



Strategická priorita **Manažment údajov**

(Verzia 1-0)



Informácia o dokumente

Názov:	Strategická priorita Manažment údajov
Stav:	Finálna verzia
Pripravil:	Pracovná skupina K9.4 Lepšie dáta
Verzia:	1.0
Dátum:	14.12.2016
Pripomienkoval:	
Dátum revízie:	20.2.2017

Distribučný zoznam

Od	Dátum	Kontakt
K9.4		

História verzií

Verzia	Dátum verzie	Pripravil/ Zmenil	Pripomienkoval	Kľúčové zmeny
0.2	14.12.2016			Prvá pracovná verzia
0.3	18.1.2017			Druhá pracovná verzia
0.9	7.2.2017			Verejná verzia
1.0	20.2.2017			Finálna verzia



Obsah

1	Úvodný pohľad	4
1.1	Cieľ dokumentu	4
1.2	Definícia strategickej priority	4
1.2.1	Kontext pre strategickú prioritu Manažment údajov	4
1.2.2	Vysvetlenie strategickej priority	4
1.2.3	Základné úlohy	7
1.2.4	Základné témy	7
1.2.5	Predstavenie konceptu	8
1.3	Analýza súčasného stavu	9
1.3.2	Riziká vyplývajúce zo súčasného stavu	13
1.3.3	Problémy, ktoré je potrebné vyriešiť	14
2	Ciele realizácie	16
2.1	Architektonické ciele	16
2.1.1	Kvalita údajov	16
2.1.2	Referenčné údaje a „jedenkrát a dost“	18
2.1.3	Moje dáta	19
2.1.4	Analytické spracovanie údajov vo verejnej správe	19
2.1.5	Prepojené dáta (Linked Data)	19
2.2	Dotknuté subjekty a ich záujmy	20
3	Organizácia	22
3.1	Zodpovednosť	22
3.2	Posúdenie kapacitnej pripravenosti	23
3.2.1	Personálna kapacita	23
3.2.2	Kapacita informačných systémov	23
3.3	Organizačné zmeny	23
4	Stratégia	26
4.1	SWOT analýza	26
4.2	Vízia	26
4.3	Výber prístupu	27
4.3.1	Kategorizácia a klasifikácia údajov	27
4.3.2	Kvalita údajov	31
4.3.3	Integrácia údajov	33
4.3.4	Moje dáta a ochrana osobných údajov (GDPR)	35
4.3.5	Analytické spracovanie údajov vo verejnej správe	36
4.3.6	Manažment prístupov a oprávnení	38
5	Architektúra riešenia	42
5.1	Princípy	42
5.2	Biznis vrstva	44
5.2.1	Prehľad požiadaviek	44
5.2.2	Prehľad procesov	44
5.3	Aplikačná vrstva	51



5.3.1	Koncept architektúry integračnej platformy	51
5.3.2	Prehľad aplikačných funkcií a komponentov	53
5.4	Technologická vrstva	56
6	Realizácia	57
6.1	Riziká a otvorené otázky	57
6.2	Plánovanie a migrácia (v súlade s akčným plánom)	58
6.2.1	Aktivity a úlohy pre centrálnu úroveň	59
6.2.2	Aktivity a úlohy pre inštitúcie verejnej správy	60
6.2.3	Identifikácia projektov, ktoré je vhodné realizovať	62
6.2.4	Požiadavky na legislatívne zmeny	65
7	Udržateľnosť a financovanie	68
8	Prílohy	70
8.1	Skratky a definície	70
8.2	Kandidáti na referenčné údaje.	72
8.3	Klasifikácia a kategorizácia údajov	74
8.4	Súčasný stav pre Modul úradnej komunikácie – dátová časť ostatných úsekov verejnej správy	78
8.5	Posudzovanie variantov a výber strategického prístupu	80
8.5.1	Referenčné údaje a „jedenkrát a dost“	80
8.5.2	Kvalita údajov	82
8.5.3	Integrácia údajov	83
8.5.4	Moje dáta a ochrana osobných údajov	86
8.5.5	Analytické spracovanie údajov vo verejnej správe	86

1 Úvodný pohľad

Nasledujúci dokument predstavuje pracovnú verziu Strategickkej priority Manažment údajov. 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. Dokument je pripravovaný na rokovanie Rady vlády Slovenskej republiky pre digitalizáciu verejnej správy a jednotný digitálny trh.

1.1 Cieľ dokumentu

Cieľom tohto dokumentu je podrobnejšie rozpracovanie strategickkej priority 6 Manažment údajov: Definovanie základných cieľov v rámci strategickkej priority, načrtnutie možných alternatív ich dosiahnutia ako aj výber najvhodnejšej alternatívy na realizáciu. Definované ciele musia priniesť pozitívny vplyv na zlepšenie dát vo verejnej správe. A to ako po stránke kvality, kvantity, tak aj pre zabezpečenie základných princípov definovaných v Národnej koncepcii informatizácie verejnej správy.

1.2 Definícia strategickkej priority

1.2.1 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,
- 6 Riadenie údajov a Big data (Manažment údajov),**
- 7 Otvorené údaje,
- 8 Vládny cloud,
- 9 Komunikačná infraštruktúra,
- 10 Kybernetická bezpečnosť.

NKIVS ku každej strategickkej prioritě informatizácie verejnej správy vysvetľuje jej cieľ, prístup k riešeniu a tiež rámcový architektonický model. Tento dokument nadväzuje na NKIVS a predstavuje návrh riešenia pre Manažment údajov vo verejnej správe a rozširuje tak kapitolu **6.2.6 Riadenie údajov a Big data**. Tento dokument ešte môže byť z pozície architektonickej kancelárie verejnej správy upravovaný a dopĺňaný aj na základe pripomienok a komunikácie s gestorm tejto strategickkej priority.

Strategická priorita Manažment údajov 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.

1.2.2 Vysvetlenie strategickkej priority

Výrazne lepšie využívanie dát vo verejnej správe predstavuje kľúčový cieľ programového obdobia 2014 až 2020. K dátam sa bude pristupovať ako k vzácnemu zdroju. Dátovej vrstve bude preto v architektúre venovaná výrazná pozornosť. Hlavným cieľom bude dátová integrácia medzi jednotlivými systémami

verejnej správy a zabezpečenie pravidelnej replikácie kvalitných a konsolidovaných transakčných dát do dátových skladov.

Nástup „Big data“ umožní verejnú správu fungujúcu na využívaní znalostí

V oblasti informačných technológií dochádza k nástupu „Big data“ – kapacity zbierať, spracovávať a analyzovať obrovské množstvá údajov nielen dávkovo, ale aj v reálnom čase. Tento fenomén transformoval mnohé oblasti ekonomiky a z revolúcie informačných technológií sa stáva revolúcia v oblasti využívania dát. V súkromnej sfére sa považuje správne využívanie analytických nástrojov za kľúčovú konkurenčnú výhodu. Vo verejnej sfére vzniká výrazný potenciál zlepšiť kvalitu politík a regulácií, ako i operatívneho rozhodovania, lepšie manažovať riziká a byť schopný flexibilnejšie reagovať, napríklad na preferencie občanov. Strategickým zámerom Slovenska je preto lepšie využívanie analytických nástrojov na podporu rozhodovania: od sledovania výkonnosti jednotlivých sektorov verejnej správy, cez vyhodnocovanie dopadov politík a regulácií, analýzu rizík po vyhodnocovanie investícií. Jedna z najväčších výziev bude transformácia kultúry inštitúcií smerom k rozhodovaniu na základe dát, dôkazov a faktov. Bude potrebné vybudovať potrebné kapacity, transformovať procesy fungovania verejnej správy a zaviesť tak riadenie verejnej správy založené na dôkazoch. Kľúčovou otázkou je, ktoré dáta je potrebné zbierať a na základe akých výstupov sa má verejná správa rozhodovať. Lepšie dáta znamenajú možnosť získavať kvalitné informácie, z nich vyplývajúce „insights“, ktoré zas slúžia ako podklady pre tvorbu znalostí a lepšie rozhodovanie. Ak sa bude vo verejnej správe lepšie rozhodovať, pozitívne sa to prejaví na výsledkoch vládnutia a stave jednotlivých sektorov verejnej správy. Je preto potrebné podrobne rozumieť logike (ontológiám) spravovaných dát, vytvárať a udržiavať dátové modely, katalogizovať metadáta a paradáta, chápať prepojenia medzi jednotlivými systémami a podporovať zvyšovanie kvality dát ukladaných v databázach.

Koncept predstavený v tejto strategickej priorite predstavuje aplikáciu myšlienky „Data-driven state“ – štáte, fungujúcom na základe využívania dát a presnom riadení celého životného cyklu dát. Podpora využívania znalostí vo verejnej správe je základným opatrením Operačného programu Integrovaná infraštruktúra pre špecifický cieľ „Umožnenie modernizácie a racionalizácie verejnej správy IKT prostriedkami“.

Ďalším dôležitým aspektom problematiky manažmentu dát je sprístupniť všetky potrebné údaje analytického alebo evidenčného charakteru, vďaka ktorým budú môcť zamestnanci poskytujúci verejné služby vykonávať svoju prácu efektívne. Údaje evidenčného charakteru ako referenčné údaje a transakčné údaje sú ukladané v agendových informačných systémoch a budú centrálne prístupné cez platformu integrácie údajov. V informačnom prostredí verejnej správy tak uplatníme princípy pre pohodlný a kontextovo-orientovaný prístup k dátam uloženým v iných systémoch, ako i definovať pravidlá, na základe ktorých informačné systémy prístupujú k dátovým zdrojom iných systémov. Riadenie dát bude tiež zabezpečovať integritu a kvalitu údajov.

Nárast používania dát zvyšuje tiež tlak na zabezpečenie súkromia a ochranu osobných údajov občanov a podnikateľských subjektov. Informačné prostredie verejnej správy musí občanom a podnikateľským subjektom zabezpečiť transparentný pohľad na využívanie údajov o nich. Na jednom mieste bude pre občana a podnikateľský subjekt a dostupný prehľad o všetkých dátových objektoch, ktoré verejná správa vo svojich informačných systémoch eviduje. Zároveň bude môcť občan a podnikateľský subjekt vidieť, kedy a kto údaje menil alebo k nim pristupoval (Pravidlá sa nastavujú podľa klasifikácie). Ako pilot pre vybrané prípady použitia sa tiež nasadí technológia pre Systémy na manažovanie osobných údajov („Personal Information Management Systems“), ktoré umožňujú ukladať osobné údaje podnikateľov a občanov do takzvaných osobných dátových úložísk. Občan alebo podnikateľský subjekt získa tak ešte väčšiu kontrolu nad svojim súkromím ako aj pohodlné nástroje na opravu alebo vymazanie údajov a udelenie súhlasu so spracovaním údajov (v prípadoch, keď je to vhodné). Pre občanov a podnikateľov sa v budúcnosti stanú ich osobné údaje ekonomickým aktívom.

Súvisiacou strategickou prioritou sú Otvorené údaje, v ktorej sa navrhne, ako sa údaje, ktoré má verejná správa k dispozícii, publikujú verejnosti vo formátoch vhodných na opakované použitie. Pre tému otvorených údajov pripravený samostatný dokument, v rámci ktorého sa okrem iného bude riešiť i téma

licenčných pravidiel, v súlade s ktorými budú údaje ako elementárne súčasti informačných zdrojov, databáz prípadne iných diel) vytvárané a následne sprístupňované. Ambíciou je zadať „štátnu licenciu“¹ alebo stanoviť licenčné pravidlá pre vytvárané údaje vo verejnej správe.

Prepojené údaje a sémantický web

Základom efektívneho zdieľania a využívania údajov je používanie princípov prepojených údajov (takzvané Linked Data). Je to technológia takzvaného Sémantického webu (web dát, nie web dokumentov), ktorý predstavuje nový generačný stupeň vývoja webu (označovaný ako Web 3.0). V Sémantickom webe predstavujú webové aplikácie regulárne znalostné systémy (podpora strojového spracovania dát s ohľadom na význam, odvodzovanie, dokazovanie), no najmä, publikované dáta predstavujú regulárny znalostný graf, čo umožňuje spájanie (linkovanie) dát naprieč rôznymi zdrojmi ako napr. databázami, resp. webovými sídlami. Základným stavebným kameňom je jazyk RDF (Resource Description Framework), ktorý poskytuje zmienenú grafovú formu reprezentácie dát, to je množinu vrcholov a hrán spájajúcich tieto vrcholy. Vrcholy predstavujú ľubovoľné entity ako *osoby*, *organizácie*, *priestorové objekty*, pričom hrany zas reprezentujú jednotlivé vzťahy medzi týmito entitami, ako napríklad *býva*, *vlastní* a podobne. Kľúčovým prvkom v Linked Data je používanie Jednotných referencovateľných identifikátorov (URI) na identifikovanie jednak uzlov, a súčasne i hrán. Jednotným referencovateľným identifikátorom je množina znakov tvoriacich jednotným spôsobom identifikátor ľubovoľných logických a fyzických objektov. Množina prvkov ktorá opisuje nejakú doménu (napríklad organizácia má sídlo, názov a podobne) sa nazýva ontológia. Používaním štandardizovaných ontológií dochádza k takzvanej integrácii v dátach, pretože pokiaľ ktorýkoľvek subjekt zverejní svoje dáta popísané štandardizovanou ontológiou, tak tieto dáta je možné okamžite spracovať. Vo všeobecnosti, grafovo reprezentované dáta majú nesporné výhody, a pri spojení so strojovým spracovaním dát s ohľadom na ich význam už predstavujú mimoriadnu pridanú hodnotu pri spracovávaní údajov. Napríklad prístup stotožňovania entít prostredníctvom sémantiky je veľmi prínosný, nakoľko sémantické databázy implementujú stotožňovanie na matematickej úrovni presnosti, pričom od okamihu prepojenia dvoch entít už databáza považuje uvedené entity ako jednu, a stotožňovanie nie je nutné samostatne implementovať.

