Strategická priorita

**Otvorené údaje**

Pracovná verzia

(Verzia 0-3)

Informácia o dokumente

|  |  |
| --- | --- |
| Názov: | Strategická priorita Otvorené údaje |
| Stav: | Pracovná verzia |
| Pripravil: | Pracovná skupina K9.4 Lepšie dáta |
| Verzia: | 0.1 |
| Dátum: | 22.03.2016 |
| Pripomienkoval: |  |
| Dátum revízie: |  |

Distribučný zoznam

| Od | Dátum | Kontakt |
| --- | --- | --- |
| K9.4 |  |  |

| Pre | Akcia\* | Dátum (do) | Kontakt |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |

\* Akcia: Schváliť, Pripomienkovať, Informovať, Realizovať, iné (uveďte)

História verzií

| Verzia | Dátum verzie | Pripravil/ Zmenil | Pripomienkoval | Kľúčové zmeny |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Obsah

[1 Úvodný pohľad 4](#_Toc478501402)

[1.1 Cieľ dokumentu 4](#_Toc478501403)

[1.2 Definícia strategickej priority 4](#_Toc478501404)

[1.2.1 Kontext pre strategickú prioritu Manažment údajov 4](#_Toc478501405)

[1.2.2 Vysvetlenie strategickej priority 5](#_Toc478501406)

[1.2.3 Základné úlohy 6](#_Toc478501407)

[1.2.4 Základné témy 6](#_Toc478501408)

[1.2.5 Predstavenie konceptu 7](#_Toc478501409)

[1.3 Analýza súčasného stavu 7](#_Toc478501410)

[1.3.1 Sémantické dátové štandardy + LOD Slovakia + data.gov.sk ontológie 7](#_Toc478501411)

[1.3.2 Modul otvorených dát (MOD) 7](#_Toc478501412)

[2 Ciele realizácie 10](#_Toc478501413)

[2.1 Architektonické ciele 10](#_Toc478501414)

[2.1.1 Komplexný pokrok v téme otvorených údajov 10](#_Toc478501415)

[2.1.2 Zlepšiť dostupnosť údajov verejnej správy vo forme otvorených údajov 10](#_Toc478501416)

[2.1.3 Zvýšiť dynamiku a prínos dátového hospodárstva 12](#_Toc478501417)

[2.2 Dotknuté subjekty a ich záujmy 13](#_Toc478501445)

[3 Organizácia 16](#_Toc478501446)

[3.1 Zodpovednosť 16](#_Toc478501447)

[3.2 Posúdenie kapacitnej pripravenosti 16](#_Toc478501448)

[3.2.1 Personálna kapacita 16](#_Toc478501449)

[3.2.2 Kapacita informačných systémov 16](#_Toc478501450)

[3.3 Organizačné zmeny 16](#_Toc478501451)

[4 Stratégia 18](#_Toc478501452)

[4.1 SWOT analýza 18](#_Toc478501453)

[4.2 Vízia 18](#_Toc478501454)

[4.3 Výber prístupu 18](#_Toc478501455)

[4.4 Prehľad najlepších skúseností 19](#_Toc478501456)

[5 Architektúra riešenia 21](#_Toc478501457)

[5.1 Popis datasetov 21](#_Toc478501458)

[5.2 Princípy 23](#_Toc478501459)

[5.3 Biznis vrstva 25](#_Toc478501460)

[5.3.1 Prehľad požiadaviek 25](#_Toc478501461)

[5.3.2 Prehľad služieb 25](#_Toc478501462)

[5.3.3 Prehľad procesov 25](#_Toc478501463)

[5.4 Aplikačná vrstva 28](#_Toc478501464)

[5.4.1 Koncept architektúry 28](#_Toc478501465)

[5.4.2 Prehľad aplikačných funkcií a komponentov 29](#_Toc478501466)

[5.4.3 Aplikačné požiadavky 29](#_Toc478501467)

[5.4.4 Aplikačné služby 30](#_Toc478501468)

[5.4.5 Aplikačné komponenty 30](#_Toc478501469)

[5.5 Technologická vrstva 31](#_Toc478501470)

[6 Realizácia 32](#_Toc478501471)

[6.1 Riziká a otvorené otázky 32](#_Toc478501472)

[6.2 Plánovanie a migrácia (v súlade s akčným plánom) 33](#_Toc478501473)

[6.2.1 Aktivity a úlohy pre centrálnu úroveň 33](#_Toc478501474)

[6.2.2 Aktivity a úlohy pre inštitúcie verejnej správy 33](#_Toc478501475)

[6.2.3 Identifikácia projektov, ktoré je vhodné realizovať 33](#_Toc478501476)

[6.2.4 Požiadavky na legislatívne zmeny 33](#_Toc478501477)

[7 Udržateľnosť a financovanie 35](#_Toc478501478)

[8 Prílohy 36](#_Toc478501479)

[8.1 Skratky a definície 36](#_Toc478501480)

[8.2 Návrh legislatívnych úprav 39](#_Toc478501481)

# Úvodný pohľad

Nasledujúci dokument predstavuje pracovnú verziu Strategickej priority Otvorené údaje. Ide o  výstup pracovnej skupiny K9.4 Lepšie dáta a je možné vidieť návrh riešení. K dokumentu je možné sa ľubovoľne vyjadrovať a rozumné návrhy budú zvážené a zapracované.

## Cieľ dokumentu

Cieľom tohto dokumentu je podrobnejšie rozpracovanie strategickej priority 7 Otvorené údaje: Definovanie základných cieľov v rámci strategickej priority, načrtnutie možných alternatív ich dosiahnutia ako aj výber najvhodnejšej alternatívy na realizáciu. Definované ciele musia priniesť pozitívny vplyv na zvýšenie dostupnosti údajov verejnej správy pre spracovanie tretími stranami na základe princípov definovaných v Národnej koncepcii informatizácie verejnej správy.

## Definícia strategickej priority

### Kontext pre strategickú prioritu Manažment údajov

Národná koncepcia informatizácie verejnej správy z roku 2016 (ďalej tiež „NKIVS“) ustanovuje 10 strategických priorít informatizácie verejnej správy:

1 Multikanálový prístup,

2 Interakcia s verejnou správou, životné situácie a výber služby navigáciou,

3 Integrácia a orchestrácia,

4 Rozvoj agendových informačných systémov,

5 Využívanie centrálnych spoločných blokov,

6Riadenie údajov a Big data (Manažment údajov),

**7 Otvorené údaje,**

8 Vládny cloud,

9 Komunikačná infraštruktúra,

10 Kybernetická bezpečnosť.

NKIVS ku každej strategickej priorite informatizácie verejnej správy vysvetľuje jej cieľ, prístup k riešeniu a tiež rámcový architektonický model. Tento dokument nadväzuje na NKIVS a predstavuje návrh riešenia pre Otvorené údaje vo verejnej správe a rozširuje tak kapitolu **6.2.7 Otvorené údaje.** Tento dokument ešte môže byť z pozície architektonickej kancelárie verejnej správy upravovaný a dopĺňaný aj na základe pripomienok a komunikácie s gestorom tejto strategickej priority.

Strategická priorita Otvorené údaje by sa mala prakticky realizovať ihneď, ako to bude možné (prvé projekty nového programového obdobia). Za základné dôvody je možné považovať najmä:

* Všetky modernizované informačné systémy verejnej správy by mali byť schopné využívať nástroje pre riadenie údajov a poskytovať svoje údaje pre účely analýz (s výnimkou osobných, citlivých a utajovaných údajov). Znamená to, že základné komponenty pre riadenie údajov a analýzu údajov musia byť pripravované pred samotnou modernizáciou informačných systémov verejnej správy.
* Zavedenie metód lepšej regulácie patrí k zásadným prioritám prebiehajúcej reformy verejnej správy, a tým pádom i Operačného programu Efektívna verejná správa. Je preto žiaduce čím skôr podporiť tieto procesy vhodnými nástrojmi informačno-komunikačných technológií.

Dôležitým predpokladom k dosiahnutiu systematického rozvoja otvorených dát, je dopracovanie nosných prvkov systému MetaIS. Medzi kľúčové považujeme doplnenie funkcionalít v oblasti štandardizácie dátových prvkov vo forme URI a ontológií. Táto zmena zaručí, že zverejnené datasety, budú môcť vychádzať zo štandardizovaných dátových prvkov, čo v konečnom dôsledku výrazným spôsobom zjednoduší integráciu a tým aj efektivitu použitia otvorených dát.

Táto strategická priorita rozširuje aplikáciu princípu pristupovania k dátam ako k vzácnemu zdroju, pretože zverejňovanie otvorených údajov umožňuje nájsť ďalšie využitie dát aj mimo prostredia verejnej správy. Základným typom zverejňovaných údajov sú takzvané informácie verejnej správy (Public Sector Information - PSI), ktoré verejné orgány vytvárajú, zbierajú alebo za ne platia – teda sú ich vlastníkom alebo správcom. Príkladom sú geografické dáta, štatistiky, informácie o počasí, údaje z výskumných projektov financovaných z verejných zdrojov a digitalizované knihy z knižníc. Priorita predstavuje plynulé pokračovanie aktivít smerujúcich k zverejňovaniu otvorených údajov naprieč verejnou správou v rámci aktuálneho Akčného plánu Iniciatívy pre otvorené vládnutie na rok 2015. Navrhuje tiež nadstavbu nad nástrojmi pre zverejňovanie otvorených dát, ktoré sa budujú v projekte eDemokracia v rámci Operačného programu Informatizácia spoločnosti. Väčšinu zo zaujímavých dát je však v súčasnosti problematické licencovať pod otvorenou licenciou umožňujúcou jednoduché a bezodplatné opakované použitie aj na komerčné účely. Dôvodom sú predovšetkým nevysporiadané autorské práva s tvorcami informačných systémov a neexistujúca transpozícia európskej direktívy na opakované využívanie informácií verejného sektora (Directive 2003/98/EC, známa ako 'PSI Directive') do samostatného komplexného zákona.

