôznych systémoch. Na realizovanie nejakej komplexnej dátovej analýzy, resp. na realizáciu verejného registra či elektronickej služby, je nutné tieto údaje veľkým úsilím najskôr integrovať, až potom je možné takého údaje využiť. Bohužiaľ, táto aktivita vyčerpá podstatné zdroje nejakého projektu. Čo ak by boli, ale údaje okamžite k dispozícii, nech sú už uložené kdekoľvek, alebo akokoľvek? To je akiste mimoriadne žiadúce. Ale je to vôbec možné?

V súčasnosti existuje obrovské množstvo údajov verejnej správy. Bohužiaľ, podstatnú väčšinu týchto údajov nevieme efektívne využiť, pretože údaje sú v rôznej kvalite roztrúsené v rôznych systémoch, súboroch, databázach, resp.

dokumentoch. Jedným z problémov je ich nejednotná reprezentácia (chýbajúci zdieľaný model) a nejednotná identifikácia (chýbajúce globálne, t.j. zdieľané identifikátory naprieč všetkými ISVS). Na základe tohto nie je možné pracovať s údajmi ako s jedným celkom, to znamená údaje nie sú prepojené. Čokoľvek je potrebné s údajmi spraviť, najskôr je potrebné takéto údaje veľkým úsilím aplikačne integrovať. Jedným z rizík takéhoto prístupu je, že kým sa takýto integračný projekt dokončí, samotný dôvod integrácie údajov sa môže zmeniť, či úplne zaniknúť.

Oveľa lepší prístup je, že nezáleží na umiestnení údajov, resp. uložení v rôznych systémoch, či v rôznej forme. Ideálne je, ak sú údaje prepojené svojou podstatou, už len vytvorením. To znamená, že používajú jednak jednotné globálne identifikátory, a súčasne jednotný model údajov. V takomto prípade, ak sa napr. chcem dozvedieť informácie o nejakom pozemku, tak o ňom viem zistiť „na jeden klik“ všetky dostupné informácie zo všetkých možných



ISVS (napr. katastrálne informácie, územno-plánovacie informácie, geografické, geologické, sociálne, kriminalitu, nehodovosť, evidovane historické udalosti a podobne). V takomto prípade nepotrebujeme míňať úsilie na integračný projekt aby sme údaje spojili a mohli niečo vytvoriť. Naopak ušetrené úsilie vieme použiť na vytvorenie pridanej hodnoty zamýšľaného systému. Inými slovami, vývoj rôznych aplikácií v takom prípade je neporovnateľne efektívnejší.

Ale je možné niečo také začať používať v prostredí veľkého množstva verejných heterogénnych systémov s veľkým množstvom

údajov? Odpoveď je áno. Odpoveďou je tzv. technológia LinkedData (prepojené údaje), ktorá zabezpečuje tzv. Sémantickú interoperabilitu. Tá je definovaná jednak zdieľaním znalostných modelov, to znamená množinou ontológií (modelov) údajov danej domény

a používaním globálnych identifikátorov. Technológia LinkedData je tvorená množinou otvorených štandardov postavených na jazyku RDF konzorcia W3C, to znamená hlavným štandardizátorom webu ako takého. V prípade keď je interoperabilita implementovaná vo

verejných údajoch, hovoríme o tzv. Linked Government Data.

Samozrejme, implementácia dátovej interoperability vo verejných údajoch nie je Slovenský nápad. Sémantický web, resp. LinkedData je používané vo verejnom sektore od počiatku prepojených

údajov. V rámci EÚ je interoperabilita samostatným programom Európskej komisie ISA² (Interoperability solutions for public administrations, businesses and citizens.). Je veľmi dôležité povedať, že používanie mnohých EÚ štandardov nie je dobrá vec len z dôvodu

medzinárodnej interoperability. Oveľa väčší prínos je postupné vyriešenie vnútro-štátnej správy údajov.