Ako už bolo spomenuté v predchádzajúcej kapitole, jedna z najväčších výziev bude transformácia kultúry inštitúcií smerom k rozhodovaniu na základe údajov, dôkazov a faktov. Pre splnenie tohto cieľa je použitie prístupu Linked Data mimoriadne vhodné. Nielen že Linked Data je nová technológia spracovávania dát na úrovni logického usudzovania, ale hlavne Linked Data ponúka efektívnu grafovú formu reprezentácie dát, to je "všetko je poprepájané so všetkým", čo je kľúčové pri potrebe zmiešania ľubovoľných dát z rôznych zdrojov a následným ľubovoľným medzi-doménovým komplexným analýzám údajov. Takéto riešenie ponúka priamu podporu rozhodovania inštitúcii na základe relevantných (najrelevantnejších) faktov. Tieto aj iné skutočnosti sú dôvodom prečo existuje veľká snaha používať princípy Sémantického webu aj vo verejných dátach, čo je označované ako Linked Government Data. V rámci Európskej únie ide o program ISA2 (Interoperability Solutions for European Public Administrations), konkrétne o skupinu SEMIC (Semantic Interoperability Community), ktorá napríklad odporúča použitie štandardizovaných ontológií (modelov) fyzickej osoby, organizácie, lokácie a ďalšie. V Slovenskej republike prebieha od roku 2013 takzvaná. Fáza 1 implementácie sémantiky (Linked Data) do štátnych dátových štandardov v podobe viacerých predložených návrhov, pričom medzi hlavné patrí:

- metodika tvorby data.gov.sk URI identifikátorov,
- celkový jednotný model údajov verejnej správy: data.gov.sk ontológie,
- metóda registrácie URI identifikátorov pre základnú sémantickú infraštruktúru údajov verejnej správy (katalógy, číselníky, dátové prvky, referenčné údaje) v MetaIS,

¹ Napríklad po vzore Veľkej Británie (Open Government License <https://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>)



- publikácia množiny otvorených dát obsahujúcich výlučne Linked Data, nazývaná LOD Slovakia (Linked Open Data Cloud Slovakia).

1.2.3 Základné úlohy

Strategická priorita Manažment údajov ponúka návod na lepšie využívanie dát vo verejnej správe a podporuje prístup k dátam ako k vzácnemu zdroju:

- Podporí sa celý životný cyklus údajov s dôrazom na riadenie kvality údajov a prístupov k údajom na základe rolí a oprávnení (vrátane životného cyklu meta-údajov, ako sú napríklad štandardizované číselníky).
- Navrhne sa vhodná kategorizácia údajov.
- Zavedú sa mechanizmy pre integráciu údajov - zavedú sa pravidlá a nástroje, pomocou ktorých budú informačné systémy pristupovať k dátovým zdrojom iného informačného systému.
- Navrhne sa systém účelného využívania dát vo verejnej správe a mechanizmus riadenia oprávnení.
- Údaje budú systematicky využívané na zlepšenie rozhodovania vo verejnej správe vďaka ich analytickému spracovaniu.
- Posilnia sa práva občanov a podnikateľských subjektov v súvislosti s používaním dát, ktoré sa ich týkajú.
- Vyberú sa vhodné štandardy pre oblasť manažmentu a interoperability údajov (štruktúra a formáty, v akých budú dáta ukladané a komunikované, ontologický model) a podporí sa tak zavedenie prepojených údajov (Linked Data).

1.2.4 Základné témy

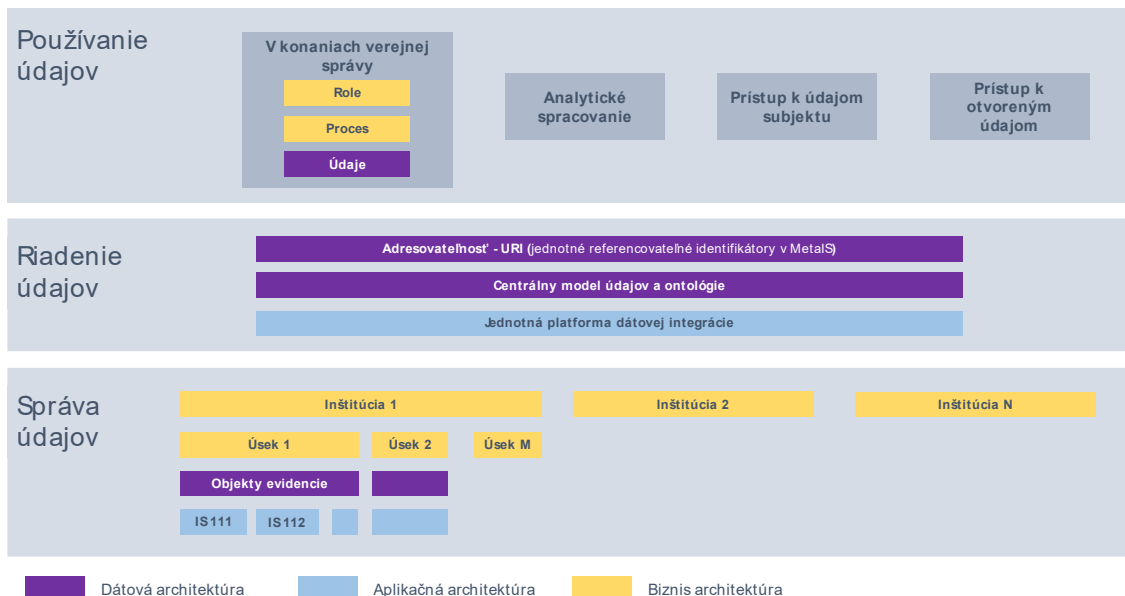
Dokument prináša odpovede v nasledujúcich témach:

- Manažment životného cyklu údajov.
- Kvalita údajov.
- Klasifikácia a kategorizácia údajov.
- Integrácia údajov.
- Prepojené údaje (Linked Data).
- Referenčné údaje a princíp jedenkrát a dosť.
- Priestorové údaje.
- Osobné údaje (služba Moje dáta).
- Analytické spracovanie údajov vo verejnej správe.

1.2.5 Predstavenie konceptu

Obrázok 1: Vysokoúrovňový pohľad na manažmentu údajov

Koncept: Manažment údajov



Údaje vo verejnej správe budú manažované systematicky. Znamená to, že vzniknú jasné pravidlá a metodiky pre Používanie údajov; Riadenie údajov a Samotnú správu údajov. Každá z týchto oblastí bude v cieľovom stave modelovaná v príslušnej vrstve architektúry (dátová architektúra, aplikačná architektúra, biznis architektúra). Jednotlivé pravidlá a metodiky musia zahŕňať aj definície jednotlivých dátových prvkov, ich použitie, spôsob odvodzovania údajov a spôsob interpretácie.

Používanie údajov

Z pohľadu používania údajov navrhované opatrenia zabezpečia:

- aby v konaniach verejnej správy boli k dispozícii všetky údaje, ktoré sú potrebné. Znamená to, že budú pre jednotlivé “Role” v “Procesoch” identifikované údaje (“Dátové objekty”) z informačných systémov verejnej správy, ku ktorým bude garantovaný prístup. Referenčné údaje budú získavané z rôznych zdrojových informačných systémov bez prekážok, čo zabezpečí realizáciu princípu “jedenkrát a dost”. Zároveň je potrebné, aby bola preukázateľne zaznamenávaná platnosť referenčných údajov v danom čase,
- aby verejná správa dokázala využívať svoje údaje pre potreby prípravy analýz (analytické spracovanie údajov), ktoré budú slúžiť ako podklad pre lepšie rozhodovanie. Počíta sa s využitím virtuálnych databáz a nástrojov pre anonymizáciu, pseudoanonymizáciu (ktoré zabezpečia, že údaje osobného charakteru bude možné používať na analytické účely bez narušenia súkromia),
- aby verejnosť mala prístup k údajom verejnej správy v otvorenom formáte vhodnom na opätovné použitie – okrem osobných údajov, citlivých údajov a utajovaných údajov (podrobnejšie sa tejto téme bude venovať nasledujúci dokument Strategická priorita Otvorené údaje),
- aby každý subjekt (občan aj podnikateľský subjekt) získal prístup k údajom, ktoré verejná správa o ňom eviduje, mohol s nimi v rámci možností manipulovať a zároveň vidieť, kto a prečo k takýmto údajom pristupoval. Manipuláciou sa rozumie najmä určovanie rozsahu zdieľania, či pristupovanie k službám súvisiacich s danými údajmi. Takúto funkcionality bude mať občan a podnikateľský subjekt dostupnú cez pripravovanú službu „Moje dáta“ (vrátane rozhrania OpenAPI).

Riadenie údajov

Riadenie údajov bude riešené centrálné, v kompetencii Úradu podpredsedu vlády SR pre informatizáciu a investície. Aby mohla byť naplnená vízia štátu fungujúcom na využívaní znalostí, sú v tomto dokumente navrhnuté centrálné riešenia, ako:

- jednotná platforma integrácie údajov (vrátane funkcií a aplikácií pre zdieľanie referenčných údajov, riadenie prístupov a oprávnení, riadenie kmeňových údajov, riadenie kvality, riadenie prístupu k osobným údajom, riešenie incidentov s údajmi a riadenie životného cyklu meta-údajov),
- jednotný centrálny model údajov verejnej správy reprezentovaný ontológiami,
- jednotná forma unikátnych identifikátorov ľubovoľných dátových údajov pre ich adresovateľnosť - každý údaj by mal byť jednoznačne adresovateľný cez URI, pričom tieto URI prideluje metainformačný systém, ktorý stráži celkovú integritu referenčných identifikátorov,
- jednotné analytické nástroje.

Centrálna úroveň tak zodpovedá za dostupnosť údajov z jednotlivých informačných systémov verejnej správy.

Správa údajov

Za správu údajov a ich kvalitu zodpovedá príslušná inštitúcia verejnej správy. V celkovom modeli údajov by mali byť zastúpené všetky relevantné inštitúcie, ktoré spravujú vlastné údaje. Jednotlivé inštitúcie sú členené na úseky verejnej správy (evidované v metainformačnom systéme verejnej správy). Každý úsek sa venuje konkrétnej kompetencii a agendám, preto na úseku verejnej správy môže byť prevádzkovaných viacero agendových informačných systémov. Pre úspešné vyriešenie otázky údajov bude na každom úseku potrebné zabezpečiť, aby:

- relevantné objekty evidencie v správe úseku boli identifikované,
- kmeňové a agendové údaje boli riadne spravované v informačnom systéme (na základe princípov pre registre),
- relevantné údaje v správe úseku boli vyhlásené ako referenčné (a tým pádom evidované v centrálnom modeli údajov a dostupné cez platformu integrácie údajov),
- relevantné údaje v informačných systémoch verejnej správy boli stotožnené s vyhlásenými referenčnými údajmi.

1.3 Analýza súčasného stavu

Momentálne sa nachádzame na začiatku procesu lepšieho využívania dát vo verejnej správe. Prijala sa legislatíva, ktorá definuje referenčné údaje a možnosti pre ich zdieľanie a vytvorené boli úvodné technologické predpoklady pre dátovú integráciu a zvyšovanie kvality údajov. Súčasný stav je tak možné charakterizovať:

- Začína sa s vyhlasovaním referenčných registrov (aktuálne sú vyhlásené 4 registre²).
- Začína sa s procesom zdieľania údajov, kde hlavným motivátorom je ročná lehota na zavedenie referencovania údajov.
- Začína sa s procesom čistenia údajov, kde motivátorom je proces stotožnenia subjektu evidencie ako nutná podmienka zdieľania dát.
- V súčasnosti v oblasti Linked Data prebieha prvá fáza implementácie sémantiky do štandardov ISVS. A to najmä štandardizácia metódy tvorby URI identifikátorov, vytvorením kľúčových ontológií (v rozsahu katalógu dátových prvkov) spolu so zaregistrovaním všetkých potrebných URI identifikátorov.

² Zdroj: <https://metais.finance.gov.sk/refregisters/list?page=1&count=20>

1.3.1.1 Životný cyklus údajov

Životný cyklus údajov vo verejnej správe nie je systematicky manažovaný

V súčasnosti každý orgán verejnej správy zbiera a spracováva svoje vlastné údaje. Údaje sú ukladané v rôznych IS, databázach s rôznymi požiadavkami na kvalitu, dostupnosť, bezpečnosť. Dáta, veľa krát redundantné, líšia sa formátom, štruktúrou, definíciou a kvalitou. Neexistuje komplexný dátový model verejnej správy, nie sú informácie, aké dáta, akým spôsobom a na aké účely sú využívané. Chýbajú interaktívne katalógy číselníkov, tried prvkov(dátových prvkov)³, slovníkov⁴, metadát a paradát. Výmena potrebných dát medzi jednotlivými orgánmi verejnej správy je vo veľkej miere riešená spôsobom „peer to peer“, teda „každý s každým“. Výnimkou sú orgány verejnej moci (pilotný rozsah 12 orgánov verejnej moci) zapojené do dátovej časti ostatných úsekov Modulu úradnej komunikácie. Pre účely manažmentu modelu údajov bol vytvorený metainformačný systém (ďalej aj MetaIS), ktorý eviduje základné číselníky.