### Vysvetlenie strategickej priority

Publikovanie otvorených údajov pre opakované použitie predstavuje samostatný špecifický cieľ programového obdobia 2014 až 2020, ktorý sa zameriava hlavne na rozsah a kvalitu publikovaných dát, ako aj na intenzitu vyžívania týchto dát podnikateľmi a kreatívnymi komunitami. Primárnym zámerom je využiť ekonomický potenciál otvorených dát verejnej správy, ktorých odhadovaná trhová hodnota v Európe je 32 miliárd EUR

Strategická priorita je v súlade s princípom pristupovania k dátam ako k vzácnemu zdroju, pretože ich zverejňovanie umožňuje nájsť ich optimálne využite aj mimo prostredia verejnej správy. Očakávané výsledky úspešnej realizácie strategickej priority sú podľa operačného programu nasledovné:

* Vďaka využívaniu otvorených dát sa výrazne zvýši miera transparentnosti verejnej správy, čím vznikne pozitívny dosah na jej efektivitu;
* Na Slovensku vznikne dostatočne silné odvetvie pracujúce s otvorenými dátami, pričom vzniknú kompetencie v progresívnych oblastiach informačných technológií, ako sú dátové analýzy;
* Inovatívne použitie otvorených dát vygeneruje pridanú hodnotu pre používateľov, ktorá sa premietne do ekonomického dopadu.
* Použitie štandardizovaných dátových schém zabezpečí efektívne použitie a integráciu

Otvorené údaje (Open Data) predstavujú údaje, ktoré sú dostupné v strojovo čitateľnom formáte a hocikto ich môže voľne využívať na komerčné aj nekomerčné účely, modifikovať, obohacovať, spájať a zdieľať maximálne s podmienkou uvedenia autora. V kontexte verejnej správy znamenajú otvorené dáta nasledovné:

* pre každú organizáciu je dostupný zverejnený zoznam jej dátových zdrojov, z ktorých možno publikovať otvorené dáta, a ich základný popis,
* údaje sú zverejnené v datasete (zostave údajov), ktorý je popísaný metadátami založených na štandardizovaných dátových prvkov uvedených v MetaIS a zaevidovaný v centrálnom katalógu otvorených údajov data.gov.sk,
* obsah dátového zdroja (samotné údaje) sú proaktívne bezplatne sprístupnené neobmedzenému okruhu záujemcov cez referenciu na miesto sprístupnenia datasetu a schémy údajov,
* údaje sú sprístupnené v štruktúrovanom a štandardizovanom tvare umožňujúcom ďalšie strojové spracúvanie, čo sa dosiahne dodržaním stanovených technologických štandardov, ktorých základom je používanie otvorených a technologicky neutrálnych riešení,
* na údaje sa možno právne spoľahnúť, pretože majú presne uvedeného pôvodcu, informáciu o aktuálnosti, o začiatku a konci účinnosti datasetu alebo jeho údajov, najmä ak je účinnosť odlišná od platnosti,
* ďalšiemu používaniu údajov (napríklad ich spracovaniu, zverejneniu, spájaniu s inými údajmi) nie sú kladené žiadne legálne prekážky. Tento fakt je jednoznačne uvedený v licencii, ktorá musí byť priradená ku každému datasetu.

### Základné úlohy

V rámci priority budú rozpracované opatrenia v nasledujúcich oblastiach.

Centrálna koordinácia publikovania otvorených dát:

* Jasne sa zadefinuje centrálna kapacita pre koordináciu a dohľad nad publikovaním otvorených dát v centrálnom katalógu data.gov.sk a opakovaným použitím otvorených dát.
* Zavedie sa dátový model verejnej správy, z ktorého budú vychádzať spoločné a doménové ontológie (spravované centrálne ontológmi v Centrálnej kapacite) spĺňajúce aj medzinárodné štandardy otvorených prelinkovaných dát (ďalej Linked dáta) a pravidlá sémantickej a syntaktickej interoperability v európskom priestore.
* Zavedú sa ďalšie automatizované metódy pre publikovanie otvorených dát vďaka implementácii Centrálneho dátového skladu.
* Zavedie sa dohľad nad kvalitou a štandardizáciou metadát a popisných štruktúr.
* Do praxe sa uvedie platforma pre štandardizované vytváranie a spravovanie aplikačných rozhraní (API) pre strojový prístup k otvoreným dátam.
* Začnú sa aplikovať nové postupy pre riadenie kvality datasetov a rizík ich publikovania, predovšetkým pre Linked dáta.
* Vytvorí sa centrálna právna kapacita pre riešenie právnych problémov so zverejňovaním a opakovaným použitím.

Zverejňovanie otvorených údajov u povinných osôb:

* Vďaka zavedeniu automatizovaných metód pre publikovanie otvorených dát budú povinné osoby značne odbremenené;
* Vyrieši sa problém efektívneho licencovania datasetov otvorených dát, najmä v oblasti informácií verejného sektora (Public Sector Information). Prideľovanie licencií bude podporovať jednoduché opakované použitie doma aj v zahraničí a v relevantných prípadoch, pri ktorých sa nejedná o PSI, bude účinne chrániť autorské práva.
* Zvýši sa kompetencia pre publikovanie prepojených údajov.
* Budú dostupné neustále sa zlepšujúce a interaktívne metodiky pre publikovanie otvorených údajov.

Zverejňovanie otvorených údajov v samospráve:

* Aplikovanie štandardizovaných formátov otvorených dát a API rozhraní do prostredia samosprávy;
* Vytvorenie centrálneho úložiska datasetov samosprávy v prostredí DCOMu s využitím centrálnych nástrojov pre publikovanie.
* Vytvorí sa dátový model samosprávy, štandardizujú sa samotné popisné štruktúry prioritných datasetov a metadáta pre samosprávu.
* Koordinovaná tvorba a aktualizácia odporúčaní a metodík pre fungovanie samospráv rôznych veľkostí tak, aby proces publikovania otvorených údajov prebiehal efektívne a výsledné dáta boli prínosom pre spoločnosť.

Podpora kreatívnych komunít a inovatívnych podnikov:

* Na základe hodnotenia spätnej väzby sa budú publikovať tie otvorené dáta a API, ktoré dokážu podniky a komunity využiť pre lepšie služby spoločnosti, zvyšovanie transparentnosti a pre rast digitálnej ekonomiky.

### Základné témy

Dokument prináša odpovede v nasledujúcich témach:

* Princípy dátového hospodárstva
* Legislatívny pohľad:
* „Open data first“ a vízia zákona o údajoch
* Pohľad architektúry:
* Čo má byť riešené na centrálnej úrovni a čo na lokálnej úrovni, aké sú kombinácie
* Pravidlá pre oprávnené používanie otvorených údajov:
* Licenčné modely pre otvorené údaje
* Ochrana súkromia a otvorené údaje:
* Rozsah zverejňovania údajov
* Anonymizácia údajov
* Procesy pre publikovanie otvorených údajov

### Predstavenie konceptu

## Analýza súčasného stavu

### Sémantické dátové štandardy + LOD Slovakia + data.gov.sk ontológie

V súčasnosti prebieha tvorba a štandardizácia sémantických dátových štandardov rešpektujúce nadnárodné EU modely a odporúčania. Agenda štandardizácie spadá pod PS1. Tieto štandardy popisujú nosné prvky a metodiky potrebné na systematické použítie linkovaných dát v priestore SR. Zároveň je rozpracovaných značné množstvo ontológií popisujúcich rôzne oblasti štátnych dát (RFO, RPO, RA, ...). Tieto svojím rozsahom ako aj významom už v súčasnosti plnohodnotne nahrádzajú a rozširujú pôvodnú ideu KDP(Katalóg dátových prvkov).

Nakoľko je problematika komplexná, predpokladá sa neustále rozširovanie a dopresňovanie štandardov.

### Modul otvorených dát (MOD)

Projekt eDemokracia momentálne nasadzuje nasledujúce komponenty, založené na osvedčených medzinárodných praktikách. Pri výbere použitých súčastí boli zároveň brané do úvahy technológie použité v európskom linked data softvérovom stacku LOD2 <http://lod2.eu>.

Tabuľka 1

| **Použité komponenty** | **Skratka** | **Funkcionalita** | **Licencia** |
| --- | --- | --- | --- |
| CKAN | CKAN | Katalóg otvorených dát | Opensource AGPL 3.0 |
| Pool Party | PPT | Ontologický subsystém, údržba riadených slovníkov, extrakcia konceptov | Proprietárna |
| Virtuoso Universal DB | Virtuoso | 70 % | Proprietárna |
| Unified Views | UV | Transformácie dát | Opensource GPL |
| XWiki |  | Crowdsourcing, komunikácia metodík | Opensource LGPL |
| Open data node | ODN | Podporná aplikácia pre povinné osoby | Opensource AGPL, LGPL |

Použité komponenty možno mapovať na nasledujúce komponenty MOD.

Tabuľka 2

| **Použitý komponent** | **Komponent MOD** |
| --- | --- |
| CKAN | čiastočne Správa katalógov, Repozitár datasetov, čiastočne Prezentačná aplikácia |
| PPT | Ontologický editor, Správa ontológií a identifikátorov |
| Virtuoso | RDF úložisko, čiastočne “SPARQL” Query engine |
| UV | Transformačný komponent, čiastočne Správa transformácií |
| ODN | Podporná aplikácia pre povinné osoby |
|  |  |

V nasledujúcej časti sú popísané kľúčové komponenty MOD, spomenuté vyššie.

Ontologický editor pre prepojené údaje

Ontologický editor umožňuje správu dátových slovníkov (ontológií).

Definície:

* Koncept – slovo alebo slovné spojenie s definovaným významom. Jednoznačne identifikované pomocou URI.

Príklady: Osoba, Povinná Osoba, Štát, IT Projekt, Bratislava (mesto), Prešov (mesto), Prešov (okres)

* Dátový slovník (Ontológia) – množina konceptov z konkrétnej domény. Medzi konceptami sú definované rôznorodé vzťahy. Každý koncept môže mať rôzne vlastnosti (atribúty).

Príklady: Geografický slovník (štáty, regióny, mestá, okresy, ...), Friend of a friend (rozšírený slovník používaný pre potrebu popisu osôb a vzťahov)

Slovník môže povoľovať jednoduché hierarchické vzťahy (nadriadený koncept – podriadený koncept), všeobecné symetrické vzťahy (koncept A "súvisí s" s konceptom B), prípadne iné typy vzťahov ("je podobný ako", "je taký istý ako", "je opak") alebo úplne arbitrárne vzťahy ("je veliteľom misie Apollo").

Ontologický editor spolu s publikačným nástrojom pre XML schémy po zvážení môže vystupovať ako úložisko ontológií a ich elementov na báze linked data.