Na Slovensku začala byť dátová interoperabilita implementovaná od roku 2016: 1) keď bol zverejnený prvý Centrálny model údajov verejnej správy založený na ontológiách (prepojené údaje), 2) keď začalo vyhlasovanie referenčných údajov vydávaním URI identifikátorov pre údaje ISVS v Centrálnom Meta-Infornačnom Systéme. Odtedy sa štandardy výrazne rozvinuli, a na pozvanie Európskej komisie boli prezentované na konferencii SEMIC 2018.

Avšak ani tento dobrý základ nemusí okamžite zaručiť úspech. Slovensko čaká obrovská výzva správne implementovať interoperabilitu spolu s princípmi Jedenkrát a dosť, Moje údaje či zvýšiť kvalitu otvorených údajov. O ďalších podrobnostiach, resp. aktuálnom prograse Vás budeme informovať v budúcnosti sériou článkov z načrtnutej oblasti dátovej interoperability.

REFERENČNÉ ÚDAJE A ČÍSELNÍKY



**Viktória
Šunderlíková**

Dátová kancelária sa chce sústrediť aj na rozvoj centrálného dátového modelu

Dátová kancelária v spolupráci s gestormi základných číselníkov v tomto roku publikovala v MetaIS niekoľko základných číselníkov. Okrem vyhlasovania základných aj nezákladných číselníkov sa chce sústrediť aj na rozvoj centrálného dátového modelu.

Dátová kancelária v spolupráci s gestormi základných číselníkov v tomto roku publikovala v MetaIS niekoľko základných číselníkov. Sú to číselníky Úradu geodézie, kartografie a katastra vzťahujúce sa k nehnuteľnostiam a tiež číselníky Katalógov odpadov, ktorých gestorom je ministerstvo životného prostredia. V súčasnosti pripravujeme ďalšie základné číselníky týkajúce sa geografického informačného systému a katastrálnych území.

Okrem vyhlasovania základných aj nezákladných číselníkov sa chce Dátová kancelária sústrediť aj na rozvoj centrálného dátového modelu, do ktorého základné číselníky neodmysliteľne patria.

Samotný centrálny dátový model je definovaný vo vyhláške o štandardoch č. 78/2020 Z.z., vo vzťahu k číselníkom centrálny dátový model existuje hlavne v MetaIS, kde sa registrujú jednotné referencovateľné identifikátory základných číselníkov a ich položiek. Na základe štruktúry jednotného referencovateľného identifikátora vieme následne umiestniť dátový prvok do schémy centrálného dátového modelu.