Dátové štandardy sú neúplné, nekonzistentné nepodporujúce Linked Data

Pre popis dát je dnes jedným z hlavných nástrojov KDP - Katalóg dátových prvkov. KDP popisuje základné dátové štruktúry ako Fyzická osoba, Právnická osoba ďalšie entity. Problémom KDP je, že jednak pokrýva len malú oblasť verejných dát, ale oveľa horšie je, že KDP je z hľadiska formálnej reprezentácie znalostí predstavuje takzvanú slabú ontológiu, z čoho jednak vyplýva nízka výrazová sila pre opis danej domény, no najmä strojové spracovanie takýchto dát s ohľadom na ich význam (odvodzovanie nových znalostí, dokazovanie) je viac než obmedzené.

Dnešné možnosti integrácie údajov

Zdieľanie údajov medzi inštitúciami verejnej správy je v princípe realizované dvoma spôsobmi:

- „Každý s každým“ rôznymi technológiami od mailových príloh až po API. OPIS projekty sa riadili metodikou usmernenia 14/2014. Programová kancelária (PgK) OPIS monitorovala maticu asi 50 x 50 integrácií. Vybrané rezorty majú nasadené integračné platformy (napríklad „ekosystém“ riadenia verejných financií MF), ktoré manažujú integráciu a redukujú počet integračných väzieb ale väčšinou len proxujú API, čiže na aplikačnej vrstve sa konzument musí prispôbovať API každého poskytovateľa údajov.
- Prostredníctvom MUK dátovej časti, ktorá ako spoločný modul zabezpečuje poskytovanie dát jednotným spôsobom. Ako produkčný bol zatiaľ oficiálne vyhlásený MUK dátová časť ostatných úsekov (november 2015). Poskytovanie dát jednotným spôsobom znamená, že konzument získava údaje štandardizovanými službami a teda sa nemusí prispôbovať API rôznych poskytovateľov. Naopak MUK sa snaží maximálne prispôsobiť poskytovateľovi údajov a jeho možnostiam, prípadne umožňuje využiť perzistentnú vrstvu použitím zápisovej služby. Súčasná legislatíva prikazuje orgánom verejnej moci použiť MUK pri prístupe k referenčným údajom a oprávňuje použiť MUK pri prístupe k iným dátam.
- Prostredníctvom výmeny správ cez MUK komunikačná časť.

Kvalita údajov nie je monitorovaná a vyhodnocovaná

Všeobecne platné závery o kvalite údajov vo verejnej správe nie je možné vyniesť. Bola pripravovaná metodika pre dátovú kvalitu⁵, jej používanie však nebolo oficiálne zavedené do praxe a navrhnuté kritériá kvality údajov sa systematicky nevyhodnocujú.

³ Napríklad <http://registre.enviroportal.sk/geo/>

⁴ Napríklad <https://www.w3.org/2015/03/inspire/>

⁵ Dokument Metodika riadenia kvality údajov vytvorená v rámci projektu IS Centrálnej Správy Referenčných Údajov, ktorý implementuje služby MUK dátovej časti ostatných úsekov

Podstata nekvality údajov pri osobách je izolovaný spôsob aktualizácie údajov v jednotlivých registroch. Každý informačný systém má svoj jednoznačný interný a vlastný identifikátor pre subjekt evidencie. Kvalita ostatných údajov je neznámou veličinou, pričom sa dá predpokladať, že vzhľadom na absenciu systematického prístupu nebude vysoká.

Proces vyhlasovania referenčných údajov bol spustený

V priebehu rokov 2015 a 2016 boli ako referenčné údaje vyhlásené prvé údaje (ako napríklad IČ DPH, dátum zaradenia uchádzača do evidencie uchádzačov o zamestnanie). Kľúčovou udalosťou bolo vyhlásenie údajov z RPO, vďaka ktorému je možné stotožňovať údaje o právnických osobách. Prechodné obdobie na stotožnenie je jeden rok od vyhlásenia referenčných údajov.

Dopyt po vyhlasovaní referenčných údajov je veľký a začínajú sa prejavovať platformové efekty. V súčasnosti boli vyhlásené referenčné údaje z RFO (7.12.2016) a nasledovať budú údaje z RA. Je dôležité uviesť, že všetky vyhlásené referenčné údaje sú vyhlásené spolu s ich URI, nakoľko sú súčasťou už vytvorených zdrojových ontológií a tieto URI sú registrované v Metainformačnom systéme.

Podrobnejšie informácie o referenčných údajoch sú k dispozícii v kapitole 4.3.1.1. Referenčné údaje a jedenkrát a dost.

„Jedenkrát a dost“

Princíp „jedenkrát a dost“ zatiaľ v praxi nefunguje úplne, napriek príprave legislatívnych a technických predpokladov. Medzi základné limity patrí aj nesúlad medzi generickou legislatívou (v tomto prípade Zákon o e-Governmente) a špecifickými zákonmi upravujúcimi fungovanie jednotlivých agend. Jednotlivé subjekty verejnej správy preto však nepracujú s aktuálnymi a rovnakými údajmi, alebo opakovane vyžadujú údaje už existujúce v inom systéme (princíp „jedenkrát a dost“ sa porušuje), prípadne riešia svoje potreby externých údajov pomocou dvojstranných neharmonizovaných dohôd o poskytovaní údajov s realizáciou individuálnych technických riešení.

Osobné údaje a služba moje dáta

K referenčným údajom, ktoré sa týkajú konkrétnej fyzickej osoby bude možné prístupit' cez službu „Moje dáta“, ktorá má byť dostupná cez elektronickú schránku. Služba momentálne nemôže poskytovať dostatočný rozsah údajov, nakoľko je potrebná úprava ISVS, ktorý bude cez predmetnú službu údaje poskytovať.

Implementácia dôsledkov smernice GDPR nie je momentálne vo verejnej správe systematicky riešená.

Analytické spracovanie údajov vo verejnej správe

Analytické jednotky sú v súčasnosti zriadené len na niektorých ministerstvách a ich výkonnosť je až na výnimky (IFP, čiastočne IVP) obmedzená, čo je dané hlavne personálnymi kapacitami a kvalitou dostupných nástrojov a dát. Metodika posudzovania vplyvov sa dodržiava formálne a analýzy nie sú podrobované kontrole kvality. Vďaka iniciatívam reformy verejnej správy vyjadrených v Operačnom programe Efektívna verejná správa dochádza postupne k zmene tohto stavu a plánuje sa posilnenie agendy lepšej regulácie a vybudovanie analytických jednotiek.

Údaje vo verejnej správe sa väčšinou nezberajú pravidelne a väčšinou nie sú konzistentné a kvalitné. Zdieľanie dát je administratívne a technicky náročné (používa sa výmena dát na pevných nosičoch ako DVD). Neexistuje predstava o dátovom modeli verejnej správy: aké dáta majú jednotlivé inštitúcie verejnej správy k dispozícii a na aké účely sa používajú. Nie sú tiež jasne vyjasnené pravidlá na ochranu osobných údajov a zavedené spoľahlivé nástroje pre ich anonymizáciu a pseudonymizáciu.

1.3.1.2 Priestorové (geografické) údaje

Priestorový objekt predstavuje abstraktné znázornenie informácie o objekte reálneho sveta súvisiacej s konkrétnou polohou alebo geografickou oblasťou v geodetických systémoch. V tejto súvislosti je každý údaj o priestorovom objekte, s priamym alebo nepriamym odkazom na jeho konkrétnu polohu alebo geografickú oblasť, priestorovým údajom.

Z praktického hľadiska je každý údaj, ktorý obsahuje informáciu o priamej alebo nepriamej lokalizácii objektu a javu, priestorovým údajom. Priamy odkaz na polohu môže byť vyjadrený formou súradníc (2D alebo 3D) a nepriamym odkazom na polohu môže byť napríklad adresa alebo uvedenie blízkych vzťahov.

Referenčné priestorové údaje

Vo všeobecnosti sú priestorové údaje svojou povahou referencované vo vzťahu k nejakému konkrétnemu priestoru. Vyššie uvedená skutočnosť však neznamená, že sú referenčnými údajmi podľa NKIVS resp. pre účely tohto dokumentu. Tento dokument sa zaoberá priestorovými údajmi a ich podmnožinou referenčnými priestorovými údajmi.

Právne predpisy ovplyvňujúce priestorové údaje

Základným právnym predpisom týkajúcim sa priestorových údajov na úrovni EÚ je Smernica Európskeho Parlamentu a Rady 2007/2/ES zo 14. marca 2007, ktorou sa zriaďuje **Infraštruktúra pre priestorové informácie v Európskom spoločenstve (INSPIRE)**. Smernica predstavuje právny rámec pre vytvorenie a prevádzkovanie infraštruktúry priestorových informácií v Európe za účelom formulovania, implementácie, monitorovania a vyhodnocovania politík spoločenstva na všetkých úrovniach a poskytovania verejných informácií v oblasti priestorových dát.

Do legislatívy Slovenskej republiky bola táto smernica transponovaná prostredníctvom zákona č. 3/2010 Z. z. o národnej infraštruktúre pre priestorové informácie v znení zákona č. 362/2015 Z. z. (Zákon NIPI) a Vyhlášky č. 12/2017 Z. z. Ministerstva životného prostredia Slovenskej republiky zo 16. januára 2017, ktorou sa mení a dopĺňa vyhláška Ministerstva životného prostredia Slovenskej republiky č. 352/2011 Z. z., ktorou sa vykonávajú niektoré ustanovenia zákona č. 3/2010 Z. z. o národnej infraštruktúre pre priestorové informácie. Transpozíciu dopĺňa séria Ďalších právnych predpisov EÚ (nariadenia a rozhodnutie). Ako odporúčané je potrebné brať do úvahy súvisiace Technické návody INSPIRE. Z ďalších právnych predpisov je potrebné zohľadniť aj zákon č. 215/1995 Z. z. o geodézii a kartografii v znení neskorších predpisov.

Základné právne predpisy dopĺňajú ďalšie právne predpisy EÚ (nariadenia a rozhodnutie) a SR (vyhláška). Ako odporúčané je potrebné brať do úvahy Technické návody INSPIRE. Z ďalších právnych predpisov je potrebné zohľadniť aj zákon č. 215/1995 Z. z. o geodézii a kartografii.

Kvalita priestorových údajov

Aj keď je problematika hodnotenia kvality priestorových údajov vymedzená medzinárodnými štandardmi (napríklad ISO 19157:2013), jej praktická aplikácia stále čelí nemalým výzvam. Systémové a harmonizované hodnotenie kvality priestorových údajov na Slovensku v súčasnosti abscentuje.

Súčasný stav priestorových údajov v SR

V Slovenskej republike sú priestorové údaje v súčasnosti vytvárané, spravované a poskytované do značnej miery neefektívne. Napriek snahe niektorých inštitúcií o dodržiavanie legislatívneho rámca a požiadaviek na vytváranie, spravovanie a zdieľanie priestorových údajov dochádza k duplicitnému i multicpltnému zberu priestorových údajov s rôznou kvalitou. Vo všeobecnosti je tento jav daný:

- Neznalosťou príslušných pravidiel, ktorými sa má riadiť vytváranie, správa poskytovanie priestorových údajov.
- Neochotou zdieľať priestorové údaje vytvorené z prostriedkov štátneho rozpočtu respektíve prostriedkov EÚ, a neochotou akceptovať priestorové údaje vytvorené iným orgánom verejnej moci.
- Nedostatkom kvalifikovaných a kompetentných ľudských zdrojov, vrátane finančných zdrojov na budovanie vlastnej či zdieľanej infraštruktúry priestorových údajov.

Uvedené dôvody majú za následok nesystémový prístup tvorby a zdieľania priestorových údajov, neefektívne vynakladanie prostriedkov štátu na vytváranie a aktualizáciu priestorových údajov a v konečnom dôsledku nedostatočný prístup k priestorovým údajom, respektíve prístup k priestorovým údajom úzko limitovaný, porušuje sa pravidlo „Jedenkrát a dost“, stráca sa možnosť reálneho

manažmentu priestorových údajov na úrovni Slovenskej republiky. Neznalosť je spôsobená aj nízkou mierou informovanosti o priestorových údajoch.

Register priestorových informácií

Register priestorových informácií (RPI) je budovaný v súlade Národnou koncepciou informatizácie verejnej správy (2008) ako základný register štátu. RPI bude po jeho naplnení zo strany povinných osôb poskytovať pre verejnú správu, odbornú verejnosť a občanov jednotný a údajovo konzistentný zdroj údajov o priestorových informáciách vo forme metaúdajov súboroch a službách priestorových údajov pre koncových používateľov. Služby RPI sú dostupné občanom, podnikateľským subjektom, verejnej správe takým spôsobom, aby používatelia RPI nemuseli získavať informácie o priestorových údajoch a službách priestorových údajov z rôznych zdrojov. RPI zjednocuje tvorbu, poskytovanie a aktualizáciu metaúdajov o priestorových informáciách, ako aj vytvorenie a aplikovanie princípov harmonizácie a interoperability priestorových informácií. Jedným z dôležitých rysov RPI pre manažment údajov je zosúladenie sa s Centrálnym metainformačným systémom verejnej správy, od ktorého bude RPI preberať pridelené Jednoznačné referencovateľné identifikátory pre priestorové údaje, čím sa dosiahne celková vzájomná interoperabilita údajov z RPI a ostatných referenčných registrov. Identifikátory tvoria základ pri validácii metaúdajov spravovaných v registri. RPI predstavuje jeden zo základných nástrojov na manažment priestorových údajov v Slovenskej republike a jeho aktívnym využívaním je možné do značnej miery eliminovať nedostatky v súčasnom stave priestorových údajov v SR.