Transformačný modul

Spravuje a vykonáva transformácie z textových formátov, z databázových zdrojov do RDF, interných relačných dát. Eviduje typy a postupy transformácií. Slúži na zaradenie rôznych zdrojov údajov do katalógu RDF alebo relačných datasetov. Umožňuje zvyšovanie kvality vstupných dát, obohacovanie datasetu, rozdielové analýzy, vyťažovanie neštruktúrovaných súborov.

Podporná aplikácia

Bude poskytovať Povinnej osobe nasledovné funkcionality:

* funkcionalita na zber (harvestovanie) a transformáciu údajov do open data formátov.
* funkcionalita na internú evidenciu zozbieraných a transformovaných dát (open data, ale aj non-open data) - interný katalóg
* funkcionalita na publikovanie (synchronizáciu) dát z internej evidencie do MODu

Na zabezpečenie týchto funkcionalít sú použité a upravené open source balíky Unified Views, CKAN a iné.

# Ciele realizácie

## Architektonické ciele

Táto kapitola obsahuje špecifikáciu architektonických cieľov , ktoré by mali byť dosiahnuté v oblasti manažmentu údajov verejnej správy do konca roku 2020. Ciele určujú konkrétnu pozitívnu zmenu, ktorá sa vo verejnej správe očakáva, ktorá bude podporovaná. Pre každý cieľ sú jasne stanovené merateľné ukazovatele, ktoré budú pravidelne vyhodnocované v rámci systematického monitoringu e-Governmentu.

Ciele sú záväzné, stav ich plnenia bude centrálne monitorovaný, posudzované projekty budú hodnotené aj z pohľadu príspevku k plneniu týchto cieľov.

### Komplexný pokrok v téme otvorených údajov

Tabuľka 3: Postavenie krajiny v téme otvorených údajov

|  |  |  |
| --- | --- | --- |
| Ciel | Ukazovateľ cieľa | Výsledok |
| Zlepšiť postavenie SR v medzinárodnom porovnávaní a stať sa lídrom v téme otvorených údajov v EU | Open Data Maturity index | Umiestnenie SR v prvej päťke |

**Open Data Maturity index** meria pripravenosť na realizáciu otvorených údajov, dopad využívania otvorených údajov a vyspelosť národného portálu pre otvorené údaje. SR dosahuje v súčasnosti v rebríčku Open data postavenie na 9. mieste a patrí do skupiny „fast-trackers“.

### Zlepšiť dostupnosť údajov verejnej správy vo forme otvorených údajov

Tabuľka 4: Dostupnosť údajov v štátnej správe

| Ciel | Ukazovateľ cieľa | Výsledok |
| --- | --- | --- |
| Zvýšiť počet inštitúcií verejnej správy, ktoré publikujú otvorené údaje | Podiel inštitúcií štátnej správy, ktoré publikujú otvorené údaje | 100 % |
| Podiel inštitúcií štátnej správe, ktoré publikujú aktuálne otvorené údaje | 75 % |
| Zvýšiť rozsah publikovaných údajov štátnej správy | Podiel **povinne zverejňovaných údajov** štátnej správy, dostupných vo forme otvorených údajov | 100 % |
| Podiel verejne prístupných údajov publikovaných ako otvorené údaje | 98 % z vhodných |
| Zvýšiť kvalitu publikovaných údajov štátnej správy | Podiel datasetov publikovaných minimálne v úrovni kvality 3★ (<http://3stardata.info/> )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ ) | 100 % |
| Podiel datasetov publikovaných minimálne v úrovni kvality 4★ (<http://4stardata.info/>) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitihttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitihttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitihttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužiti | 90 % |
| Podiel datasetov publikovaných minimálne v úrovni kvality 5★ (<http://5stardata.info/> ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie http://5stardata.info/ ) s vysokým potenciálom na znovupoužitie | 70 % |
| Umožniť používanie otvorených údajov na právne účely | Podiel **prioritných datasetov**, ktoré sú použiteľné na právne účely | 100 % z vhodných |
| Podiel údajov publikovaných ako otvorené údaje, ktoré sú použiteľné na právne účely | 90 % z vhodných |

Tabuľka 5: Dostupnosť údajov v samospráve

|  |  |  |
| --- | --- | --- |
| Ciel | Ukazovateľ cieľa | Výsledok |
| Zvýšiť počet inštitúcií samosprávy, ktoré publikujú otvorené údaje | Podiel inštitúcií samosprávy, ktoré publikujú otvorené údaje | VÚC 100 %  Mestá 100 %  Obce 70 % |
| Podiel inštitúcií samosprávy, ktorých dáta sú publikované ako **publikačné minimum pre samosprávu** | VÚC 100 %  Mestá 100 %  Obce 90 % |
| Zvýšiť kvalitu publikovaných údajov samosprávy | Podiel datasetov publikovaných v úrovni kvality 3★ (<http://3stardata.info/> )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ )http://3stardata.info/ ) | 100 % |
| Podiel datasetov publikovaných vo formáte 4★ (<http://4stardata.info/>) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitiehttp://4stardata.info/) s vysokým potenciálom na znovupoužitie | 80 % |
| Podiel datasetov publikovaných vo formáte 5★ (<http://5stardata.info/>) s vysokým potenciálom na znovupoužitie http://5stardata.info/) s vysokým potenciálom na znovupoužitie http://5stardata.info/) s vysokým potenciálom na znovupoužitie http://5stardata.info/) s vysokým potenciálom na znovupoužitie http://5stardata.info/) s vysokým potenciálom na znovupoužitie http://5stardata.info/) s vysokým potenciálom na znovupoužitie http://5stardata.info/) s vysokým potenciálom na znovupoužitie http://5stardata.info/) s vysokým potenciálom na znovupoužitie | 60 % |

Prioritné datasety

Prioritné datasety sú najžiadanejšie datasety od verejnosti, ktoré sú určené na základe pravidelného prieskumu medzi verejnosťou a zainteresovanými subjektami. Prioritné datasety budú schválené v procese manažmentu údajov. Zoznam je dynamicky aktualizovaný každoročne, pričom cieľom je jeho redukcia. (podľa OGP doplníme, kto je zodpovedný v akom čase)

Publikačné minimum pre samosprávu

Publikačné minimum pre samosprávu predstavuje zoznam datasetov podľa jednotlivých kategórií miest a obcí, ktoré by mala prioritne publikovať inštitúcia samosprávy. Tento zoznam datasetov definuje aj požiadavky na kvalitu, aktuálnosť a formát údajov (pre jednotlivé typy datasetov). Ako také bude toto publikačné minimum definované Úradom podpredsedu vládu SR pre informatizáciu a investície v spolupráci so splnomocnencom vlády SR pre rozvoj občianskej spoločnosti a predsedom Združenia miest a obcí Slovenska do 30. júna 2017, pričom publikačné minimum bude aktualizované.

Vysoký potenciál na znovupoužitie

Podiel inštitúcií verejnej správy, ktoré publikujú otvorené údaje

Ukazovateľ vyjadruje podiel inštitúcii verejnej správy na celkovom počte inštitúcií, ktoré publikujú otvorené dáta, ktoré sú zaevidované v centrálnom katalógu otvorených údajov. Dôležité je, aby inštitúcie dodržiavali procesy publikovania otvorených údajov.

Podiel inštitúcií verejnej správy, ktoré publikujú aktuálne otvorené údaje:

Aktuálne údaje sú údaje, ak respektíve ak sa údaje zhodujú s platnými údajmi v informačnom systéme verejnej správy a informácia o tejto zhode je dostupná respektíve ak požadovaná frekvencia zverejňovania datasetu je dodržiavaná.

Podiel údajov publikovaných ako otvorené údaje, ktoré sú použiteľné na právne účely

Dataset je použiteľný na právne účely v prípade, ak má zmysel na základe neho tvoriť rozhodnutia, ktoré ovplyvňujú vzťahy subjektov v občianskom alebo obchodnom práve, alebo sú informácie z datasetu relevantné v konaniach orgánov verejnej moci, respektíve ide o referenčné údaje.

Počet datasetov publikovaných vo formáte s vysokým potenciálom na znovupoužitie

Vysokým formátom na znovupoužitie sú myslené prepojené údaje: na identifikáciu entít sa používajú Jednotné referencovateľné identifikátory v podobe URI, jednotlivé údaje sú popísané prostredníctvom odporučených ontológiú EÚ na popis údajov respektíve prostredníctvom národných ontológií, ak také nie sú k dispozícií. Údaje tak vedia byť strojovo spracované s ohľadom na význam.

### Zvýšiť dynamiku a prínos dátového hospodárstva

Tabuľka 6: Dátové hospodárstvo

|  |  |  |
| --- | --- | --- |
| Ciel | Ukazovateľ cieľa | Výsledok |
| Zvýšiť dynamiku a prínosnosť dátového hospodárstva | Počet stiahnutí otvorených údajov ročne | 950 000 |
| Počet aplikácií, ktoré kreatívne využívajú otvorené údaje | 200 |
| Zlepšiť zapojenie Slovenska do jednotného digitálneho trhu (DSM) | Podiel datasetov dostupných v anglickom jazyku | 30 % |
| Katalóg datasetov predložený v anglickom jazyku | 100 % |

Počet stiahnutí otvorených údajov

Ide o stiahnutie datasetov cez katalóg otvorených údajov aplikáciami alebo používateľmi, evidované na platforme pre otvorené údaje.

Počet aplikácií, ktoré kreatívne využívajú otvorené údaje

Aplikácia počítaná v prípade, ak je registrovaná na platforme data.gov.sk a pre svoju funkčnosť využíva dostupné údaje verejnej správy cez centrálny katalóg.



## Dotknuté subjekty a ich záujmy

V tejto časti sú identifikované dotknuté subjekty, ktorých sa daná strategická priorita (alebo jej časť) týka a ich relevantné záujmy.

Štúdia otvorených dát (The Open Data Study (2010)) zistila, že na úspešnú implementáciu programu otvorených dát verejnej správy (OGD) sú nevyhnutné tri kľúčové skupiny stakeholderov, alebo inak povedané tri vrstvy spoločnosti:

* Vplyvná a aktívna občianska spoločnosť, ktorá vytvára tlak zdola na zmenu tradičných postupov informovania verejnosti a vytvára inovatívne služby, produkty a webové stránky, ktoré ukazujú spôsob využitia OGD.
* Zamestnanci verejnej správy, ktorí vnímajú OGD ako spôsob zvýšenia efektivity a transparentnosti, predstavujú strednú vrstvu.
* Vysoko postavení politickí lídri ako ministri tvoria tretiu vrstvu a tvoria víziu otvoreného vládnutia.