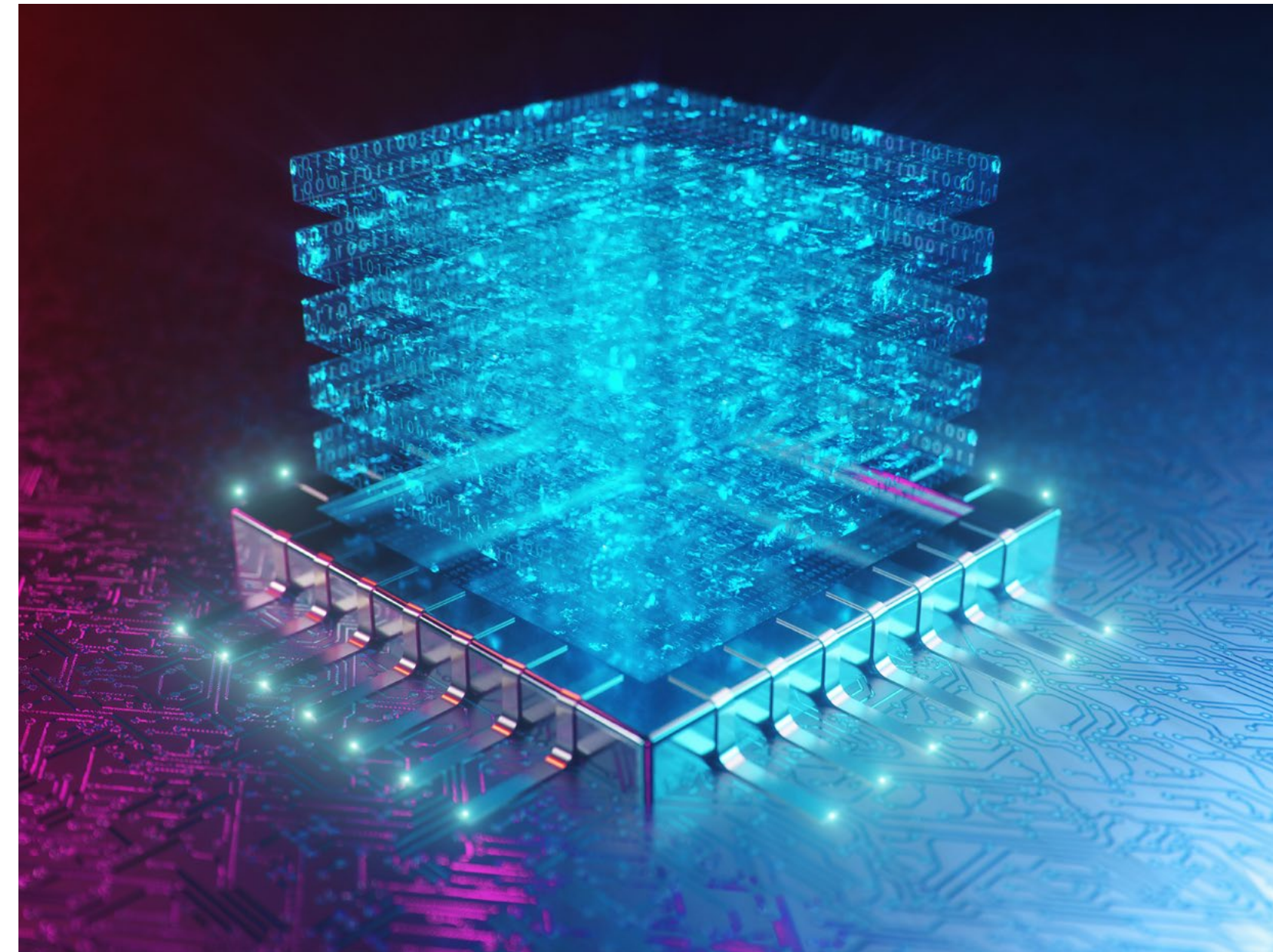
Samotné číselníky sa registrujú ako dataset a majú jednotnú štruktúru referencovateľného identifikátora v tvare <https://data.gov.sk/set/codelist/CL0XXXX>. Jednotné referencovateľné identifikátory číselníkov sa odlišujú iba kódom číselníka. Napr. číselník životné situácie bude mať jednotný referencovateľný identifikátor v tvare <https://data.gov.sk/set/codelist/CL000646>.

Číselník okruhy životných situácií bude mať jednotný referencovateľný identifikátor v tvare <https://data.gov.sk/set/codelist/CL000647>.

Navonok nám táto štruktúra URI veľa nenapovie o zaradení datasetu (číselník) do centrálného dátového modelu. Nehovorí nám nič o tom, do ktorej ontológie alebo k akému dátovému prvku daný číselník patrí.

Registrátor datasetu číselníka však pri registrácii určuje ontológiu a dátový prvok, ku ktorému ho zaraďuje. Napr. číselník životná situácia patrí do ontológie egovernment entít a k dátovému prvku životná situácia.