Stav budovania Národnej infraštruktúry pre priestorové informácie podľa príslušnej legislatívy

Povinné osoby zodpovedné v zmysle zoznamu povinných osôb⁶ za jednotlivé témy podľa príloh zákona č. 3/2010 Z. z. nespĺňajú implementáciu v plnom rozsahu a nedodržiavajú stanovený časový harmonogram. Podrobnejšie informácie o stave budovania Národnej infraštruktúry priestorových údajov sú k dispozícii prostredníctvom SK INSPIRE Monitoringu a Reportingu⁷.

1.3.2 Riziká vyplývajúce zo súčasného stavu

Ak by sa súčasný stav nezmenil, je možné očakávať nasledujúce riziká.:

- **Verejná správa nebude mať k dispozícii kvalitné údaje** –, ich nejednoznačnosť, nekonzistentnosť znižuje výkonnosť verejnej správy, znižuje sa miera transparentnosti, zvyšuje sa nedôveryhodnosť verejnej správy a hrozia nesprávne rozhodnutia.
- **Neexistencia dostatočného počtu referenčných údajov respektíve pomalé vyhlasovanie referenčných údajov** – údaje nebude možné referencovať, čo vytvára riziko trvalej a systémovej nekvality údajov a riziko nedostatočnej miery interoperability, efektívneho opätovného využitia referenčných údajov.
- Požiadavky verejnosti a orgánov EÚ na dostupnosť údajov a ochranu osobných údajov nebudú naplnené, vrátane možných sankcií EÚ za neplnenie súvisiacich legislatívnych povinností.
- **Verejná správa nevyužije potenciál**, ktorý prinášajú nové technológie spracovania údajov a „Big Data“ – z údajov nevznikne vzácny zdroj a Slovenská republika nevytvorí jedinečné inovačné prostredie v oblasti dátových analýz.
- **Rozhodnutia vo verejnej správe nebudú podporované kvalitnou analýzou** a založené na dôkazoch. Zámery reformy verejnej správy v oblasti lepšieho návrhu politík a regulácií sa nepodariť podporiť IT nástrojmi – regulačné prostredie v Slovenskej republike ostane suboptimálne.

⁶ http://inspire.enviroportal.sk/povinne_osoby/

⁷ <http://inspire.enviroportal.sk/implementacia/monitoring-a-reporting>

1.3.3 Problémy, ktoré je potrebné vyriešiť

Súčasný stav zo sebou nesie značné množstvo problémov, ktoré je potrebné systematicky riešiť. Za hlavné problémy, ktoré má dokument ambíciu adresovať je možné považovať:

- Momentálne chýba systematický manažment údajov vo verejnej správe – nie sú jasne definované požiadavky pre centrálnu úroveň a jednotlivé inštitúcie verejnej správy.
- Pre manažment údajov nie sú definované jasné a dostačujúce štandardy.
- Legislatívne problémy (najmä nekonzistencia jednotlivých legislatívnych aktov).
- Inštitúcie verejnej správy nedisponujú presnými znalosťami o svojich povinnostiach v súvislosti s manažmentom údajov.

Absencia systematického manažmentu údajov

V rámci verejnej správy v súčasnosti neexistuje systém tvorby a zdieľania spoločných údajov, jednotlivé subjekty nadbytočne požadujú údaje ktoré už existujú v iných ISVS, alebo sa izolovane snažia získať údaje dvojstrannými dohodami s vlastníkmi údajov. Vzhľadom na množstvo uzlov kde údaje vznikajú ako i množstvo uzlov pre distribúciu, je dvojstranná výmena (na úrovni dvoch subjektov) neefektívna, nevyhovuje platnej legislatíve (e-Government), vyvoláva náročný spôsob riešenia nezrovnalostí. Aktuálny stav preto nezabezpečuje jednotlivým subjektom verejnej správy riadený prístup ku aktuálnym údajom. Riadenie a distribúciu spoločných zdieľaných údajov je často ponechaná na individuálne úsilie jednotlivých subjektov.

Súčasný mechanizmus pre zabezpečenie integrácie na úrovni údajov je relatívne komplikovaný – pre každý register je potrebné vypracovávať samostatnú Dohodu o integračnom zámere. Bolo by dobré, ak by si táto požiadavka maximálne zjednodušila.

Nie sú definované jasné dátové štandardy

Ak zhrnieme poznatky o stave uplatňovania dátových štandardov, je možné konštatovať:

- Neexistencia jednotného dostupného modelu údajov verejnej správy (data.gov.sk ontológie).
- Nekvalita základu ISVS: chýba celkový model zverejnených dát.
- Duplicitné/prekrývajúce sa identifikátory používané v číselníkoch.
- Nedefinované životné cykly pre meta-údaje a číselníky.
- Sprístupnenie potrebných údajov od ich držiteľov v štandardizovanom formáte a forme.
- Nejednoznačné chápanie terminológie a nedostatočný výklad najdôležitejších legislatívnych noriem.
- Problémom je i potreba udržať už existujúce projekty minimálne 5 rokov.

Legislatívne problémy

Súčasný zákon o e-Governmente je pre oblasť zdieľania údajov a jedenkrát a dosť postačujúci a dobre vypracovaný s výnimkou veľkého počtu častí MUK. Celá časť o objektoch evidencie, registroch, referencovaní a stotožňovaní je rozpracovaná v dostatočnej miere. **Problém je však s nevykonaním transpozície zákona o e-Governmente do legislatív poskytovateľov.** Napríklad orgán verejnej moci poskytujúci dotačné schémy zo zákona vyžaduje potvrdenie o nedoplatku žiadateľa na dani. Zákon o e-Governmente síce prikazuje dáta poskytnúť ale ak v Daňovom poriadku nie je OVM uvedený ako subjekt prelamujúci daňové tajomstvo tak Finančné riaditeľstvo údaje nemôže poskytnúť. Podobné platí pre osobné údaje. Zákon o e-Governmente má rovnakú právnu silu a teda nemôže systémovo prelamovať ostatné zákony explicitne chrániace dáta. Problémom tiež je, že osobitné predpisy dávajú povinnosť účastníkom konania/povinným osobám predkladať údaje, ktoré sa nachádzajú v referenčných registroch, alebo ktoré majú OVM k dispozícii. Osobitné predpisy však môžu tento problém vyriešiť tak, že po súhlase účastníka konania/povinnnej osoby sa príslušné údaje sprístupnia.



Legislatíva sa môže ukázať ako problém aj pri kvalite údajov. Ak legislatíva nejakého registra určuje spôsob aktualizácie registra, ktorý nie je kompatibilný s princípom referencovania, tak len zákon o e-Governmente nemusí stačiť na vyčistenie registra.

Nedostatočné znalosti povinností OVM

Inštitúcie verejnej správy zatiaľ plnohodnotne neabsorbovali a nevedia systémovo uchopiť plnenie zákona o e-Governmente. Nemajú informáciu, že vo vládnom cloude existujú pre nich dostupné služby pre čistenie dát a hlavne pre stotožňovanie. Inštitúcie tiež nemajú systematickým spôsobom zmapované dátové toky. A ani nevedia, ako sa k problému metodicky postaviť.

Chýba tu koordinačný a hlavne zrozumiteľný metodický usmerňujúci mechanizmus a kontrola plnenia a vymožitelnosť zákonných noriem.

2 Ciele realizácie

2.1 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 stanovené pre zabezpečenie dostatočnej kvality údajov, rozšírenie zoznamu referenčných údajov, zabezpečenie princípu "jedenkrát a dost", transparentné sprístupnenie údajov klientom - moje dáta, lepšie analytické spracovanie údajov vo verejnej správe a prepojené údaje (Linked Data). Ciele sú záväzné, stav ich plnenia bude centrálné monitorovaný, posudzované projekty budú hodnotené aj z pohľadu príspevku k plneniu týchto cieľov.

2.1.1 Kvalita údajov

Tabuľka 1: Návrh cieľov pre kvalitu údajov

Cieľ	Ukazovateľ cieľa	Výsledok
Zabezpečenie dostatočnej dátovej kvality v ISVS organizácie	Kvalita údajov (pre jednotlivé organizácie verejnej správy) meraná parametrami: <ul style="list-style-type: none"> – Presnosť, – Komplexnosť, – Unikátnosť, – Aktuálnosť, – Integrita, – Strojová spracovateľnosť, – Konzistentnosť, – Správnosť. 	99 % inštitúcií dosiahlo navrhnuté prahové hodnoty

Tabuľka 2: Návrh parametrov pre meranie kvality údajov

Ukazovateľ	Hodnoty	Výsledok (Ambícia)
Presnosť (čistota) – miera, akou objekt evidencie reprezentuje reálny svet.	Percento objektov evidencie v ISVS, ktoré majú zistené chyby	<1 %
Komplexnosť - Komplexnosť údajov znamená, že všetky údaje z objektu evidencie, považované alebo označené za povinné, sú prítomné v dátovom prvku.	Percento objektov evidencie v referenčných registroch, ktoré majú prítomné všetky údaje vyžadované agendou.	>97 %
	Percento objektov evidencie v ISVS, ktoré majú prítomné všetky údaje vyžadované agendou.	>70 %



Ukazovateľ	Hodnoty	Výsledok (Ambícia)
	Relevantné objekty evidencie musia byť určené v MetaIS pre každý ISVS.	
Unikátnosť - Vyhodnotenie duplicity údajov vo vzťahu k jednotnému referencovateľnému identifikátoru.	Percento viacnásobného výskytu subjektu evidencie v jednom objekte evidencie	0 %
Aktuálnosť - Údaje sú časovo príslušné a považované za aktuálne.	Percento objektov evidencie, ktorých dátum aktualizácie je rovnaký, ako dátum vzniku relevantnej skutočnosti, ku ktorej sa aktualizácia viaže.	>80 %
Referenčná integrita	Percento údajov v objekte evidencie, ktoré sú stotožnené s relevantnými referenčnými údajmi	100 %
	Percento subjektov evidencie, ktoré majú jednoznačný referencovateľný identifikátor	100 %
Strojová spracovateľnosť - (Možnosti spracovania dát plynúce zo zdrojového formátu dát)	<p>Úroveň 0*: Údaje nie sú poskytované v elektronickej forme.</p> <p>Úroveň 1*: Údaje sú dostupné vo webovom prostredí, avšak formát dát nemá štruktúrovaný obsah a nie je poskytovaný v otvorenom formáte. (napr. PDF)</p> <p>Úroveň 2*: Údaje sú dostupné vo webovom prostredí; obsah je štruktúrovaný tak, že umožňuje automatizované spracovanie, avšak nie je poskytovaný v otvorenom formáte. (napr. XLS)</p> <p>Úroveň 3*: Minimálna úroveň pre publikovanie podľa štandardov pre ISVS: údaje sú dostupné vo webovom prostredí; obsah je štruktúrovaný tak, že umožňuje automatizované spracovanie v otvorenom formáte, avšak nie s ohľadom na význam (napr. CSV, XML).</p> <p>Úroveň 4*: Na identifikáciu entít sa používajú Jednotné referencovateľné identifikátory v podobe URI. Jednotlivé údaje sú popísané prostredníctvom národných ontológií (data.gov.sk). Tým sa dáta vedia strojovo</p>	<p>>70 % objektov evidencie na úrovni 5</p> <p>>90 % objektov evidencie na úrovni 4</p>

Ukazovateľ	Hodnoty	Výsledok (Ambícia)
	spracovať s ohľadom na význam (napr. RDF, OWL). Úroveň 5*: Národné entity sú prepojené na externé globálne štandardizované entity (napr. odporúčené ontológie EÚ na popis dát).	
Konzistentnosť – vzájomné logické vzťahy v rámci objektu evidencie sú správne a v súlade s biznis pravidlami.	Percento objektov evidencie, ktorých údaje spĺňajú logické kritériá konzistentnosti údajov.	100 %
Správnosť - zhoda údajov s kritériami, ktoré stanovujú formát dát	Percento objektov evidencie so správnymi údajmi	100 %

2.1.2 Referenčné údaje a „jedenkrát a dost“

Tabuľka 3

Ciel	Ukazovateľ cieľa	Výsledok
Rozšírenie zoznamu referenčných údajov ⁸	Podiel referenčných údajov, na celkovom množstve údajov (ktoré boli stanovené, že budú vyhlásené ako referenčné)	100 %
Zabezpečenie princípu "jedenkrát a dost"	Percento agend, ktoré nevyžadujú od občanov a podnikateľských subjektov údaje, ktorými už verejná správa disponuje	100 %
Zabezpečenie využívania referenčných údajov v praxi	Podiel inštitúcií verejnej správy, ktoré využívajú referenčné údaje	100 %

⁸ Zoznam kandidátov na referenčné údaje vznikne kombináciou: analýzy požiadaviek inštitúcií verejnej správy na referenčné údaje, analýzy formulárov elektronických služieb a analýzy procesov obsluhy klientov od SVS. Zoznam bude dostupný na webovom sídle informatizacia.sk.