Ako v rámci spomenutej štúdie odpozoroval Sir Tim Berners-Lee: *„Úspešná realizácia OGD musí začať zhora, v strede aj zospodu!”* Tento prístup je všeobecne označovaný za trojstupňový prístup implementácie OGD a zahŕňa angažovanie sa všetkých vrstiev dotknutých subjektov.

Úrad podpredsedu vlády SR pre informatizáciu a investície

* Mať politický mandát, právnu a odbornú kompetenciu (v oblasti licencií, legislatívy, technológií a dátových analýz) riešiť problémy a spory pri publikovaní otvorených dát, napríklad v oblasti licencií alebo platnosti dát,
* Zabezpečiť vysoký počet užitočných datasetov otvorených dát vo vysokej kvalite, ktorú budú intenzívne využívané inovatívnymi podnikmi a komunitami,
* Efektívne vykonávať dohľad nad zverejňovaním otvorených dát a dodržiavaním štandardov a najlepších praktík.
* Zastrešenie tvorby a štandardizácie dátových prvkov, ontológií a URI v systéme MetaIS.

Úrad splnomocnenca vlády pre rozvoj otvorenej spoločnosti

* Aby dochádzalo k plneniu úloh plánu iniciatívy pre otvorené vládnutie (OGP) a ich koordinácii s aktivitami e-Govermentu.

Správca informačného programu verejnej správy

* Mať možnosti pre plnenie povinností súvisiacich s rolou správcu informačného systému verejnej správy a poskytovateľa údajov (vrátane jasného návodu pre postup, ako zverejňovať otvorené údaje),
* Mať k dispozícii dostatok prostriedkov a kapacít na tvorbu a publikovanie kvalitných údajov.

Ministerstvá, samospráva a ďalšie štátne organizácie

* Získavať konštruktívnu spätnú väzbu od verejnosti a neziskového sektora,
* Legitimizovať pred verejnosťou svoje rozhodnutia a schopnosť zavádzať pozitívne opatrenia do praxe,
* Zaviesť princípy štátu, fungujúceho na základe využívania dát v praxi,
* Vykonávať dohľad nad efektívnosťou opatrení,
* Mať kontrolu nad efektívnosťou vynakladaných zdrojov a investícií,
* Efektívne plniť povinnosti vyplývajúce zo Zákona o slobodnom prístupe k informáciám.

Úrad vlády SR

* Efektívne implementovať Akčný plán Iniciatívy pre otvorené vládnutie na rok 2017 až 2019 a viesť centrálny katalóg otvorených dát prostredníctvom agentúry NASES.
* Komunikovať s tretím sektorom.

Verejnosť

* Mať zabezpečenú ochranu súkromia,
* Mať jednoduchý, zrozumiteľný a centrálny prístup ku všetkým otvoreným dátam verejnej správy a samosprávy,
* Efektívne aplikovať práva vyplývajúce zo Zákona o slobodnom prístupe k informáciám,
* Zvyšovať transparentnosť vo verejnej správe a v procesoch verejného obstarávania,
* Mať prehľad o predpokladaných dopadoch politík a regulácií,
* Mať prehľad o zdôvodnení investícii vo verejnom záujme,
* Mať prehľad o výkone verejnej správy v jednotlivých sektoroch (až na úroveň jednotlivých inštitúcií) – vrátane ratingov,
* Možnosť participatívnej účasti.
* Mať prístup k navrhovaným štandardom prostredníctvom zverejňovania v systéme MetaIS
* Mať prístup k definíciám schválených štandardov

Občianske komunity a inovatívne komunity

* Okrem vyššie uvedeného, chcú mať k dispozícií kvalitné údaje v strojov čitateľnom formáte a nástroje (platformy a aplikačné rozhrania), aby mohli vytvárať aplikácie klasifikované do troch základných kategórií:
* Transparentnosť a zodpovednosť, ktoré nútia verejnú správu deklarovať svoju zodpovednosť tým, že informácie a procesy budú transparentné. Projekty v tejto oblasti často smerujú k poradnej činnosti komunít a k ovplyvňovaniu politík a rozhodovania.
* Interakcia občanov a verejnej správy, ktorá uľahčuje komunikáciu a dáva jej nový rozmer, tieto projekty majú za cieľ nahradiť tradičné kanály a spôsoby komunikácie elektronickými verejnými službami, ktoré sú orientované na používateľa a poskytujú vyšší používateľský komfort.
* Digitálne nástroje pre každodenný život občanov, ktoré ich život uľahčujú, tieto projekty sa často orientujú na malé, ale dôležité aspekty nášho každodenného života, s cieľom poskytnúť riešenia jednoduché na použitie.

Táto skupina subjektov sa považuje vždy za pionierov v oblasti rozvoja programov OGD, hlavne vďaka implementácií pilotných projektov preukazujúcich úžitok zo zverejňovania otvorených dát. Odporúča sa tiež orientovať na vytváranie takzvaných verejno-komunitných partnerstiev, ktoré budú experimentovať s novými typmi spolupráce medzi týmito veľmi rozdielne fungujúcimi subjektami[[1]](#footnote-2).

Podnikateľské prostredie

* Mať dostupné nástroje, API a otvorené dáta a pre poskytovanie inovatívnych komerčných služieb,
* Regulácie, ktoré podporujú dátové hospodárstvo a využitie údajov (vrátane jasného modelu licencií),
* Mať prístup ku kvalitným analýzam a prognózam o vývoji ekonomiky a o plánovaných opatreniach a obstarávaniach verejnej správy,
* Zvyšovať transparentnosť verejnej správy pre zlepšovanie podnikateľských príležitostí,
* Možnosť participatívnej účasti na príprave regulácií.
* Možnosť realizovať rizikové inovácie financované štátom.

# Organizácia

Návrh modelu zodpovedností a potrebné organizačné zabezpečenie vrátane organizačných zmien.

## Zodpovednosť

Zodpovedný gestor

Gestorom strategickej priority Otvorené údaje je je **Úrad podpredsedu vlády SR pre informatizáciu a investície** –

* Je nevyhnutné vytvoriť Centrálnu kapacitu pre koordináciu publikovania otvorených dát vo verejnej správe a v samospráve.
* Dopracovanie Metainformačného systému do podoby, ktorá zabezpečí systematický, udržateľný a efektívný rozvoj otvorených linkovaných dát

Zapojené organizácie

Zoznam organizácií, ktoré budú dotknuté realizáciou priority:

* **NASES** - prevádzkuje Centrálny katalóg otvorených údajov, ďalej „centrálny katalóg“, na Ústrednom portáli verejnej správy. Centrálny katalóg je informačný systém verejnej správy, ktorý obsahuje údaje o informáciách sprístupňovaných na opakované použitie a otvorených údajoch,najmä ich obsah, meta-údaje, podmienky použitia a obmedzenia použitia.
* Zodpovedá za dereferenciáciu URI
* **Úrad splnomocnenca** **Úrad vlády SR** - **Všetky organizácie verejnej správy a samosprávy**, pretože na základe pripravovaného právneho stavu majú byť sprístupnené povinnými osobami všetky údaje, okrem tých, pre ktoré to špecificky určitý zákon vylučuje.

## Posúdenie kapacitnej pripravenosti

### Personálna kapacita

Jednotlivé zložky štátnej správy nie sú z hľadiska personálnej kapacity na zmeny pripravené.

### Kapacita informačných systémov

Informačné systémy verejnej správy sú vo všeobecnosti schopné integrácie údajov. Je však potrebná ich modifikácia, najmä v súvislosti:

* s implementáciou princípov pre prepojené údaje (Linked Data),
* efektívnym zabezpečením využívania mechanizmov správy a zdieľania údajov, najmä riadenie kvality údajov a využitie platformy integrácie údajov.

## Organizačné zmeny

Aby bolo možné realizovať strategickú prioritu Open Data, sú navrhované nasledujúce organizačné zmeny:

* Doplnenie Dátovej kancelárie verejnej správy o nové kompetentcie, ktorá bude mať na starosti správu dátovej vrstvy, ontológií **a centrálnu koordináciu zverejňovania otvorených údajov v správnom formáte**. Vhodné by bolo tiež disponovať právnymi odborníkmi v oblasti dát a licencovania.
* Zriadenie Centrálnej kapacity pre koordináciu publikovania otvorených údajov v štátnej správe a samospráve.
* Zriadenie Centrálnej kapacity pre koordináciu publikovania otvorených údajov v samospráve

Nové role:

Na centrálnej úrovni a lokálnej úrovni budú menované nasledujúce role:

Tabuľka 7: Prehľad organizačného zabezpečenia

| Úroveň | Zodpovednosť | Rola | Počet |
| --- | --- | --- | --- |
| Centrálna  (ÚPPVII) | * Celková zodpovednosť za otvorených údajov vo verejnej správe | CDO | 1 |
| Väčšia inštitúcia[[2]](#footnote-3) | * Životný cyklus údajov * Kvalita údajov | Vlastník údajov | >1 (pre každý segment údajov) |
| Dátový kurátor | >1 (pre každý segment údajov) |
| * Analytické spracovanie údajov | Dátový analytik | 1 |
| Menšia inštitúcia | * Životný cyklus údajov * Kvalita údajov | Vlastník údajov | 1 |
| Dátový kurátor | 1 |

# Stratégia

Účelom tejto kapitoly je výber strategického prístupu, ktorý nás privedie k dosiahnutiu cieľov a želaného zlepšenia. Jednotlivé témy, ako referenčné údaje, kvalita údajov sú podrobnejšie vysvetlené v kontexte slovenského prostredia. Identifikované sú základné varianty, ako k problematike pristúpiť, vrátane výberu vhodného riešenia.