Bežný používateľ vidí príslušnosť k dátovému prvku podľa URI položiek číselníka, ktoré však už majú úplne štruktúru. Napríklad položky číselníka životné situácie budú mať URI v tvare https://data.gov.sk/id/egov/life-situation/{kod_polozky}. Z danej URI vidíme, že položky číselníka životná situácia patria



do ontológie egovernment entity a k dátovému prvku životná situácia. Rovnako je to aj s položkami číselníka okruhy životnej situácie, ktorá má tvar https://data.gov.sk/id/life-situation/scope/{kod_polozky}. Položky tohto číselníka patria do

rovnakej ontológie, ale k inému dátovému prvku.

Základné číselníky sú štandardizované prvky, preto je dôležité ich čo najširšie použitie na dosiahnutie čo najväčšej miery štandardizácie a tým zároveň aj interoperability v informačných

systemoch štátnej správy. Zároveň sú súčasťou aj centrálného dátového modelu, čo dáva potenciál aj pre zmapovanie ich využitia v jednotlivých informačných systémoch štátnej správy prostredníctvom jednotlivých dátových modelov.

REFERENČNÉ ÚDAJE



**Jozef
Antoš**

V čom spočíva význam referencovania údajov

Úspešným vyhlásením vybraných údajov za referenčné dosiahneme len polovicu cieľa. Konečný úspech totiž tkvie v referencovaní. Prečo má referencovanie údajov taký veľký význam? Prečo je klasifikované ako legislatívna povinnosť pre subjekty verejnej správy? Prečo je vnímané ako nutná, a nie chcená povinnosť? Predchádzajúca veta je rečníckou otázkou predkladaného textu. Súčasná legislatíva nám jasne definuje povinnosti a práva vzťahujúce sa k používaniu hodnôt referenčných údajov vrátane procesu referencovania vo verejnej správe.

Realizácia referencovania údajov a naplnenie jeho významu sú podmienené vyhlásením nových údajov za referenčné. Za referenčné by sa mali vyhlasovať práve tie údaje, ktoré využívajú subjekty verejnej správy pri

realizácii princípu „jedenkrát a dosť“. Vyhlásenie nových skupín údajov za referenčné by malo byť silno motivujúce z dôvodu jednoznačného poskytovania a využívania údajov z pohľadu garancie ich úplnosti a správnosti a zároveň ich spoľahlivého použitia na právne účely pri výkone verejnej moci. Koncept referenčných údajov ponúka nástroje na sprehladnenie a zefektívnenie vzájomného vymieňania údajových tokov vo verejnej správe. K naplneniu želaného stavu by mali prispieť aj jednotlivé subjekty verejnej správy tým, že budú pri používaní hodnôt referenčných údajov postupovať v zmysle príslušných ustanovení zákona o e-Governmente.

Konkrétne to znamená, že subjekty verejnej správy, ktoré vo svojich registroch okrem iných údajov evidujú aj tie vedené v niektorom z referenčných registrov a vyhlásené za referenčné údaje, sú povinné referencovať tieto údaje voči údajom vedeným v referenčných registroch. Zákon o e-Governmente

v § 49 definuje referencovanie ako použitie údajov objektu evidencie v podobe, ktorá je zhodná s podobou, v ktorej je referenčný údaj vedený v referenčnom registri. Používanie je pritom definované ako používanie hodnôt referenčných údajov medzi registrami automatizovaným spôsobom, tzn. dátovou integráciou bez použitia ustanovenia o poskytovaní elektronických odpisov z ISVS a nutnosti elektronického podpísania použitých referenčných údajov. V praxi by malo dochádzať k priebežnému referencovaniu. Orgány verejnej moci (OVM) sú na účely udržiavania aktuálneho stavu registrov, ktoré vedú, oprávnené referencovať hodnoty referenčných údajov aj priebežne pri ich zmenách v referenčnom registri. Povinnosť referencovania sa nevzťahuje len k samotnej evidencii údajov v registroch, ale predovšetkým k výkonu verejnej moci. Ak OVM pri úradných konaniach zisťujú, preukazujú alebo inak spracúvajú údaje, ktoré sú referenčnými údajmi, sú

povinné referencovať príslušné hodnoty referenčných údajov.