2.1.3 Moje dáta

Tabuľka 4

Ciel	Ukazovateľ cieľa	Výsledok
Sprístupnenie údajov klientom, ktoré sa vo verejnej správe o nich evidujú	Podiel dát o občanovi sprístupňované pre službu „moje dáta“	95 %
	Podiel dát o podnikateľskom subjekte sprístupňované pre službu „moje dáta“	95 %

2.1.4 Analytické spracovanie údajov vo verejnej správe

Tabuľka 5

Ciel	Ukazovateľ cieľa	Výsledok
Zlepšiť rozhodovanie vo verejnej správe	Počet úsekov verejnej správy, v ktorých je rozhodovanie podporované analytickými systémami	77
	Vytvorené analytické jednotky	99 %
Zvýšiť dostupnosť dát pre analytické spracovanie	Podiel dát prístupných na analytické spracovanie (vo formáte umožňujúcom strojové spracovanie)	90 %

2.1.5 Prepojené dáta (Linked Data)

Tabuľka 6

Ciel	Ukazovateľ cieľa	Výsledok
Zlepšenie interoperability údajov zavedením sémantických dátových štandardov	Metóda tvorby URI identifikátorov pre entity ISVS (ontológie a ich entity, individuá, datasety, verzionovanie URI)	100 %
	Vytvorenie jednotného popisu údajov verejnej správy prostredníctvom ontológií v rozsahu katalógu dátových prvkov, referenčných registrov, základných číselníkov a entít MetaIS s registráciou daných URI v MetaIS	100 %
	Metóda a proces prideľovania URI identifikátorov pre údaje ISVS v súvislostiach s povinnosťami zainteresovaných subjektov. Referenčné URI identifikátory pre ľubovoľné entity	100 %

	ISVS vydáva výhradne Centrálny metainformačný systém.	
Zvýšenie dostupnosti údajov vo forme otvorených a prepojených dát	Publikácia otvorených prepojených dát v rozsahu referenčných registrov, základných číselníkov a entít MetaIS	90 %
	Zabezpečenie dereferenciácie URI nad sprístupnenými dátami	100 %

2.2 Dotknuté subjekty a ich záujmy

V tejto časti sú identifikované dotknuté subjekty, ktorých sa daná strategická priorita týka. Manažment údajov riešia inštitúcie naprieč celou verejnou správou, vo všetkých agendových informačných systémoch a v neposlednom rade aj verejnosť. Jej význam nespočíva len v údajoch ako takých, ale aj v kvalite a rýchlosti rozhodnutí, ktoré kvalitné, spoľahlivé, dostupné, aktuálne, referencovateľné údaje umožňujú realizovať.

Úrad podpredsedu vlády SR pre informatizáciu a investície

- Zaviesť manažment údajov ako aktív (vzácných zdrojov) do praxe.
- Zabezpečiť účinné nástroje, ktoré umožnia naplnenie požiadaviek e-Governmentu, princíp „jedenkrát a dost“.
- Mať celkový prehľad o kvalite údajov vo verejnej správe a ako i o využívaní údajov.
- Zvýšiť efektívnosť využívania finančných prostriedkov verejnej správy (racionálna prevádzka e-Governmentu):
 - odstránením vytvárania "peer to peer" rozhraní,
 - odstránením vytvárania duplicitných agend,
 - odstránením manuálneho zadávania referencovateľných údajov,
 - zvýšením kvality údajov.
- Riadenie plnenia úloh v oblasti manažmentu údajov, vrátane kontrolných mechanizmov.

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).
- Plniť povinnosti súvisiace s rolou správcu informačného systému verejnej správy a poskytovateľa údajov.
- Mať jednoduchý prístup k zdieľaným údajom v roli konzumenta.
- Mať prístup k nástrojom pre správu kvality údajov.
- Mať spoľahlivé údaje, informácie slúžiace k analýzam súčasného stavu spoločnosti a verejnej správy vo vybraných agendách, k vyhodnocovaniu dopadu existujúcich opatrení, politik a regulácií, k vytváraniu prediktívnych modelov.

Verejnosť

- Mať prístup k informáciám o relevantných objektoch evidencie týkajúcich sa používateľ a vrátane informácií o stave spracovania podania.
- Mať k dispozícii nástroje pre správu osobných údajov.
- Jednoduché používanie služieb verejnej správy - princíp "jedenkrát a dost" je všade zavedený.



- Mat' prehľad o vykonávaných analýzach, dopadoch politik a regulácií a o zdôvodnení investícií vo verejnom záujme.
- Transparentný prístup k informáciám - mat' prehľad o výkone verejnej správy v jednotlivých sektoroch (až na úroveň jednotlivých inštitúcií) – vrátane ratingov.

Ministerstvo vnútra SR

- Mat' prehľad o kvalite údajov evidovaných v systémoch miestnej štátnej správy.
- Koordinácia zavádzania princípov štátu, fungujúcom na základe využívania dát, do praxe (v rámci reformy verejnej správy).
- Mat' prehľad o výkone a kvalite služieb miestnej štátnej správy.
- Mat' prehľad o výkone a kvalite jednotlivých sektorov verejnej správy.
- Zavedenie IT nástrojov, ktoré podporia lepšie politiky, regulácie a služby (a podporia reformu verejnej správy) – jednoduchšie dosiahnutie cieľov OP EVS.

Úrad splnomocnenca vlády pre rozvoj otvorenej spoločnosti

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

Ministerstvo financií SR

- Vykonávať dohľad nad efektívnosťou politik.
- Vykonávať dohľad nad efektívnosťou investícií.
- Mat' prehľad o dopadoch opatrení na verejné financie a štátny rozpočet.

Úrad vlády SR

- Mat' prehľad o stave napĺňania strategických priorít.
- Mat' k dispozícii analytické podklady pre návrh vládnych stratégií.

Ministerstvo hospodárstva SR

- Vytvorenie efektívneho regulačného rámca pre podporu podnikateľského prostredia.
- Koordinácia procesu hodnotenia vybraných vplyvov a vplyvov na podnikateľské prostredie, podľa kompetencie.

Analytické jednotky verejnej správy

- Mat' kapacity vytvárať kvalitné sektorové analýzy s čo najlepším využitím dát (vrátane mandátu a odbornej nezávislosti ako aj možnosti nárokovať si prístup k dátam).
- Vykonávané analýzy budú integrálnou súčasťou návrhu stratégií, politik a regulácií (relevancia analýz v procese, nie tvorba „do zásuvky“).
- Získať rešpekt v odbornej komunite a uznanie verejnosti.
- Mat' k dispozícii užitočné nástroje a IT infraštruktúru, kvalitné dáta a modely, aby bolo možné pohodlne vytvárať analytické produkty.

Podnikateľské prostredie

- Mat' prístup ku kvalitným analýzám a prognózam o vývoji ekonomiky.
- Nastavenie regulačných rámcov tak, aby sa podporovali inovácie a aby bola znížená regulačná a administratívna záťaž a vylepšené podnikateľské prostredie v súlade s agendou lepšej regulácie.
- Možnosť participatívnej účasti na vyhodnocovaní vplyvov.

3 Organizácia

Návrh modelu zodpovedností a potrebné organizačné zabezpečenie vrátane organizačných zmien.

3.1 Zodpovednosť

Zodpovedný gestor

Zodpovedným gestorom strategickej priority Manažment údajov je **Úrad podpredsedu vlády SR pre informatizáciu a investície** – zodpovedá za referenčnú architektúru informačných verejnej správy, e-Government komponentov, kde vrstva údajov predstavuje kľúčový komponent. Úrad tiež spravuje dátovú časť ostatných úsekov Modulu úradnej komunikácie.

Zapojené organizácie

Zoznam organizácií, ktoré budú dotknuté realizáciou priority:

- **Ministerstvo vnútra SR** – prevádzkuje časť dátového centra e-Governmentu, prevádzkuje dátovú časť vnútornej správy modulu MUK a má v kompetencii kľúčové informačné systémy verejnej správy.
- **NASES** – prevádzkuje Ústredný portál verejnej správy.
- **Úrad vlády SR** – je správcom Ústredného portálu verejnej správy.
- **Štatistický úrad SR** – je ústredný orgán štátnej správy zodpovedajúci za výkon štátnej štatistiky. Má dlhodobé skúsenosti s metodikou zberu, spracovania a poskytovania štatistických údajov spolu s vyhodnocovaním ich kvality.
- **Všetky orgány verejnej správy** - Realizáciou priority manažment údajov budú dotknuté všetky orgány verejnej správy, ktoré poskytujú svoje údaje iným orgánom verejnej správy alebo verejnosti, minimálne v rozsahu zavedenia manažmentu kvality údajov a meta-údajov manažmentu a rovnako aj orgány verejnej moci ktoré na výkon činností vo svojej kompetencii potrebujú vstup údajov
- **Ministerstvo životného prostredia SR** - Primárna úloha Ministerstva životného prostredia Slovenskej republiky spočíva v sprístupnení infraštruktúry priestorových údajov, koordináciou spolupráce medzi jednotlivými orgánmi štátnej moci, kontrolou dodržiavania legislatívy, princípov a pravidiel v oblasti priestorových údajov, s nezastupiteľnou úlohou Úradu geodézie, kartografie a katastra SR pri budovaní referenčných údajov pre tematické priestorové údaje.

V oblasti analytického spracovania údajov sú kompetencie rozdelené nasledovne (ďalšie kompetencie môžu byť definované):

- Gestorom témy Analýza dát je **Úrad podpredsedu vlády SR pre informatizáciu a investície** – zodpovedá za vytvorenie rámca pre lepšie využívanie dát, riadenie dát a vytvorenie analytickej vrstvy.
- Za koordináciu strategických priorít vlády SR zodpovedá Úrad vlády SR.
- Za proces koordinácie RIA a nastavenie modelov pre hodnotenie vplyvov na podnikateľské prostredie zodpovedá Ministerstvo hospodárstva SR.
- Za návrh a tvorbu modelov pre vyhodnocovanie vplyvov na rozpočet zodpovedá Ministerstvo financií SR.
- Za návrh a tvorbu modelov pre vyhodnocovanie vplyvov na malých a stredných podnikateľov – takzvaný MSP Test zodpovedá Slovak Business Agency (SBA).
- Za návrh a tvorbu modelov pre vyhodnocovanie vplyvov na životné prostredie zodpovedá Ministerstvo životného prostredia SR.
- Za návrh a tvorbu modelov pre vyhodnocovanie sociálnych vplyvov zodpovedá Ministerstvo práce, sociálnych vecí a rodiny SR.

- Za koordináciu vyhodnocovania vplyvov politík zodpovedá Inštitút finančnej politiky (IFP) MF SR.
- Za návrh a tvorbu modelov pre vyhodnocovanie investícií zodpovedá Ministerstvo financií SR.
- Za manažment kvality verejnej správy a koordináciu merania výkonnosti jednotlivých sektorov verejnej správy zodpovedá Analyticko-metodická jednotka MV SR.
- Za koordináciu analytických jednotiek zodpovedá Ministerstvo vnútra SR.

3.2 Posúdenie kapacitnej pripravenosti

3.2.1 Personálna kapacita

Jednotlivé zložky štátnej správy nie sú z hľadiska personálnej kapacity na zmeny (ktoré vyvolá naplnenie cieľov a princípov konceptu Manažmentu údajov) pripravené.

3.2.2 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.

3.3 Organizačné zmeny

Je potrebné vytvoriť nové role s definovanými právami a zodpovednosťami v jednotlivých organizáciách verejnej správy:

- Vytvorenie Dátovej kancelárie verejnej správy - nadrezortného tímu zodpovedného za Manažment údajov (Data governance) teda za celý životný cyklus dát, v ktorom budú minimálne zástupcovia vlastníkov dát a dátoví kurátori. Tento tím bude dohliadať hlavne na to, aby sa zapisovanie dát do databáz a systémov uskutočňovalo na základe prijatých štandardov, teda podľa biznis pravidiel, dátových definícií a obmedzení na dátovú integritu definovaných v dátovom modeli a ontológiách – čím sa zabezpečí vysoká kvalita a integrita dát. Centrálna Dátová kancelária bude mať na starosti koordináciu segmentových dátových kancelárií, ktoré predstavujú dátové kancelárie na úrovni orgánov verejnej správy. Metodicky riadi dátových kurátorov, dátových špecialistov, vlastníkov dát a dátových analytikov jednotlivých orgánov verejnej správy.
- Posilnenie kapacít na úrovni jednotlivých inštitúcií verejnej správy (zriadenie segmentových dátových kancelárií). Na základe veľkosti a charakteru danej inštitúcie je potrebné menovať role so stanovenými zodpovednosťami za jednotlivé aspekty manažmentu údajov. Väčšie inštitúcie v prípade potreby rozdelia svoje údaje do segmentov. Je však dôležité, aby bola stanovaná jasná zodpovednosť na úrovni inštitúcie.
- Vytvorenie alebo posilnenie analytickej jednotky na každom ministerstve a ostatných dôležitých inštitúciách verejnej správy. Analytické jednotky budú realizovať analýzy a produkovať analytické produkty, ktoré budú dohodnuté. Časť analytických jednotiek bude mať na starosti jednotlivé súčasti RIA.

Nové role:

Na centrálnej úrovni a lokálnej úrovni budú menované nasledujúce role:

- Programový manažér je zodpovedný za program nového prístupu k dátam a za jednotlivé projekty, ako aj za dennodenné aktivity vzhľadom na rozsah, rozpočet a časové obmedzenia. Tieto role tiež spolupracujú s vlastníckmi dát za účelom zadefinovania požiadaviek na kvalitu údajov.
- Poradca organizačnej zmeny pomáha organizáciám porozumieť hodnote a dopadu nového prostredia verejnej správy založeného na rozhodovaní na základe objektívnych dát a na riadení ich

kvality. Adresuje tiež akékoľvek problémy, ktoré môžu byť spojené s touto zmenou a šíri najlepšie praktiky vedúce k úspešnej implementácii zmeny.