## SWOT analýza

Tabuľka 8: SWOT analýza pre manažment údajov

|  |  |
| --- | --- |
| Silné stránky | Slabé stránky |
| * Spustenie portálu data.gov.sk * Implementácia projektu Elektronické služby Úradu vlády SR - eDemokracia a otvorená vláda * Rastúci počet datasetov verejnej správy * Zabezpečené financovanie ďalších aktivít v oblasti otvorených údajov v rámci OP II a OP EVS * Podpora vlády SR a vytvorenie samostatného ÚOŠS, ktorý komplexne riadi informatizáciu * Zapojenie odbornej verejnosti a mimovládnych organizácií do riešenia problematiky otvorených údajov * Postavenie SR v rámci krajín EÚ | * Chýbajúca komplexná legislatíva a metodická podpora pre oblasť otvorených údajov * Chýbajúce jednotné pravidlá a princípy pre poskytovanie otvorených údajov v rámci verejnej správy * Nízky počet zverejnených datasetov v pomere k informáciám, ktorými verejná správa disponuje * Chýbajúci komplexný prehľad o existujúcich údajoch vo verejnej správe * Chýbajúca kontrola poskytovania otvorených údajov a nedostatočná vymáhateľnosť * Nedostatočné kapacity na strane verejnej správy * Nedostatočná komunikácia medzi verejnou správou a socioekonomickými partnermi * Nekvalitné alebo chýbajúce metadáta * Nejednotné používanie slovníkov * Nízka kvalita zverejnených datasetov * Chýbajúca osveta voči občanom a podnikateľom * Nedostatky týkajuce sa používateľskej prívetivosti a odozvy |
| Príležitosti | Hrozby |
| * Prínos otvorených údajov do rozvoja ekonomiky štátu * Zvýšenie konkurencieschopnosti SR * Zvýšenie transparentnosti a dôveryhodnosti verejnej správy * Zvýšenie spokojnosti občanov s verejnou správou * Väčšie zapojenie verejnosti do správy veci verejných * Nové služby pre občanov zo strany verejnej správy * Nové služby a produkty zo strany súkromného sektoru * Podpora inovatívnych služieb * Podpora začínajúcich podnikateľov * Zlepšenie procesov vo verejnej správe na základe nových údajov * Zlepšenie politík a legislatívy na základe nových údajov * Možnosť znovupoužitia údajov, resp. využitia existujúcich údajov bez potreby vytvárať duplicitné údaje * Nové údaje a podnikateľské možnosti získané kombináciou a spojením verejných a súkromných údajov * Vytvorenie centrálneho portálu, kde by mohli podnikatelia zverejňovať aplikácie založené na otvorených údajoch | * Odmietavý postoj úradníkov verejnej správy * Zmeny politiky v oblasti otvorených údajov v rámci politického cyklu * Nejednotná politika v oblasti otvorených údajov * Nejednotná licenčná politika * Chýbajúce štandardy * Nedostatok finančných zdrojov na realizáciu zverejňovania otvorených údajov obzvlášť u menších úradov * Zverejňovanie len otvorených údajov, ktoré nemajú potenciál na znovupoužitie * Ekonomický model založený na spoplatnení otvorených údajov (priamy príjem z otvorených údajov vs. nepriamy príjem v podobe vyšších daňových príjmov) * Nedostupnosť otvorených údajov, resp. ich obmedzená prístupnosť * Nepripravenosť podnikateľov na využívanie otvorených údajov * Slabá odozva verejnej správy na podnety verejnosti * Nedostatočná ochrana osobných a citlivých údajov * Nedostatočná aktualizácia datasetov * Neschopnosť verejnej správy poskytovať nadstavbové služby |

## Vízia

Slovensko bude transformované na „Data-driven state“ – štáte, fungujúcom na základe využívania dát a presnom riadení celého životného cyklu dát Štát bude prijímať rozhodnutia na základe najlepších znalostí, ktoré sú k dispozícii. Aby bolo možné dosiahnuť takýto stav:

* Údaje budú spravované ako vzácny zdroj.
* Vytvoria sa podmienky pre maximálne využitie potenciálu, ktorý je možné vyťažiť z údajov.
* Verejná správa bude schopná komplexne pokryť životný cyklus údajov.

Znamená to tiež, že pre jednotlivé oblasti rozhodovania budú k dispozícií modely, dáta a nástroje, ktoré umožnia vytvoriť analýzy, na základe ktorých sa podporí rozhodovanie. Znamená to tiež transformáciu fungovania organizácií a procesov vo verejnej správe tak, aby boli tieto analýzy efektívne používané a zároveň zverejňované vo vhodnej vizuálnej podobne pre verejnosť.

Treba zaviesť jednotné riešenie, ktoré však nebude narúšať autonómiu jednotlivých inštitúcii verejnej správy v oblasti tvorby údajov a riadenia prístup k nim.

## Výber prístupu

Prístup k systematickému riešeniu Open Data je možné zhrnúť ako:

* Implementovať lessons learned z projektu eDemokracia,
* Ako súčasť reformy verejnej správy, vytvorí sa centrálna kapacita schopná centralizovane a systematicky riadiť a koordinovať zverejňovanie a používanie otvorených dát vo verejnej správe a v samospráve:
* Manažment dátového modelu verejnej správy a samosprávy(vrátane ontológie a definície prepojení),
* Spolupráca pri návrhu analytických produktov, definícia požiadaviek na dátové zdroje,
* Riadenie zberu dát z informačného prostredia verejnej správy a samosprávy,
* Licencovanie dátových zdrojov pre otvorené dáta.
* Na jednotlivých ministerstvách budú zriadené analytické jednotky zodpovedné za konkrétne oblasti verejnej správy (politiky, regulácie, výkonnosť, investície) a so schopnosťami zverejňovať otvorené dáta (obdobné kapacity budú vybudované aj pre samosprávu).
* Podporované budú segmentové iniciatívy s cieľom komplexného zabezpečenia publikovania v segmente a v samospráve:
* Identifikácia datasetov, ktoré sa budú zverejňovať v rámci segmentu,
* Nastavenie procesov zverejňovania dát,
* Zabezpečenie, že datasety sú pravidelne publikovanie cez portál „data.gov.sk“.
* Jednotlivé informačné systému verejnej správy budú pripájané k Centrálnemu dátovému skladu cez dátovú integráciu a ETL procesy.

### Otvorené údaje a prístup k dátovému hospodárstvu

Ak chceme, aby naše dátové hospodárstvo vytváralo rast a pracovné miesta, musia sa využívať dáta. Ale na to, aby sa využívali, je tiež potrebné, aby boli dostupné a aby sa analyzovali.

Dátové hospodárstvo predstavuje súbor ekonomických aktivít, ktoré súvisia so zberom, spracovaním, analyzovaním údajov a tvorbou vyplývajúcej pridanej hodnoty. Hlavným motorom dátovej ekonomiky v súčasnosti je spracovanie veľkých údajov a aplikácia umelej inteligencie, ako i fenomén digitálnych platforiem, ktoré vytvárajú nové a disruptujú tradičné biznis modely.

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| Údaje sa zverejnenia ako otvorené v maximálne možnom rozsahu - štát "nezaujíma" použitie údajov, to je úlohou dátového hospodárstva. Autorské práva sú riešené otvorenou licenciou | X | * Schválené už v Stratégií XY, Akčnom pláne OGP a NKIVS a v smernici PSI |
| Údaje by sa nemali používať na komerčné účely. |  |  |
| Dátové hospodárstvo nie je podporované otvorenými údajmi. |  |  |

### Ako vytvoriť funkčnú komunitu

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| Dopytové výzvy na tému otvorené údaje, kde žiadateľom respektíve partnerom budú neziskové organizácie a MSP - stimulácia vytvárania aplikácií verejnými zdrojmi | ? | * Je potrebné správne nastaviť kritéria výberu * Zamerať sa aj na hodnotné aplikácie pre spoločnosť (nie len komerčné) |
| Urýchlené zverejnenie zaujímavých a kvalitných otvorených údajov (prioritných datasetov) | X |  |
| Medzinárodná spolupráca a integrácia s nadnárodnými bázami údajov | X | * Pozitívny príklad sú aktivity v oblasti priestorových údajov |
| Organizácia hackathonov a komunitných podujatí | X |  |
| Fórum, komunikácia a právna istota (aj zber spätnej väzby) | X |  |
| Podpora spracovania otvorených údajov univerzitami. Výskum, s ktorými výskumníci pracujú by mali byť otvorené |  | * Je nevyhnutné zapojiť univerzity do témy otvorených údajov |
| Podpora komunity zo strany štátu nie je potrebná |  |  |
| Grantová agentúra a pôžičkový systém |  |  |

### Model pre „governance“ otvorených údajov

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| „laissez-faire":   * bez spoločnej dátovej politiky, publikovanie na základe dobrovoľnosti a spolupráce |  |  |
| Hierarchický model („command and control“):   * centrálne určovanie politík a lokálne publikovanie údajov na základe nastavených cieľov |  |  |
| Zmiešaný model:   * centrálne určovanie dátovej politiky a koordináciu * **iniciatíva a zodpovednosť** na strane inštitúcií * možnosť využívať centrálne nástroje pre identifikáciu a transformáciu údajov * možnosť využívať centrálne nástroje pre publikovanie údajov * možnosť využívať centrálne dátové služby | X | * v prípade štátnej správy |
| Centralistický model:   * centrálne určovanie politík * a **centrálna autorita zodpovedná** za publikovanie údajov |  |  |
| Model „outsourcing“:   * publikovanie údajov môže byť "outsourcované" súkromným "data providerom" | X | * V prípade samosprávy |
|  |  |  |

### Spôsob zabezpečenia kvality otvorených údajov

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| Kvalita otvorených údajov bude riešená rovnakými nástrojmi ako Manažment údajov (ÚPPVIaI resp. NASES) | X |  |
| Kvalita bude posudzovaná nezávislou komisiou (aj automatizovane) – kompetencia by bola komunite |  |  |
| Kvalita bude posudzovaná nezávislou komisiou (aj automatizovane) – nezávislý Úrad |  |  |
| Kvalita bude zvyšovaná službami pre spätnú väzbu a jej zakomponovanie do procesov Manažmentu údajov | X |  |

### Spôsob prístupu (access) k otvoreným údajom

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| prostredníctvom štandardizovaného aplikačného rozhrania (API) | X | Preferované |
| použitím news-feeedu (RSS/Atom) | X |  |
| cez SPARQL uzol | X |  |
| dávkové súbory | X |  |
| CD 😊 |  |  |
|  |  |  |

### Výber formátov

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| JSON, CSV, XML |  |  |
| Prepojené údaje („Linked-data“) |  |  |
| Priestorové údaje |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

### Miera centralizácie

* Čo má byť riešené na centrálnej úrovni a čo na lokálnej úrovni, aké sú kombinácie

| Variant | Riešenie | Zdôvodnenie |
| --- | --- | --- |
| Model údajov | * Jeden centrálny: evidovaný v metaIS |  |
| Manažment identifikácie | * Jeden centrálny: evidovaný v MetaIS |  |
| Katalóg pre otvorené údaje | * Jeden centrálny „data.gov.sk“ |  |
| Dátové úložisko | * Kombinácia: centrálne úložisko môže existovať, ale inštitúcie majú kompetenciu rozhodnúť o použití nástrojov, na základe: kritérií efektivity a kritérií kvality |  |
| Nástroje pre integráciu, transformáciu a publikovanie údajov | * Kombinácia: centrálny nástroj môže existovať, ale inštitúcie majú kompetenciu rozhodnúť o použití nástrojov, na základe: kritérií efektivity a kritérií kvality – pričom je možné využívať aj riešenia hybridného cloudu |  |

### Výber licenčných modelov

Licencia otvorených údajov musí byť vybraná povinne.