Referencovaniu údajov predchádza stotožnenie údajov, ktorým sa rozumie jednoznačné priradenie hodnôt údajov k subjektu evidencie v registri (t.j. register vedený subjektom verejnej správy) a následné logické priradenie k tomu istému subjektu evidencie v referenčnom registri prostredníctvom identifikačného údajov, ktorým je identifikátor osoby alebo iný identifikátor referenčného údajov určený správcom referenčného registra. Účelom stotožnenia údajov je jednoznačné potvrdenie vzťahu medzi objektom evidencie a subjektom evidencie v registri a identifikovanie subjektu evidencie v registri a v referenčnom registri, a to najmä na účely následného bezchybného referencovania. Proces stotožnenia vlastných údajov s referenčnými údajmi môžeme označiť ako inicializačné referencovanie údajov.

Vyššie popísané skutočnosti nám naznačujú význam celého konceptu stotožnenia a referencovania údajov.

Zároveň nám napovedajú nutnosť zakotvenia povinnosti referencovania v legislatíve. Povinnosť zdieľania referenčných údajov automatizovaným spôsobom prispieva k zamedzeniu vzniku chybovosti údajov pri rôznych „polomanuálnych“ spracovaniach. Dochádza k logovaniu jednotlivých úkonov referencovania, čím sa zaznamenáva v prípade potreby dokazovania čas úkonu a údaje, ktoré boli referencované. Z legislatívy taktiež vyplýva ohlasovanie chybných údajov správcovi takéhoto referenčného registra. Výsledkom je zvýšenie dátovej kvality údajov zdieľaných v rámci verejnej správy pri povinnosti používať referenčné údaje.

Azda najväčšou pridanou hodnotou referenčných údajov stále zostáva to, že v sebe nesú domnienku správnosti a zabezpečenie právnej relevantnosti poskytovaných údajov, t.j. pri použití referenčných údajov je zabezpečená ochrana konzumenta.

OTVORENÉ ÚDAJE



**Matej
Kučera**

Podporujeme zverejňovanie najžiadanejších datasetov

Zverejňovanie otvorených údajov je jednou z priorít v oblasti informatizácie v Programovom vyhlásení vlády. Cieľom je dosiahnuť čo najľahší prístup k nim.

Koncept otvorených údajov vychádza z predpokladu, že ústredné orgány štátnej správy zhromažďujú, spravujú a aktívne sprístupňujú obrovské množstvo verejných údajov v rôznych oblastiach, napr. doprava, kultúra, financovanie, veda a výskum, počasie, životné prostredie, geografické údaje či rôzne štatistiky (nie údaje tajné, dôverné, citlivé, osobné a pod.).

Samotné sprístupňovanie týchto údajov je jednou z priorít v oblasti informatizácie Programového vyhlásenia vlády Slovenskej republiky na obdobie rokov 2020 – 2024. Jedným z cieľov je dosiahnuť jednoduchý prístup k informáciám

a údajom, čo spočíva práve v ich sprístupňovaní.

Zo všetkých vyššie spomenutých údajov, ktoré by mali byť aktívne sprístupňované sú niektoré verejnou žiadané viac, niektoré menej. S týmto súvisí dôležitá téma s názvom „Najžiadanejšie datasety“.

Úrad splnomocnenca vlády SR pre rozvoj občianskej vybavenosti robí prieskum dopytu verejnosti najžiadanejších datasetov otvorených údajov. Spomínaný prieskum sa realizuje každoročne v rámci Akčného plánu Iniciatívy pre otvorené vládnutie a môžete v ňom vidieť [datasety, ktoré sú najžiadanejšie](#) v spojitosti s jednotlivými orgánmi verejnej moci (OVM).