- Vlastník údajov – má zodpovednosť za aktuálnosť a dostupnosť údajov a právomoc na rozhodovanie o prístupe k dátam, k ich distribúcii, definíciám ako aj o nastavení správnych kritérií kvality dát.
- Dátový architekt - poskytuje informačnú a metodickú podporu vlastníčkovi dát pri pretavení jeho požiadaviek do dátového modelu a do procesov zberu a distribúcie dát..
- Dátový analytik – zodpovedá za využitie údajov za účelom analýz, tvorby dôkazov a podkladov pre lepšie rozhodovanie vo verejnej správe.
- Dátový kurátor – riadi a stará sa o údaje, **avšak údaje mu nepatria**. Tvorí rozhranie medzi biznis a BI tímom. Dáva odporúčania týkajúce sa prístupu k dátam, distribúcií, vytvára definície údajov a ich klasifikáciu, usmerňuje aktivity na vytváranie metadát a zvyšovanie kvality dát, identifikuje, zbiera a pomáha pri stanovení priorit požiadaviek na dáta a informácie. Venuje sa organizácii a integrácii dát z rôznych zdrojov, anotácii dát predovšetkým pre ich archiváciu a sekundárne použitie a ich publikácii a prezentácii.
- Dátový špecialista – má zodpovednosť za uloženie dát, ich archiváciu, backup, zabráňuje poškodeniu dát, neoprávnenému prístupu k dátam.

Tabuľka 7: Prehľad organizačného zabezpečenia

Úroveň	Zodpovednosť	Rola	Počet
Centrálna (ÚPPVII)	– Celková zodpovednosť za manažment údajov vo verejnej správe	CDO	1
	– Vzdelávanie a podpora	Poradca pre vzdelávanie	1
	– Program Manažment údajov	Programový manažér	1
		Poradca organizačnej zmeny	1
	– Kvalita údajov	Hlavný dátový kurátor	1
		Dátový kurátor	2
	– Analytické spracovanie údajov	Dátový architekt	1
		Dátový analytik	1
	– Priestorové údaje	Dátový kurátor	1
Ministerstvo vnútra SR	– Program Manažment údajov	Programový manažér	1
	– Životný cyklus údajov	Vlastník údajov	>1 (pre každý segment údajov)
	– Kvalita údajov	Dátový kurátor	>1 (pre každý segment údajov)



Úroveň	Zodpovednosť	Rola	Počet
	– Analytické spracovanie údajov	Dátový analytik	3
Väčšia inštitúcia ⁹	– 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

⁹ Skutočne potrebné kapacity sa môžu líšiť na základe rozsahu agendy jednotlivých inštitúcií verejnej správy (vo vzťahu k spracovaniu údajov)

4 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.

4.1 SWOT analýza

Tabuľka 8: SWOT analýza pre manažment údajov

Silné stránky	Slabé stránky
<ul style="list-style-type: none"> Časť technológie (nástroje pre prístup k údajom, kvalitu údajov a správu kmeňových údajov) a metodiky dodaná v rámci MUK dátová časť ostatných úsekov (pozri tiež prílohu 8.2) Overenie funkčnosti MDM v podmienkach SR na pilotnej vzorke inštitúcií Pochopenie dôležitosti manažmentu dát vyjadrené prostredníctvom NKIVS 	<ul style="list-style-type: none"> Nepripravenosť väčšiny inštitúcií na systematický manažment údajov Nedostatočná osвета De jure chýbajúca metodika pre riadenie kvality údajov Nízka kvalita údajov Správcovia referenčných registrov nevedia ako vykladať zákon (SLA, personálne zabezpečenie a podobne) Neexistujúce personálne a procesné zabezpečenie Nejednotné riešenie Modulu úradnej komunikácie – dátová časť Zložité integrácie v rámci G2G a nejednotný spôsob prístupu k údajom (každý má vlastné objekty, prvky, identifikátory, API)
Príležitosti	Hrozby
<ul style="list-style-type: none"> Vyčistenie dát v štátnej správe Zníženie objemu dát zapisovaných do jednotlivých informačných systémov Spokojnosť verejnosti so znížením administratívnej záťaže Spokojnosť verejnosti s kvalitou poskytovaných údajov Zlepšenie kvality regulačného prostredia, vďaka lepšiemu využívaniu dát Vytvorenie jednotného dátového modelu a vybudovanie jednotnej dátovej platformy 	<ul style="list-style-type: none"> Prílišné očakávania a následné sklamanie vedúce k zastaveniu programu Zablokovanie procesu MDM z dôvodu komplikovaného manažmentu identifikátorov (SIFO, SID) Ukončenie technologických projektov bez následného využitia pre samotnú činnosť riadenia kvality dát Nezískanie kvalitných ľudských zdrojov pre prácu s dátami v štátnej správe Politické vplyvy

4.2 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 vytážiť z údajov.
- Verejná správa bude schopná komplexne pokryť životný cyklus údajov.

Znamená to tiež, že pre jednotlivé oblasti rozhodovania budú k dispozícii 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 podobe pre verejnosť.

Treba zaviesť jednotné riešenie, ktoré však nebude narušovať autonómiu jednotlivých inštitúcií verejnej správy v oblasti tvorby údajov a riadenia prístupu k nim.

4.3 Výber prístupu

4.3.1 Kategorizácia a klasifikácia údajov

Aby bolo možné nastaviť jednotné pravidlá pre prácu s údajmi, je potrebné vychádzať z jasnej klasifikácie. Klasifikácia údajov umožňuje údaje kategorizovať podľa vybraných kritérií tak, aby bolo možné nastaviť efektívne a účelné pravidlá pre prácu s takýmito údajmi a aby bol pokrytý potrebný rozsah údajov. V prostredí verejnej správy navrhujeme jednotnú kategorizáciu, ktorá rozlišuje podľa charakteru údajov, podľa vlastníctva údajov, podľa pôvodu údajov, podľa zdroja, podľa typu a formy, podľa úrovne strojového spracovania a podľa typu interoperability zverejnených dát. Návrh kategorizácie je možné nájsť v prílohe 8.3 Kategorizácia a klasifikácia údajov.

Pravidlá, ktoré budú nastavené v súvisiacich metodikách budú vychádzať z tejto kategorizácie. Dôležitými kategóriami údajov sú:

- referenčné údaje (pre zabezpečenie princípu jedenkrát a dost'),
- infoúdaje,
- údaje v dokumentoch,
- priestorové údaje.

Údaje predstavujú reprezentáciu faktov, ako čísla, text, obrázky, video-súbory. Fakty sú v údajoch vyjadrené, uložené, zachytené. Na základe údajov a ich kontextu je možné odvodiť informácie, pričom kontextom je formát údajov, význam údajov, relevancia údajov a časová platnosť údajov.

Každá inštitúcia verejnej správy pracovala s vlastnou doménou údajov a pre spoločné využitie týchto údajov a ich zdieľanie naprieč verejnou správou je potrebné nastavenie dvoch základných mechanizmov:

- Riešenie spoločných slovníkov a ontologického modelu (metadát, číselníkov a podobne) – aby si informačné systémy verejnej správy medzi sebou rozumeli.
- Zavedenie princípov „Master Data Managementu“ (správy kmeňových údajov) - aby sa používali rovnaké hodnoty pre identifikované kmeňové záznamy – aby sa naprieč informačným prostredím používali:
 - kompletné, to znamená obsahujú všetky dostupné informácie existujúce vo verejnej správe;
 - časovo aktuálne, to znamená. vyjadrujú poslednú platnú informáciu zo zdroja dát;
 - preukázateľné a konzistentné.

4.3.1.1 Referenčné údaje a jedenkrát a dost'

Referenčné údaje a číselníky tvoria kontext pre agendové údaje. Referenčnosť údajov nie je nutnou podmienkou na plnenie princípu „jeden-krát a dost'“, na druhej strane vyhlásenie údajov za referenčný dáva OVM jasný postup a časové rámce pre plnenie princípu 1x a dost' vrátane právnej istoty.

Splnením tejto podmienky je možné nastaviť princíp jedenkrát a dost' – požiadavku aby pri interakcii s verejnou správou sa od žiadateľa vyžadovali len údaje, ktoré sú nové a verejná správa nimi ešte nedisponuje. Na to, aby daný údaj nemusel byť požadovaný je potrebné zabezpečiť:

- aby bolo jasné, kto je za údaj zodpovedný a v ktorom objekte evidencie je údaj zaradený,
- aby bol vybraný subjekt evidencie jednoznačne identifikovateľný a stotožnený,
- aby bol vybraný objekt evidencie prístupný pre použitie.

Referenčné údaje tak slúžia na dva účely:

- synchronizáciu údajov a stotožňovanie subjektov evidencie (inštitúcia, ktorá referencuje referenčné údaje má právnu istotu pri ich používaní),
- výmenu a zdieľanie údajov (faktov) medzi inštitúciami verejnej správy pre potreby výkonu verejnej moci.

Základné číselníky (code list) – sú údaje, ktoré sa používajú na kategorizáciu a klasifikáciu iných údajov. Hodnota je väčšinou vyberaná zo zoznamu, ktorý sa mení len vo výnimočných prípadoch.

Kmeňové údaje (master data) – na rozdiel od číselníkov nie sú obmedzené množinou povolených hodnôt. Kmeňové dáta predstavujú kontext pre transakčné údaje a sú relatívne stále. Za základné typy kmeňových údajov môžeme považovať:

- Osoby (parties) – fyzické a právnické, organizácie, vrátane ich rolí, ako konateľ, študent, občan, pacient, majiteľ a podobne.
- Produkty a služby.
- Lokality a adresy.
- Finančné štruktúry.

Objekt evidencie (business information unit) – je ucelená množina údajov o evidovanom subjekte, ktorá je predmetom evidovania orgánom verejnej moci v rámci jeho pôsobnosti a ktorá je jednoznačne identifikovaná identifikátorom objektu evidencie. Objekt evidencie má jasnú štruktúru a môže tak obsahovať referencované údaje z referenčných registrov.

Referenčné údaje – Referenčný údaj je údaj, ktorý je vyhlásený za referenčný údaj. Referenčný údaj je súčasť objektu evidencie s jasným významom. Použitie referenčného údajá garantuje právnu istotu v konaní pre konzumenta. Referenčný údaj môže byť aj zložený dátový prvok. Každý referenčný údaj má jasnú štruktúru a prvky tejto štruktúry môžu odkazovať na iný referenčný údaj. Dôležitým aspektom je v tomto prípade vyjadrenie vzťahu, ktorý takýto referenčný údaj reprezentuje – to znamená jednoznačný referencovateľný identifikátor, nie konkrétne údaje, ktoré sú referenčné v inom registri.

Kandidátom na referenčný údaj je taký údaj, ktorý potrebuje iný orgán pre svoj výkon kompetencií než len ten, ktorý ho vytvára a používa v rámci vlastných kompetencií svojim informačným systémom. Pri stanovovaní kandidátov na referenčný údaj je tiež potrebné vyhodnotiť charakter údajá z pohľadu toho, či je to údaj popisujúci stav subjektu evidencie (osoba, vec, právo, povinnosť alebo skutočnosť), alebo údaj má transakčný charakter evidujúci napríklad aktivity alebo plnenia si zákonných povinností subjektu evidencie (napríklad daňové priznania, colné vyhlásenia a podobne).

Používaný centrálny systém riadenia referenčných (zdieľaných) údajov verejnej správy, obsahuje nástroje pre prístup k referenčným údajom vo verejnej správe. Zdieľané agendové údaje sa nachádzajú vo forme agendových zdrojových referenčných údajov (čo sú údaje na strane zdrojových systémov) a vo forme agendových odvodených referenčných údajov (čo sú údaje na strane používateľských systémov).

Postup vyhlasovania referenčných údajov je momentálne určený v metodickom usmernení MF/012027/2016-171¹⁰. Pre dosiahnutie cieľov je okrem zabezpečenia potrebných organizačných kapacít potrebná aj voľba proaktívneho prístupu pri vyhlasovaní referenčných údajov. Samotný plán vyhlasovania tak bude evidovaný v MetaIS a vznikne s ním s ohľadom na naplnenie cieľa pre „jedenkrát a dost“ na základe:

- sledovanie požiadaviek inštitúcií verejnej správy,
- analýzy zverejnených formulárov na slovensko.sk,
- výstupov aktivít sekcie verejnej správy.

Prvá verzia plánu vyhlasovania referenčných údajov bude pripravená do 30.4.2017 a plán bude aspoň raz ročne aktualizovaný.

Agendové údaje: popisujú jednotlivé udalosti a fakty v rámci danej agendy verejnej správy, ktoré sa evidujú v informačnom systéme verejnej správy, napríklad rozhodnutia, daňové priznania, colné deklarácie, žiadosti, platobné príkazy, zoznamy neplatičov a podobne. Agendové údaje majú vždy časovú dimenziu. Niektoré agendové údaje sa môžu stať referenčnými.

Pre rezorty bude aktualizované metodické usmernenie (takzvaná „kuchárka“) o referenčných údajoch. Metodické usmernenie bude obsahovať aj odporúčanie ohľadom zmeny osobitných predpisov upravujúcich agendy orgánov verejnej moci tak, aby tieto nebránili v realizácii princípu „jedenkrát a dost“.

Princípy a pravidlá pre referenčné údaje budú aplikované pre všetky údaje zdieľané medzi orgánmi verejnej moci používané pre ich konanie. Katalóg pre údaje bude vedený v MetaIS.

Základné rozhodnutia

Vymedzenie referenčných údajov

Vymedzenie konceptu referenčných údajov odpovedá na otázku, ktoré údaje môžu byť referenčné. Zvolená možnosť predpokladá univerzálne využitie mechanizmu pre zdieľanie údajov. Ako referenčné údaje budú chápané všetky údaje, ktoré vyjadrujú jedinečný fakt – na základe našej klasifikácie pôjde o kmeňové (evidenčné) a agendové údaje. Znamená to, že za referenčné údaje budú považované aj údaje o nedoplatkoch respektíve o poberaných dávkach pre subjekt. Takéto vymedzenie umožní flexibilný návrh objektov evidencie a výrazne rozšíri priestor pre zdieľanie údajov.