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| Creative Commons CCZero (CC0) | X - default | * Odporúčanie EK |
| Open Data Commons Public Domain Dedication and Licence (PDDL) |  |  |
| Creative Commons Attribution 4.0 (CC-BY-4.0) | X | * Iba na základe vážneho zdôvodnenia |
| Open Data Commons Attribution License (ODC-BY) |  |  |
| Creative Commons Attribution Share-Alike 4.0 (CC-BY-SA-4.0) |  |  |
| Špeciálna licencia (na mieru) |  |  |

* Navrhnúť mechanizmus, kto by mal zverejňovať otvorené údaje v prípade previazaných zdrojov údajov (zdrojové registre vs referenčné registre)

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| Zdrojový register |  |  |
| Referenčný register |  |  |
| Platforma integrácie údajov |  |  |

### Ochrana súkromia a otvorené údaje

Rozsah zverejňovania údajov

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| Zverejňujú sa anonymizované údaje | X |  |
| Osobné údaje a citlivé údaje sú vypustené z datasetu |  |  |
| Osobné údaje a citlivé údaje sú vypustené z datasetu, okrem výnimiek:   * dostupné vo verejných listinách, * údaje štátnych zamestnancoch z dôvodu transparentnosti, * .... |  |  |

Anonymizácia údajov

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| Agregácia údajov |  |  |
| „Top and bottom coding“ |  |  |
| Odstránenie osobných údajov z datasetov |  |  |
| Použitie pseudo-identifikátorov (crypographic hash) | X |  |
| Zahmlenie údajov (odchýlky, aby bolo možné štatistické spracovanie) | X |  |

Prepájanie údajov, vo vzťahu k osobným údajom

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| Prepájanie osobných údajov je limitované |  |  |
| Prepájanie osobných údajov je v zodpovednosti osoby, ktorá prepája údaje | X |  |

### Spôsob zabezpečenia dôveryhodnosti

| Variant | Výber | Zdôvodnenie |
| --- | --- | --- |
| Podpisovanie KEP |  |  |
| „blockchain“ | X |  |
| Bez špecifických technických opatrení |  |  |

## Prehľad najlepších skúseností

Nasledujúca kapitola uvádza prehlaď najlepších skúseností (best-practice) zo zahraničia, ktoré sú relevantné pre strategickú prioritu, respektíve sa z nich vychádza.

Portály pre otvorené údaje

* Open data portály v EU, UK a USA,
* Medzinárodný projekt COMSODE,
* European Public Sector Information (PSI) Platform.

Prepojené údaje (Linked dáta)

**TODO**

Ontológie

* Ontológie - SEMIC odporučené ontológie pre popis Fyzickej osoby, Organizácie, Lokácie a podobne.

Dátová kvalita

* Metodika Five Star Open Data.
* Vychádza sa z odporúčaní the **Open Data Foundation, the DDI Alliance**, the International Household Survey Network, ktoré sa zameriavajú na kvalitu dát aj z pohľadu sekundárneho využitia.

Metadáta

* Kľúčové je propagovať dôležitosť kvalitnej dokumentácie dát pomocou štandardov na metadáta na základe najlepších praktík. Metadáta sa musia štandardizovať naprieč celou verejnou správou jednotným prístupom podporujúcim rôzne štandardizované metamodely (ISA2 odporučené ontológie, INSPIRE), pričom kľúčovým prvkom v ich správe je Centrálny metainformačný systém. Zakomponovanie popisu metadát do každej .tapy životného cyklu – je potrebné dokumentovať v okamihu udalosti Metadáta možno ďalej využiť na automatické generovanie štatistických skriptov alebo vytváranie dokumentácie, čo znižuje celkové náklade a zvyšuje kvalitu.

# Architektúra riešenia

V kapitole sú navrhnuté princípy, ktoré budú vyžadované od riešení vo verejnej správe z pohľadu otvorených údajov.

## Popis datasetov

## Na popísanie datasetu publikovaného v minimálnej kvalite 4★ bude použitý medzinárodný štandard DCAT (Data Catalog Vocabulary) , ktorý poskytuje prostriedky pre definíciu katalógov, datasetov a ich distribúcii v čase. S ohľadom na interoperabilitu a odporúčania Európskej komisie budú požadované vlastnosti publikovaného datasetu formálne vyjadrené prostredníctvom tzv. Aplikačného profilu pre publikáciu otvorených dát verejnej správy Slovenskej republiky DCAT-AP-SK. Štandardizácia uvedených požadovaných vlastností datasetu obsahujúceho dáta s vysokou interoperabiliou je v súčasnosti predmetom pracovnej skupiny PS1.

Detailná špecifikácia datasetov a prislúchajúcich metaúdajov bude štandardizovaná v rámci pracovnej skupiny PS1 a zverejnená v podrobnej metodike.

## Princípy

Zoznam architektonických princípov, na ktorých je riešenie postavené:

## Biznis vrstva

### Prehľad požiadaviek

Všeobecné požiadavky

* Publikovanie dát bude efektívne, jednoduché a transparentné, a čo v najvyššej možnej miere automatizované.
* Údaje budú identifikované pomocou URI, pričom údaje centrálnej povahy budú popísané prostredníctvom Centrálneho modelu údajov.
* Otvorené dáta z nových alebo modernizovaných systémov budú publikované vo vysokej kvalite interoperability spĺňajúcej úroveň minimálne 4★ pričom pre údaje centrálnej povahy je to 5★.
* Otvorené dáta budú v cieľom stave publikované vo vysokej kvalite spĺňajúcej úroveň 4★ alebo 5★, nebude sa brániť publikovaniu na úrovni 3★.
* Podnikatelia, verejnosť a kreatívne komunity sa budú môcť jednoducho zapojiť do procesu otvoreného vládnutia.

Požiadavky na centrálne riešenie

* Riešenie obsahuje centrálne automatické transformácie dát.
* Riešenie umožňuje jednoducho prideliť licenciu dátovému zdroju.
* Existuje centrálny dátový model verejnej správy vrátane ontológií pre otvorené dáta.
* Bude efektívne monitorovaný dopad opatrení otvorených dát a v prípade nedosahovania cieľov budú dostupné nápravné opatrenia.
* Datasety je možné referencovať, zdieľať, vyhľadávať, komentovať a sú dostupné cez jednotný prístup.
* Budú dostupné neustále sa zlepšujúce a interaktívne metodiky pre publikovanie otvorených dát.
* Pomoc s interpretáciou pre identifikáciu vhodných datasetov.

Požiadavky na inštitúcie verejnej správy

* Implementácia procesov pre publikovanie otvorených údajov (v rámci čoho je kľúčový zber požiadaviek a ich vyhodnocovanie)

### Prehľad služieb

Zoznam služieb (aké služby budú k dispozícii pre otvorené údaje):

* Vyhodnocovanie využívania API
* Vyhodnocovanie sťahovania datasetov
* Vytvorenie prieskumu verejnej mienky o poskytovaných datasetoch
* Riadenie vzťahu s neziskovým sektorom a verejnej participácie
* Riešenie právnych sporov

### Prehľad procesov

V nasledujúcej kapitole sú identifikované základné procesy pre správu otvorených údajov, ktoré je potrebné nastaviť: definícia, ako sa budú procesy vykonávať a kto bude za nich zodpovedný.

Obrázok 1: Biznis procesy pre otvorené údaje



* Informovanie verejnosti prostredníctvom otvorených údajov

Tabuľka 9: Procesy pre životný cyklus údajov

| Proces | Centrálne | Lokálne |
| --- | --- | --- |
| * Riešenie vecí verejných |  |  |
| * Informovanie verejnosti prostredníctvom otvorených údajov |  |  |
| * Pripomienkovanie a hodnotenie datasetov |  |  |
| * Poskytovanie otvorených údajov |  |  |
| * Príprava plánu pre publikovanie údajov | * Prieskum v rámci OGP * Je nevyhnutné vykonávať prieskumy v rámci kreatívnych komunít a inovatívnych podnikov, o ktoré dáta by mali najväčší záujem, a podľa toho prioritizovať v pláne publikovanie. Odporúča sa tiež sledovať Globálny index otvorených dát <http://index.okfn.org/place/>, ktorý formou crowd-sourcingu vyhodnocuje rozsah, kvalitu a užitočnosť datasetov zverejňovaných naprieč verejnými správami. | * Reakcia na priority prieskumu OGP * Návrh vlastných priorít |
| * Spracovanie údajov pre publikovanie | * Metodika | Realizácia:   * Príprava prístupu k dátovým zdrojom, extrakcia dát, * Transformácia dát na otvorené datasety, prípadne linked dáta, * Popis datasetov a dátových zdrojov metadátami, * Pridelenie licencie. |
| * Publikovanie údajov na portáli. | * Metodika * Správa katalógu | Zverejnenie datasetov a ich sprístupnenie cez API a centrálny katalóg |
| * Aktualizácia publikovaných údajov. |  |  |
| * Archivácia publikovaných údajov |  |  |
| * Koordinácia spracovania otvorených údajov |  |  |
| * Správa centrálnych katalógov | * Otvorených údajov; * API; * Aplikácií nad otvorenými údajmi |  |
| * Návrh a spravovanie aplikačných rozhraní |  |  |
| * Dohľad nad kvalitou a dodržiavaním štandardov | pre validovanie schém údajov samosprávy možno použiť <http://validator.opendata.esd.org.uk/contracts>. Možno vydávať certifikáty kvality <https://certificates.theodi.org> |  |
| * Riadenie rizík | Je nevyhnutné nastaviť kompetencie a zodpovednosti za riadenie rizík a realizáciu plánovaných aktivít pre mitigovanie rizík. Pre každé riziko vedené v registri rizík musí byť definované, či je možné sa riziku vyhnúť, mitigovať ho alebo preniesť zodpovednosť na tretiu stranu napríklad prostredníctvom poistenia. Riadenie rizika však nesmie vyžadovať viac zdrojov ako očakávaná strata, spôsobená s istou pravdepodobnosťou a dopadom rizika. Preto pri výbere vhodných stratégií je nevyhnutné zobrať do úvahy aj očakávaný dopad a pravdepodobnosť rizika |  |
| * Správa URI identifikátorovov | Realizácia | Používanie |
| * Správa centrálneho modelu údajov | Realizácia | Používanie |

## Aplikačná vrstva

V kapitole je uvedený prehľad aplikačných komponentov, ich funkcií a požiadaviek na nich.