Dátová kancelária na základe dohody s Úradom splnomocnenca vlády SR pre rozvoj občianskej vybavenosti začala organizovať stretnutia s dátovými kurátormi jednotlivých OVM, ktorých hlavným cieľom je, aby boli najžiadanejšie datasety dostupné pre všetkých a aby boli publikované na portáli <https://data.gov.sk/>. Na stránke Dátovej kancelárie nájdete [prehľad uskutočnených stretnutí](#),

ktoré obsahujú aj samotný zápis zo stretnutia, kde sa môžete dozvedieť o priebehu zasadnutia, aké úlohy vyplynuli zo stretnutia pre jednotlivých účastníkov, poprípade kedy bude vybraný najžiadanejší dataset zverejnený na vyššie spomínanom portáli.

Pred samotným stretnutím s vybraným OVM na tému zverejňovania najžiadanejších datasetov Dátová kancelária v dostatočnom predstihu diskutuje s odborníkmi na danú problematiku na [platforme slovensko.digital](#). Okrem iného má diskusia za cieľ bližšiu konkrétnosť a špecifikáciu požiadaviek na datasety, ktoré obsahujú veľké množstvo údajov, ako aj nové návrhy datasetov, ktoré v prieskume chýbajú. Následne všetky tieto požiadavky odkomunikujeme priamo s dátovým kurátorom na stretnutí.

Takýto postup sa nám osvedčil a Dátová kancelária v ňom plánuje pokračovať aj naďalej. Výsledkom diskusií a stretnutí je, že splnenie cieľa o krajine s jednoduchým prístupom k informáciám a údajom sa začína rýchlejšie naplňať.

PETER BARLÍK

Peter Barlík nastúpil do Dátovej kancelárie v auguste 2020 a venuje sa téme analytických údajov. S dátovou analytikou prišiel do kontaktu už počas vysokoškolského štúdia regionálnej geografie na Prírodovedeckej fakulte Univerzity Komenského v Bratislave, kde bola jeho výskumná činnosť zameraná na analýzy priestorových dát s využitím v záchranných zložkách (regionalizácia a optimalizácia obvodov policajného, hasičského

a záchranného zboru) a maloobchodu (lokalizácia maloobchodných prevádzok, analýza spotrebiteľského správania, geomarketing a pod.)

Po skončení doktorandského štúdia v roku 2017 Peter Barlík nastúpil do spoločnosti Instarea (Big data monetization company) na pozíciu Data Scientist, kde mal skvelú možnosť

analyzovať telekomunikačné dáta a pretvárať ich do zmysluplných priestorových analýz pre súkromnú aj verejnú sféru (optimalizácia pobočkovej siete, dochádzkové vzorce obyvateľov, priestorové správanie obyvateľstva a pod.). Popri tejto pozícii nadobudol aj manažérske zručnosti, keďže viedol team softvérových developerov a zodpovedal za vývoj a stabilitu aplikácie na ciele marketingové oslovenie klientov.

Zručnosti získané v predchádzajúcom zamestnaní by rád využil aj v Dátovej kancelárii, kde zastupuje tému analytických údajov s jeho víziou štátu, ktorého rozhodnutia sa budú vykonávať na základe kvalitných a dostupných dát a vhodne zvolených analytických metód. „Verím, že k tomu dopomôžeme z pozície Dátovej kancelárie, ale aj pripravovaného projektu konsolidovanej analytickej vrstvy určeného pre analytické jednotky v štáte,“ hovorí Peter Barlík.

Mimo práce Peter Barlík rád trávi čas v prírode, a to na snehu, turistikou, pri rybolove alebo len tak s priateľmi.



MINISTERSTVO
INVESTÍCIÍ, REGIONÁLNEHO ROZVOJA
A INFORMATIZÁCIE
SLOVENSKEJ REPUBLIKY



datalab.digital
datalab.community

Ministerstvo investícií, regionálneho rozvoja a informatizácie Slovenskej republiky / Štefánikova 15 / 811 05 / Bratislava / datalab@vicepremier.gov.sk