Nástroje pre „Jeden-krát a dost“

Pri výbere nástrojov pre naplnenie princípu „jeden-krát a dost“ budú:

- Preferované metódy zdieľania údajov (referenčných) – jedenkrát a dost sa zabezpečí tak, že v procesoch vybranej osoby nebude nutné dokladať už známe fakty, ale systém (alebo úradník) získa príslušnú informáciu cez platformu pre zdieľanie údajov. Ak sa takýto údaj použije ako podklad pre ďalšie rozhodnutie, je dôležité, aby bola zaznamenaná informácia o hodnote údajov v čase konania.
- Zasielanie údajov na základe splnomocnenia - v obmedzených prípadoch (ak údaje nie sú k dispozícii referenčné údaje), na základe vyjadrenia vôle subjektu o ktorom sú údaje spracúvané.
- Ad-hoc výmena údajov bude podporovaná do vzniku mechanizmu na zdieľanie rozhodnutí.

Identifikácia fyzických osôb

V celkovom modeli údajov sú dôležité práve identifikátory (najmä fyzických osôb), ktoré jednoznačne identifikujú subjekty evidencie (kmeňové záznamy). V prípade identifikátora fyzických osôb sa očakáva realizácia opatrení, ktorých výsledkom má byť zavedenie jednoznačného univerzálne použiteľného

¹⁰ http://www.informatizacia.sk/ext_dok--metodicke_usmernenie_na_zaradovanie_ru_do_zoznamu/22954c

identifikátor, ktorý nebude obsahovať ďalšie charakteristiky dotknutej osoby (BIFO). Pri ďalšom postupe sa očakáva:

- zmiešané použitie rodného čísla a BIFO (pričom tieto identifikátory môžu byť aj vo formáte URI) - v jednej množine objektov môžu takto súčasne byť prítomné záznamy s rodným číslom aj BIFO (máme na mysli stav viditeľný na rozhraní platformy zdieľaných údajov, nie internú realizáciu).
- prehodnotenie plánu implementácie sektorových identifikátorov, najmä vo väzbe na ciele a metódy zdieľania údajov verejnej správy.
- použitie sektorových identifikátorov spôsobom, ktorý uprednostňuje spoločný sektor pre celú verejnú správu, pričom nové sektory môžu byť založené v prípade odôvodnených výnimiek. Systém segmentov sa určí na základe zoznamu segmentov, ktorý odôvodnený cost-benefit analýzou a analýzou rizík.

4.3.1.2 Infoúdaje

Údaje, poskytujúce informačný obsah sú veľkou kategóriou, ktorá nie je žiadnym spôsobom uchopená, spravovaná a štandardizovaná. Patria sem napríklad obsahové časti informačných webstránok. Je potrebné, aby vznikla kategorizácia typu publikovanej informácie (popis agendy, zverejnený dokument na úradnej tabuli, popis životnej situácie, informácia o inštitúcii a podobne) a definovanie povinnej sady metaúdajov o takejto informácii, centrálne evidovaných a ukladaných. Takisto z dôvodu jednotného vzhladu stránok inštitúcií je potrebné, aby bola pre každú kategóriu zadefinovaná základná štruktúra infoobsahu (takzvaný Corporate design).

Uplatnenie týchto princípov umožní synchronizáciu obsahu medzi jednotlivými portálmi, čím sa zabezpečí jednotnosť poskytovanej informácie, zjednoduší vyhľadávanie a poskytne lepšiu informovanosť voči verejnosti a tým zlepši informatizáciu spoločnosti.

4.3.1.3 Údaje v dokumentoch

Dokumentmi pre potreby tohto materiálu sa rozumejú primárne elektronické správy, ktoré nie sú vo svojej celej štruktúrovanej podobe zapisované do evidencií a registrov (to je nie sú to štruktúrované správy, ktorých jednotlivé atribúty sa zapisujú ako položky do databáz). Typickými dokumentmi sú podania, rozhodnutia a podobne. Väčšina dokumentov spadá pod správne konanie, sú však dokumenty, ktoré spadajú aj pod iné spôsoby spracovania.

Spravidla sa dokumenty na strane inštitúcie verejnej správy evidujú v registratúrach, ktoré o nich vedú základné údaje a zaraďujú podľa vecnej a agendovej príslušnosti do spisov. Napriek tomu, že legislatíva týkajúca sa registratúr a vedenia a archivácie spisov je dobre spracovaná, metadáta o jednotlivých dokumentoch nie sú navzájom kompatibilné z dôvodu rôznorodosti výrobcov elektronických systémov pre vedenie registratúr a spisovej služby. Navyše ešte stále (najmä v prípade menších inštitúcií) je vykonávaná registratúrna činnosť výlučne v papierovej podobe alebo v takzvanej hybridnej podobe (časť v papierovej a časť v elektronickej podobe).

Rozhodnutia vo verejnej správe

Je potrebné, aby vznikli centrálné štandardy pre dátové prvky (metadáta) dokumentov a povinnosť viesť spisy výlučne v elektronickej podobe. V budúcnosti to uľahčí a zrýchli postupovanie dokumentov, zjednoduší vedenie odosielateľov a adresátov formou identifikátorov a podobne.

Často sú úradné dokumenty vydávané jednou inštitúciou používané ako vstup pre podania pre iné inštitúcie. V zmysle princípu „jedenkrát a dosť“ je potrebné zaviesť centrálnu evidenciu vydávaných úradných dokumentov tak, aby podávajúcí nemusel prikladať k podaniu samotný úradný dokument ale aby si tento dokument dohľadal. Je potrebné navrhnúť jednotný a efektívny spôsob vzájomného zdieľania elektronických úradných dokumentov. Tento prístup má byť postavený na čo najširšom využití existujúcich komponentov a princípov zdieľania údajov verejnej správy.

Rozhodnutia v našom kontexte predstavujú špecifické dokumenty, ktoré vyjadrujú výsledok konaní orgánov verejnej moci. Spomedzi dokumentov by práve rozhodnutia mali byť ako prvé strojovo spracovateľné (100 percent na úrovni metadát, 60 percent na úrovni rozhodnutí ako celku). Funkčný mechanizmus pre zdieľanie rozhodnutí by mal:

- uchovávať metadátá o rozhodnutiach,
- zaviesť princíp a štandardy zdieľania a referencovania elektronických spisov a ich jednotlivých položiek.

4.3.1.4 Priestorové údaje

Spôsob dosiahnutia cieľového stavu v doméne priestorových údajov:

- Definovať pravidlá na identifikáciu priestorových údajov a referenčných priestorových údajov.
- Spresniť zodpovednosť za vytváranie a aktualizáciu priestorových údajov a referenčných priestorových údajov.
- Identifikovať relevantné priestorové údaje a referenčné priestorové údaje a zabezpečiť ich dokumentáciu prostredníctvom meta-údajov.
- Vytvárať, spravovať, zdieľať priestorové údaje a referenčné priestorové údaje harmonizovaným spôsobom pre dosiahnutie požadovanej miery interoperability.
- Vytvoriť a prevádzkovať pre priestorové údaje a referenčné priestorové údaje štandardizované aplikačno programové rozhrania a súvisiace služby.
- Zabezpečiť harmonizované zdieľanie a integráciu priestorových údajov a referenčných priestorových údajov na národnej úrovni v rámci eGovernmentu aj v medzinárodnom kontexte (ISA2, INSPIRE).
- Využiť existujúce priestorové údaje ISVS.
- Definovať pravidlá pri budovaní nových IS spravujúcich resp. využívajúcich priestorové údaje.

4.3.2 Kvalita údajov

Kvalita údajov má priamy dopad na možnosti využitia údajov v oblastiach a ovplyvňuje:

- kvalitu rozhodovania vo verejnej správe,
- spoľahlivosť navrhovaných politík a regulácií vo verejnej správe,
- jednoduchosť a komfort služieb pre podnikateľov a občanov,
- ako aj na platnosť prípadného vedeckého výskumu, ktorý by sa vykonával nad údajmi verejnej správy.

Navrhujeme zaviesť systematické meranie kvality s využitím kritérií:

- **Presnosť** (čistota) – miera, s akou objekt evidencie reprezentuje reálny svet, vyjadrená zhodou s referenčnými údajmi.
- **Kompletnosť** – Kompletnosť údajov znamená, že všetky údaje z objektu evidencie, považované alebo označené za povinné, sú prítomné v dátovom prvku.
- **Aktuálnosť** – Údaje sú časovo príslušné a považované za aktuálne.
- **Unikátnosť** – Vyhodnotenie duplicity údajov vo vzťahu k jednoznačnému referencovateľnému identifikátoru.
- **Referenčná integrita** – Údaje v objekte evidencie sú referencované s referenčnými údajmi v podobe, v akej sú evidované v referenčných registroch. Referenčné údaje z objektu evidencie sú stotožnené so subjektom evidencie.
- **Strojová spracovateľnosť** – možnosti spracovania údajov plynúce z formátu reprezentácie dát ako napr. spájanie rôznych dát z rôznych zdrojov, či spracovanie dát s ohľadom na ich význam.

- **Konzistentnosť** – vzájomné logické vzťahy v rámci objektu evidencie sú správne a v súlade s biznis pravidlami.
- **Správnosť** - zhoda údajov s kritériami, ktoré stanovujú formát dát.

Kvalitu údajov chápeme v dvoch rozmeroch: ako celkovú kvalitu dát v informačnom prostredí verejnej správy (vrátane možnosti vynucovania zverejňovania dát príslušnými inštitúciami verejnej správy) a v druhom rozmere ide o kvalitu údajov evidovaných v jednotlivých ISVS.

Pre splnenie kritérií kvality bude potrebné nastaviť procesy riadenia a správy dátovej kvality: okrem samotného sprístupnenia údajov, bude potrebné aj zrozumiteľne zdokumentovať dátové štruktúry, proces tvorby dát, štatistické metodológie (ak boli použité), dátové zdroje, kontext a mnoho ďalších aspektov. Kvalita údajov má aj mäkké aspekty ako je prehľadnosť a zrozumiteľnosť ich prezentácie, preškolenie zamestnancov, metodické pokyny a podobne.

Problém kvality údajov začína byť vypuklý hlavne v kontexte ich zdieľania. Konkrétnym príkladom problematiky kvality údajov je proces stotožnenia, to je potvrdenia si, že rôzne dáta v rôznych IS sa viažu práve k jednej „stotožnenej“ entite.

Základné rozhodnutia

Prístup k riadeniu kvality údajov

Pre riadenie kvality údajov vo verejnej správe sa bude používať hierarchický prístup: z centrálnej úrovne sa nastaví metodika, parametre, procesy a systém monitoringu. Samotná zodpovednosť za výkon týchto procesov a za kvalitu údajov v jednotlivých úsekoch správy ostane decentralizovaná (na úrovni inštitúcií verejnej správy).

Služby pre kvalitu údajov

Prioritne sa budú používať spoločné (centrálne) nástroje a služby pre zlepšovanie kvality údajov (služby dátovej kvality). Navrhne a do praxe sa zavedie štandardný katalógu služieb pre kvalitu údajov. V kontexte vládneho cloudu je použitie spoločných metód a nástrojov ekonomické.

Vo vybraných prípadoch (ako sú napríklad dôležité a kritické údaje, špecifické legislatívne požiadavky alebo v prípade, ak sú už nástroje k dispozícii a nasadené) sa nevyklučuje použitie lokálnych nástrojov, ak to bude potrebné a ekonomicky výhodné.

Monitoring kvality údajov

Zavedie sa mechanizmus merania a monitoringu stavu dátovej kvality v duchu hesla: „Ak chceme monitorovať a zlepšovať dátovú kvalitu, musíme sa ju najskôr naučiť merať.“ Ide o kľúčový predpoklad, aby sme vedeli vyhodnocovať úspešnosť realizácie jednotlivých opatrení na zvyšovanie dátovej kvality. Najvyššou prioritou bude byť meranie a monitorovanie stavu dátovej kvality pre údaje v referenčných registroch. Realizujú sa opatrenia, ako:

- centrálny zber KPI kvality údajov,
- zverejňovanie kvality údajov pre jednotlivé inštitúcie.

Štandardy

„Základné pojmy dátovej kvality“ a „Atribúty a kritériá DQ“ boli pre tento dokument navrhnuté na základe metodiky Data management association (DAMA International©) - Guide to The Data Management Body of Knowledge (DAMA-DMBOK Guide). V prípade priestorových údajov je potrebné v oblasti kvality zabezpečiť súlad s relevantnými štandardmi, ktoré sú výstupom

štandardizačnej komisie ISO/TC 211 – Geographic information/Geomatics¹¹ ako aj požiadavkami na kvalitu údajov zadefinovanú v kontexte INSPIRE¹².

4.3.3 Integrácia údajov

Integráciu údajov je možné chápať ako prepojenie agendových systémov, ich sekundárnych registrov a evidencií na úrovni údajov (najmä integrácia na úrovni jednotlivých dátových úložísk, synchronizácia, aktualizácia a referencovanie). Cieľom integrácie údajov je v čo najväčšej miere uľahčiť tok dát a informácií. K tomu je nevyhnutné identifikovať požadované dátové toky a poskytnúť nástroje pre referencovanie údajov, synchronizáciu údajov, riešenie konfliktov a podobne.

Vďaka integráciám údajov, údaje o subjekte, ktoré vlastní a spravuje iný orgán verejnej správy, si inštitúcie poskytujú navzájom bez potreby interakcie s klientom (FO/ PO) podľa pravidiel:

- Vytvorenie jednotného dátového modelu údajov verejnej správy (ontológie) + používanie URI identifikátorov.
- Pre každý proces bude navrhnutý biznis kontext: zoznam typov dátových objektov (z iných inštitúcií verejnej správy), ku ktorým je možné priamo pristupovať – na základe tohto kontextu budú nastavené oprávnenia pre prístup k údajom.
- Biznis kontext bude centrálne spravovaný.
- Subjekt bude informovaný o výmene údajov s ním súvisiacich medzi orgánmi verejnej moci.
- Vo vybraných situáciách klient bude môcť dávať súhlas na zdieľanie údajov (podľa zákona).
- Automatizácia zdieľania údajov medzi jednotlivými informačnými systémami verejnej správy, odstránenie manuálnych činností pracovníkov verejnej správy (prepisovanie údajov z jednej aplikácie do druhej), ktoré môžu byť zautomatizované.