### Koncept architektúry

### Prehľad aplikačných funkcií a komponentov

Obrázok 2: Prehľad aplikačných funkcií a komponentov



### Aplikačné požiadavky

**Zoznam požiadaviek:**

* Prideľovanie URI pre akékoľvek entity ISVS je riešené centrálne prostredníctvom Centrálneho metainformačného systému
* Riešenie centralizuje katalógy pre otvorené dáta, aplikácie a API a štandardizuje popisné metadáta,
* Riešenie podporuje ontológie ako aj centrálny a ad hoc dátový model,
* Nové dátové prvky je možné navrhovať dynamicky,
* Riešenie poskytuje nástroje na čistenie dát, obohacovanie dát, prelinkovanie dát a dátové transformácie,
* Riešenie obsahuje základné štatistické a analytické nástroje (aj nad RDF úložiskom),
* Riešenie podporuje distribuovanú prácu s dátami,

**Prepojené údaje:**

* Riešenie poskytuje nástroje interaktívnych vizualizácií pre linked dáta,
* Riešenie využíva URIs ako mená vecí,
* Riešenie využíva HTTP URIs, aby si používatelia mohli dané mená vyhľadať,
* Ak si niekto vyhľadá URI, riešenie poskytne užitočné informácie s využitím štandardov (RDF, SPARQL),
* Riešenie umožní pridať linky na ďalšie URIs, aby používatelia mohli objaviť ďalšie veci a súvislosti,
* Riešenie zabezpečí jednoznačné označenie objektov a zdrojov,
* Riešenie využije štruktúru webu,
* Riešenie uľahčí získavanie informácií o menovanom objekte alebo zdroji, pričom sa tieto informácie dajú ľahko parsovať,
* Riešenie uľahčí prelinkovanie akéhokoľvek súvisiaceho objektu alebo zdroja.

### Aplikačné služby

Zoznam služieb:

* Poskytovanie raw datasetov
* Poskytovanie obohatených datasetov
* Poskytovanie Linked Data
* Extrahovanie dát z interných a externých systémov
* Spracovanie dát na otvorené dát (konverzie a transformácie, čistenie, anonymizácia, obohacovanie, prelinkovanie a mnohé ďalšie)
* Ukladanie publikovaných datasetov a metadát
* Poskytovanie datasetov a služieb prostredníctvom open API
* Vizualizovanie dát

### Aplikačné komponenty

Zoznam základných komponentov je nasledovný.

#### Centrálne komponenty

Centrálny katalóg

Obsahujúci štandardizované metadáta o položkách na [data.gov.sk](file:///C:/Users/Juraj/Downloads/data.gov.sk)

* Otvorené údaje (datasety);
* API ( bude nový komponent, ako inšpirácia slúži <https://www.data.gov/developers/apis>);
* Aplikácie nad otvorenými dátami,
* Funkcionalita pre hodnotenie datasetov, diskusie a prieskumy.

Úložisko otvorených údajov (dátový sklad otvorených údajov)

* Katalóg metadát dátových zdrojov,
* Repozitár datasetov,
* RDF úložisko,
* Archív otvorených údajov (nový komponent).

Centrálne služby pre prácu s dátami

* Blockchain“ mechanizmus:
* Služby pre prepojené údaje:
* Vizualizačné služby:
* Správa aplikačných rozhraní pre tvorbu, publikovanie a spravovanie open API:
* Register referencovateľných identifikátorov:
* Služby pre priestorové údaje: správa ontológií
* Centrálny model údajov VS

Kontajner aplikácii obsahujúci aplikácie vytvorené kreatívnymi komunitami a podnikateľmi

Podporná aplikácia pre inštitúcie verejnej správy

**Podporná aplikácia pre povinné osoby**, obsahujúce hlavne:

* Nástroje na dopytovanie a transformáciu dát (ETL nástroje),
* Publikovanie datasetu do centrálneho katalógu,
* Nástroje na vizualizáciu otvorených dát

Služby dátovej integrácie

* Služby dátovej integrácie (CSRÚ)

Komponenty na strane inštitúcie verejnej správy

* Lokálne úložisko datasetov

## Technologická vrstva

Nové a aktualizované informačné systémy budú prevádzkované vo vládnom cloude. Služby integrácie údajov budú nasedené vo forme iPaaS, pričom k dispozícii budú centrálne poskytované služby dostupné cez definované aplikačné rozhrania a podporné služby realizované formou SaaS.

# Realizácia

Kapitola sa venuje návrhu praktických krokov, ktoré zabezpečia realizáciu navrhnutej stratégie a riešenia pre Manažment údajov v praxi. Na úvod sú zhrnuté potrebné legislatívne požiadavky a zmeny, ktoré bude nutné presadiť. Nasleduje pomenovanie hlavných rizík a otvorených otázok, s ktorými bude nutné sa vysporiadať. Pomenované sú míľniky, ktoré je potrebné dosiahnuť podľa NKIVS. Nasleduje identifikácia kľúčových aktivít, ktoré je vhodné vykonať na centrálnej a lokálnej úrovni. Inštitúcie verejnej správy tak získavajú jednoduchý prehľad čo môžu v nasledujúcom období očakávať a čo by mali vykonať (vrátane požiadaviek na KRIS z pohľadu Manažmentu údajov).

## Riziká a otvorené otázky

Tabuľka 10: Prehľad rizík/otvorených otázok a spôsobov ich zmiernenia

| Riziko/otvorená otázka | Spôsob zmiernenia |
| --- | --- |
| Návrh dátového modelu bude príliš komplexný a nepoužiteľný v praxi |  |
| Nepodarí sa získať otvorenú licenciu pre kľúčové datasety zaujímavé pre inovatívne podniky ako meteorologické a geografické informácie – v tomto smere je dôležité vytvoriť a aplikovať Open Data Ready Metodiku, popísanú v časti Legislatívne požiadavky |  |
| O datasety a open API nebude záujem (verejnosti, občianskych komunít a podnikateľov) |  |
| Kvalita dát bude nízka |  |
| Ako vytvoriť efektívnu platformu s uspokojivou SLA pre zapojenie kreatívnych komunít a inovatívnych podnikateľov? |  |
| Ako zabezpečiť ochranu osobných údajov a údajov týkajúcich sa národnej bezpečnosti? |  |
| Ako získať plnú podporu zo strany autorov informačných systémov pre opakované používanie PSI pre komerčné aj nekomerčné účely? |  |
| Ako navrhnúť prioritizáciu vzniku dátových skladov, pripájanie informačných systémov do dátového skladu? |  |
| Ako získať dostatočné kapacity, schopné zverejňovať a spravovať otvorené dáta a riadiť inovačné snahy? |  |

## Plánovanie a migrácia (v súlade s akčným plánom)

Obrázok 3: Akčný plán podľa NKIVS

Kľúčové míľniky

* Otvorené údaje (do konca roka 2019).

### Aktivity a úlohy pre centrálnu úroveň

Tabuľka 11: Aktivity a úlohy pre centrálnu úroveň (ktorú reprezentuje ÚPPVII)

| Aktivita | Dátum |
| --- | --- |
|  |  |

### Aktivity a úlohy pre inštitúcie verejnej správy

Nasledujúce kroky

Tabuľka 12: Aktivity a úlohy pre inštitúcie verejnej správy

| Aktivita | Dátum |
| --- | --- |
|  |  |

### Identifikácia projektov, ktoré je vhodné realizovať

Nasledujúci zoznam predstavuje projekty, ktoré sa javia ako vhodné, aby bolo možné dosiahnuť ciele definované v tomto dokumente. Pre identifikované projekty sa očakáva, že gestor posúdi ich vhodnosť, vypracuje reformný zámer (ak je potrebné) a štúdiu uskutočniteľnosti. Rozhodnutie o realizácii projektu sa vykoná práve na základe štúdie uskutočniteľnosti.

Programový zámer: Manažment údajov

Tabuľka 13

| Projekt | Možný zdroj financovania | Aktivity | Garant |
| --- | --- | --- | --- |
|  |  |  |  |

### Požiadavky na legislatívne zmeny

Legislatíva

Zoznam požiadaviek na legislatívne zmeny, ktoré sú potrebné pre realizáciu strategickej priority (vrátane podzákonných predpisov):

| Legislatívny predpis | Legislatívna požiadavka | Garant | Termín ukončenia |
| --- | --- | --- | --- |
| Metodika „Open Data Ready“ | Ide o súbor kritérií podstatných pre možnosť publikácie otvorených dát automaticky informačnými systémami povinnej osoby. Metodika obsahuje podrobný popis týchto kritérií, postupy ich dosahovania a postup vyhodnotenia pre plnenie týchto kritérií. Kľúčovým kritériom metodiky je vysporiadanie si autorských práv údajov uložených v databázach informačných systémov s ich dodávateľom ako aj získanie súhlasu dodávateľa na zverejňovanie údajov uložených v databázach ako otvorených údajov pod licenciou CC0. Dodržiavanie metodiky Open Data Ready bude podmienkou pre čerpanie financií zo štrukturálnych fondov EÚ. |  |  |
| Nový zákon o údajoch vo verejnej správe | Nový komplexný zákon, ktorý prehľadne transponuje európsku direktívu (Directive 2003/98/EC, známa ako 'PSI Directive'). | ÚPPVII | Február 2018  Účinnosť zmien je žiadúca od Januára 2019 |
| Aktualizácia Výnosu 55/2014 o štandardoch ISVS | Upravuje výnos na základe trendov rozvoja otvorených dát, predovšetkým v oblasti linked dáta a ontológií. |  |  |
| Aktualizácia Autorského zákona | Bude obsahovať presné vymedzenie používania licencií pre PSI, pričom primárne sa odporúča používať CC0 a následne CC-BY. |  |  |

Požiadavky, ktoré by mali byť zabezpečené v rámci Operačného programu Efektívna verejná správa

* Návrh a nastavenie nových procesov pre manažment údajov vo verejnej správe a samospráve pre potreby otvorených údajov.
* Podpora zriadenia a vybudovanie kapacít Dátovej kancelárie verejnej správy pre problematiku otvorených údajov.
* Zefektívnenie komunikácie s občianskymi kreatívnymi komunitami a inovatívnymi podnikateľmi.
* Návrh opatrení pre podporu občianskych komunít a inovatívnych podnikateľov.

# Udržateľnosť a financovanie

Požiadavky na udržateľný prístup k projektom

Pre zmysluplné nastavenie projektov v rámci programu Otvorené údaje je potrebné:

Dopady navrhovaných zmien

Navrhnuté zmeny budú realizované ako program, financovaný z veľkej miery z Európskych štrukturálnych fondov. Vďaka tejto investícii dokážeme:

* Vybudovať potrebné kapacity (ľudské zdroje) a priniesť „know-how“ v oblasti Manažmentu údajov a dátovej vedy.
* Realizovať zmeny v procesoch a prijať potrebné legislatívne opatrenia.
* Implementovať potrebné nové komponenty a upraviť existujúce informačné systémy v zmysle navrhnutých princípov a požiadaviek.

Odporúčania z ohľadom na udržateľnosť

Aby riešenie dlhodobo fungovalo, odporúčame:

* Zabezpečiť dostatočné zdroje potrebné na publikovanie otvorených údajov (do X miliónov EUR ročne).
* Systematicky podporovať personálne obsadenie a kvalitu ľudského kapitálu (odborné vzdelávanie a kurzy, osobné ohodnotenie).
* Systematicky budovať „know-how“ inštitúcie (vrátane medzinárodnej spolupráce na úrovni EÚ).
* Neustále aplikovať najnovšie metódy a technológie (princípy kontinuálneho zlepšovania).

# Prílohy

Ako prílohy sú v dokumente zahrnuté:

* Skratky a definície.
* Súčasný stav pre Modul otvorených údajov.
* Návrh legislatívnych úprav.

## Skratky a definície

Definície

| Odborný termín | Definícia |
| --- | --- |
| Ontológia | Ontológia umožňuje špecifikovať koncepty a vzťahy medzi nimi, ktoré spoločne charakterizujú istú doménu záujmu, a to otvoreným a zrozumiteľným jazykom.. Napríklad v oblasti biológie môže zadefinovať tri triedy ako Zvieratá, Ryby a Cicavce, pričom Zvieratá sú nadradenou triedou. Každá ryba je tým pádom aj zvieraťom. Všeobecnou vlastnosťou triedy Zvieratá môžu byť napríklad Zmyslové orgány. Oko bude potom podriadenou vlastnosťou Zmyslových orgánov. |
| Otvorené údaje (Open Data) | Otvorené údaje (OpenData) je názov pre paradigmu, v ktorej je možné s údajmi, ktoré sú vytvárané a spracovávané v dátových zdrojoch verejnej správy, voľne pracovať. To znamená, že údaje musia byť dostupné, zrozumiteľné a práca s nimi musí byť možná bez obmedzení.  V kontexte verejnej správy to znamená nasledovné:   * Otvorené údaje môžu byť zverejňované v rôznych úrovniach kvality podľa takzvaného Tim Berners-Lee modelu označovaného ako five star model * pre každú organizáciu je dostupný zverejnený zoznam jej dátových zdrojov a ich základný popis, * obsah dátového zdroja (samotné údaje) sú proaktívne bezplatne sprístupnené neobmedzenému okruhu záujemcov, * údaje sú sprístupnené v tvare umožňujúcom ďalšie automatizované spracúvanie, čo sa dosiahne dodržaním stanovených technologických štandardov, ktorých základom je používanie otvorených a technologicky neutrálnych riešení, * ďalšiemu používaniu údajov (napr. ich spracovaniu, zverejneniu, spájaniu s inými údajmi) nie sú kladené žiadne legálne prekážky. |
| Public Sector Information (Informácie verejného sektora) | PSI sú definované v Directive 2013/37/EU, ktorá určuje jednotný trh pre informácie verejnej správy.  Informácie verejného sektora zahŕňajú akýkoľvek obsah bez ohľadu na médium (v papierovej podobe alebo elektronickej forme alebo založený na zvukových, vizuálnych alebo audiovizuálnych záznamoch) vytvorení verejným sektorom v rámci jeho agendy a kompetencií s výnimkou:   * dokumentov verejnoprávnych médií, * dokumentov vzdelávacích výskumných inštitúcií (školy, univerzity, knižnice a podobne), * dokumentov kultúrnych inštitúcií (múzea, knižnice, archívy, orchestre, opery, balety, divadlá). |
| Transformácia | Predstavuje sled transformačných predpisov reprezentujúcich celý proces spracovania otvorených dát. Predpisy umožňujú rozdelenie transformácie na spracovacie jednotky (data processing unit DPU) a ich reťazenie. Výsledkom sú spracované alebo obohatené otvorené dáta s príslušnými metadátami a prípadný záznam v katalógu. |
| Big Data | Pojem Big Data predstavuje obrovské množstvo dát rôzneho typu získavaných s vysokou rýchlosťou z vysokého počtu rôznych typov zdrojov. Big Data sa vyznačujú 4 charakteristikami: Objem, Rôznorodosť (napríklad videá, tweety, dáta zo senzorov), Rýchlosť (spracovávanie tokov dát v reálnom čase) a Hodnovernosť (s množstvom zbieraných dát sa spája neistota o ich presnosti). Spracovávanie takého množstva rôznorodých dát zbieraných v reálnom čase si vyžaduje nové nástroje a metódy ako výkonné procesory, nové softvérové riešenia a algoritmy. |
| Dataset | Ucelená a samostatne použiteľná skupina súvisiacich údajov vytvorených a udržiavaných na určitý účel a uložených spoločne podľa rovnakej schémy. |
| Dátový zdroj | Pôvodné miesto evidencie datasetu. Dátové zdroje verejnej správy sú vytvárané za účelom podpory služieb verejnej správy, služieb vo verejnom záujme alebo verejných služieb. Jednotlivé agendy súvisiace s výkonom verejnej správy môžu obsahovať viaceré dátové zdroje. Charakteristickým znakom dátového zdroja je jeho samostatná použiteľnosť (výpovedná hodnota) nezávisle od iných dátových zdrojov. Ďalšími znakmi popisujúcimi dátový zdroj sú názov, účel (zameranie), typy spracúvaných údajov a vzťahy medzi nimi, formát údajov a podobne – tzv. metadáta. Dátovým zdrojom môžu byť aj zariadenia produkujúce dáta v internete vecí, senzory alebo aplikačné rozhrania sociálnych sietí. |
| Metadáta | Štruktúrované údaje obsahujúce informácie o primárnych údajoch – dátovom zdroji, pričom primárne údaje spravidla reprezentujú určitý hmotný objekt alebo nehmotný objekt. sú určené najmä na vyhľadávanie, katalogizáciu a využívanie primárnych údajov. |
| Mikrodáta | V kontexte štatistiky predstavujú mikrodáta informáciu na úrovni individuálnych respondentov, ktorí sa zúčastňujú prieskumov alebo sčítaní ľudu. |
| Paradáta | Paradáta sú dáta dokumentujúce proces zberu údajov. V prípade vykonania prieskumu môže ísť o informácie zahŕňajúce časy, kedy sa vykonávali pohovory, dĺžku trvania pohovorov, spôsob komunikácie, ochotu respondenta a mnohé ďalšie. |
| Prístup cez aplikačné rozhranie - API | Štandardizovaný spôsob vzájomnej strojovej komunikácie, charakteristický určením koncového bodu poskytovania služby, množinou funkcií v tomto bode dostupných, ich parametrami a formátmi vymieňaných správ. |
| Jednotný referencovateľný identifikátor | Jednotný referencovateľný identifikátor (URI) je množina znakov tvoriacich jednotným spôsobom identifikátor ľubovoľných logických a fyzických objektov. |
| Surové dáta | Surové dáta, tiež nazývané primárne dáta, sú dáta získané priamo z dátového zdroja, ktoré nie sú spracované na použitie pre určitý konkrétny účel (ešte neprešli spracovaním ani žiadnou inou manipuláciou). |

Skratky

| Skratka | Popis |
| --- | --- |
| BIA | Hodnotenie vplyvov regulácie na podnikateľské prostredie (Business impact assessment) |
| CDS | Centrálny dátový sklad verejnej správy |
| CKS | Centrálny konsolidačný systém |
| CSRÚ | Centrálna správa referenčných údajov |
| EIA | Hodnotenie vplyvov na životné prostredie (Environmental impact assessment) |
| ETL | (Extract, Transform, Load) |
| FIA | Hodnotenie vplyvov na rozpočet (Financial impact assessment) |
| IČ DPH | Identifikačné číslo pre daň z pridanej hodnoty |
| IAM | Systém na správu identít a oprávnení (Identity and Access Management) |
| IFP | Inštitút finančnej politiky |
| IKT | Informačno-komunikačné technológie |
| ISVS | Informačný systém verejnej správy |
| IVP | Inštitút vzdelávacej politiky |
| KDP | Katalóg dátových prvkov |
| KPI | Kľúčový výkonnostný ukazovateľ (Key performance indicator) |
| MDM | Správa kmeňových údajov (Master data management) |
| MetaIS | Metainformačný systém |
| MSP | Malí a strední podnikatelia |
| MUK | Modul úradnej komunikácie |
| OP EVS | Operačný program Efektívna verejná správa |
| OP II | Operačný program Integrovaná infraštruktúra |
| OVM | Orgán verejnej moci |
| PgK | Programová kancelária |
| RA | Register adries |
| RIA | Hodnotenie vplyvov regulácie (Regulatory impact assessment) |
| RFO | Register fyzických osôb |
| RPI | Register priestorových informácií |
| RPO | Register právnických osôb |
| SIA | Hodnotenie sociálnych vplyvov (Social impact assessment) |
| SSO | Jednotné prihlásenie (Single Sign-On) |
| URI | Jednotný referencovateľný identifikátor (Uniform resource identifier) |
| ÚPPVII | Úrad podpredsedu vlády SR pre informatizáciu a investície |

## Návrh legislatívnych úprav

1. Aknowledgement to the European Public Sector Information (PSI) Platform [↑](#footnote-ref-2)
2. Skutočne potrebné kapacity sa môžu líšiť na základe rozsahu agendy jednotlivých inštitúcií verejnej správy (vo vzťahu k spracovaniu údajov) [↑](#footnote-ref-3)