Integrácia údajov bude riadená centrálne a bude definovaná v dátovej architektúre. Bez integrácie údajov a architektúry by nebolo možné zabezpečiť požadovanú kvalitu údajov a neodstránili by sa duplicity a nekonzistentnosti.

Uprednostňujeme platformový prístup k dátovej integrácii: vytvorenie prostredia, v ktorom si poskytovatelia údajov a konzumenti údajov môžu zdieľať objekty evidencie riadeným (kontrolovaným spôsobom). Platforma by mala byť otvorená a poskytnúť možnosti pre pripojenie informačného systému cez nástroje integrácie údajov, možnosti pre konzumentov údajov so zabezpečením jednotného prístupu k objektom evidencie a ako i konsolidovaným údajom, možnosti pre nasadzovanie aplikácií (ako čistenie údajov, komplexné služby transformácie dát, dočasné ukladanie dát, tvorba jednoduchého registra, ...). Využitie týchto "aplikácií" je v kompetencii jednotlivých poskytovateľov, alebo konzumentov údajov. Pri prístupe k údajom je potrebné zabezpečiť dve spôsoby interakcie:

- cez aplikačné rozhrania systému (prioritne),
- cez používateľské rozhrania respektíve používateľský prístup na systém, prostredníctvom GUI (voliteľne).

Za základné otázky, ktoré je potrebné vyriešiť je výber prístupu a konceptu, nastavenie pravidiel pre spúšťanie udalostí a vyriešenie ukladania dát, ako vstup pre návrh architektúry.

Platforma integrácie údajov bude otvorená aj pre systémy tretích strán (zo súkromného sektora). Štát si berie za úlohu vytvoriť nástroje pre jednotný prístup k údajom subjektu (na základe súhlasu osoby ku ktorej sa moje údaje viaže a ich implementáciu na základe konceptu OpenAPI. Znamená to, že súkromné inštitúcie môžu znížiť administratívnu povinnosť pri svojich službách, keď budú mať k dispozícii overené referenčné údaje (respektíve keď poskytnú svoje údaje do platformy pre ďalšie použitie).

¹¹ http://www.iso.org/iso/home/store/catalogue_tc/catalogue_tc_browse.htm?commid=54904

¹² http://inspire.ec.europa.eu/documents/INSPIRE_/JRC83209_Online_Data_quality_in_INSPIRE.pdf

Kľúčové bude nastaviť dostatočný manažment oprávnení pre prístup k údajom verejnej správy za podmienok ochrany osobných údajov.

Základné rozhodnutia

Obrázok 2: Prehľad dôležitých otázok v oblasti integrácie údajov



Koncept integrácie údajov

Základnou myšlienkou pre riešenie otázky integrácie údajov je využitie jednotnej platformy a vytvorenie spoločného mechanizmu integrácie údajov vo verejnej správe. Znamená to, že sa zlúčia všetky dátové časti modulu úradnej komunikácie od jednej. Gestorom tohto modulu bude ÚPPVII. Modul bude slúžiť na prístupovanie všetkých údajov v podobe objektov evidencie v príslušnom kontexte.

Logika integrácie údajov

Interpretácia údajov bude postavená na ontológiách evidovaných v centrálnom modeli údajov verejnej správy. Takýto prístup výrazne uľahčí spracovanie údajov na strane konzumentov.

Údaje budú centrálne modelované ako prepojitelné (Linked Data). Samotná výmena údajov bude prebiehať prostredníctvom štandardizovaných štruktúrovaných objektov.

Spôsob riešenia komunikácie pri integrácii údajov

Navrhnuté riešenie bude flexibilné a podporí nasledujúce vzory pre komunikáciu pri dátovej integrácii, pričom kritérium je vhodnosť:

- Asynchrónna komunikácia - bude preferovaná, umožňuje rýchlu optimalizáciu procesov a ďalšieho rozvoja riešenia. Je potrebné, aby sa zosúladiła platnosť údajov pri zmenách vo všetkých procesoch verejnej správy.
- Real-time (synchrónna komunikácia) – je ideálne z pohľadu konzumenta – bude použitá ako preferovaná v rámci platformy integrácie údajov všade tam, kde je to z technických a kapacitných dôvodov možné.
- Dávkové spracovanie (asynchrónna výmena veľkého množstva údajov, platných k určitému času) - bude používaná len v nevyhnutných prípadoch, najmä pre veľmi veľké objemy údajov z dôvodu časovej efektívnosti.
- Virtualizácia dát - bude používaná najmä pre komplexnú integráciu údajov z viacerých zdrojov a typov (pri službe „moje dáta“, analytickom spracovaní údajov).

Koncept správy kmeňových údajov - MDM

Ako koncept správy kmeňových údajov bude použitá Federácia údajov: vznikne jednotný pohľad na dáta dostupný pre konzumentov údajov, ktorý si zabezpečia stotožnenie vo vlastnom informačnom systéme. Výhodou je dostatočná miera autonómie pre správcov informačných systémov so zaručením dostatočnej miery kvality údajov.

Ukladanie údajov (v jednotnej platforme integrácie údajov)

Údaje budú ukladané v zdrojových informačných systémoch. Pre konkrétne prípady použitia na obmedzenú dobu bude podporované dočasné ukladanie údajov v rámci jednotnej platformy integrácie údajov. Vybudované budú tiež podporné komponenty poskytujúce služby perzistencie údajov pre poskytovateľov a konzumentov údajov

Podporované technológie

Dátová integrácia bude využívať technológiu ESB – to znamená, že dátovej integrácii bude dedikovaný vyhradený server, ktorý bude manažovať prenos informácií medzi informačnými systémami verejnej správy. Vytvára to jasnú požiadavku na aplikačnú architektúru a integráciu a orchestráciu.

Zaradenie integrácie údajov – komunikácia s registrom

V rámci riešenia budú oddelené funkcie na zdieľanie údajov od zápisových funkcií a ostatných integrácií. Funkcie na sprístupňovanie údajov majú podstatne odlišný charakter od zápisových funkcií, najmä z pohľadu riadenia prístupu, rozsahu subjektov ktoré majú služby využívať, špecifickosti pre doménu, závislosti na procesoch. Preto je správne vytvoriť pre zdieľanie údajov špecifické pravidlá (vrátane aplikačnej podpory), oddelene od zápisových funkcií.

4.3.4 Moje dáta a ochrana osobných údajov (GDPR)

Všeobecné nariadenie (EU) 2016/679 - GDPR o ochrane osobných údajov posilní dôveru v digitálne služby, nakoľko bude chrániť jednotlivcov v súvislosti so spracovaním osobných údajov všetkými spoločnosťami, ktoré ponúkajú svoje služby na európskom trhu (aj inštitúcie verejnej správy). Zavedie sa moderný regulačný rámec v súlade s nariadením pre kategorizáciu údajov, ochranu osobných a citlivých údajov a pre efektívne využívanie dát z verejného aj súkromného sektora, ktoré dokážu pozitívne prispieť k verejnému záujmu. Občania postupne získajú právo kontrolovať tok svojich dát, ktoré vlastnia, a udeľovať súhlas na ich spracovanie (aj nad rámec spracovania ktoré sú povinní strpieť na základe požiadaviek záväzných právnych predpisov).

V prostredí e-Governmentu sa vybuduje riešenie pre manažment osobných údajov (Personal Information Management System), ktoré naštartuje akceptáciu trhom a spotrebiteľmi a podporí tak vznik ďalších komerčných riešení a aplikácií. Zabezpečí sa tak aby každý občan a podnikateľský subjekt mal transparentný prístup k dátam, ktoré verejná správa o ňom eviduje.

Občan alebo podnikateľský subjekt má tiež právo vedieť, kto v rámci verejnej správy jeho údaje použil a za akým účelom. Aby bolo možné získať takéto údaje (takzvané transparentné logy), je potrebné, aby všetky informačné systémy verejnej správy boli schopné evidovať a poskytovať logy prístupov v štandardizovanom formáte.

Dôležitou vlastnosťou moderných systémov pre správu osobných údajov je manažovanie súhlasov na manipuláciu s údajmi. Občan a podnikateľský subjekt budú schopní rozšíriť prednostne nastavené oprávnenia, napríklad súhlasiť s analytickým spracovaním údajov. Dôležitým dátovým objektom je i splnomocnenie

Služba „moje dáta“ bude musieť vyriešiť aj fungovanie v systéme sektorových identifikátorov. Služba „moje dáta“ tvorí jeden z bodov pre interakciu medzi verejnou správou a súkromným sektorom.

Základné rozhodnutia

Spôsob skladania údajov pre „moje dáta“

Pohľad na údaje, ktoré sa týkajú subjektu (občana a podnikateľa) bude vytváraný na základe dopytu subjektu. Údaje tak budú v centrálnom riešení ukladané len dočasne. Údaje budú získavané zo zdrojových registrov prostredníctvom platformy integrácie údajov na základe požiadavky subjektu (vrátane riešenia prístupu k asynchrónne poskytovaným údajom). Údaje budú k dispozícii lokálne v osobnej aplikácii subjektu a v jednotlivých zdrojových informačných systémoch.

Prístup k službe „moje dáta“

Služba „moje dáta“ bude prístupná:

- ako centralizovaná aplikácia (vrátane podpory mobilnej verzie),
- cez slovensko.sk (po prihlásení sa do portfólia klienta),
- cez otvorené API.

Rozsah služby „moje dáta“

Vďaka službe budú vo verejnej správe implementované požiadavky GDPR. Služba bude postavená ako otvorená.

Typy údajov v službe „moje dáta“

Rozsah údajov bude narastať postupne:

- v prvom kroku pribudnú údaje, ktoré sú dostupné cez platformu zdieľania údajov a týkajú sa subjektu (evidencie),
- údaje o rozhodnutiach, týkajúcich sa subjektu,
- údaje o osobných preferenciách subjektu (v súvislosti s používaním elektronických služieb a interakciou s verejnou správou) a kontaktné údaje,
- údaje o aktívnych konaniach (a ich stave),
- údaje o notifikáciách a správach od inštitúcii verejnej správy,
- transparentný pohľad na prístup k údajom subjektu (kto pristupoval k údajom, za akým účelom) a manipuláciu s údajmi subjektu,
- prehľad a definovanie splnomocnení (pre každý objekt evidencie bude možné nastaviť splnomocnenia a súhlasy subjektu pre manipuláciu a pristupovaním k údajom – ak to má význam).

Funkcie pre službu „moje dáta“

Postupne budú nasadené funkcie, ktoré vytvoria pridanú hodnotu nad samotnými údajmi:

- overenie platnosti údajov – subjekt bude môcť nahlásiť problém v prípade, že evidované údaje sa mu nejaví platné,
- prístup k službám – ktoré sa týkajú objektu evidencie.

4.3.5 Analytické spracovanie údajov vo verejnej správe

Implementáciu konceptu „Data-driven state“ (teda štátu, ktorý funguje na základe využívania dát) do praxe si vyžaduje výrazné zlepšenie využívania a spracovania údajov na analytické účely. Pre jednotlivé oblasti rozhodovania budú zavedené 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 podobe pre verejnosť.

Za základné oblasti, v ktorých sa bude využívať systematická analýza dát, je možné považovať:



- **Strategické priority** – príprava stratégií, analýza súčasného stavu a porovnávanie variantov a následné sledovanie naplňovania strategických priorít,
- **Politiky** – vyhodnocovanie dopadov politík,
- **Regulácie** – Hodnotenie dopadov regulácií (Regulatory impact assessment - RIA), ktoré sa ďalej delí na:
 - hodnotenie vplyvov na podnikateľské prostredie (Business impact assessment - BIA) – vrátane sledovania dopadov na malé a stredne podniky (MSP Test),
 - hodnotenie vplyvov na životné prostredie (Environmental impact assessment - EIA),
 - hodnotenie vplyvov v sociálnej oblasti (Social impact assessment - SIA),
 - hodnotenie vplyvov na rozpočet verejnej správy (Financial impact assessment - FIA),
 - hodnotenie vplyvov na informatizáciu spoločnosti,
 - hodnotenie vplyvov na verejnú správu.
- **Investície** – vyhodnocovanie dopadov projektov a verejných investícií,
- **Výkonnosť a kvalita** – vyhodnocovanie výkonu verejnej správy pri poskytovaní služieb v jednotlivých sektoroch,
- **Podpora rozhodovania v operatívnych procesoch** (analýzy rizík, „fraud management“ a podobne).

Prístup k realizácii systematického zlepšenia využívania dát je nasledujúci:

- **Analytické jednotky:** posilnia sa organizačné kapacity – na centrálnej úrovni Dátová kancelária verejnej správy, na úrovni inštitúcií sa posilnia a vybudujú analytické jednotky (z OP EVS). Analytické jednotky sú špeciálne organizačné útvary v rámci verejnej správy (spravidla súčasti jednotlivých ministerstiev vlády SR), ktoré sa venujú tvorbe analýz.
- **Systematický prehľad o analýzach vo verejnej správe:** pre jednotlivé sektory verejnej správy, úseky verejnej správy a oblasti analýz bude možné definovať množinu analytických produktov, ktoré sa budú vytvárať, pričom sa očakávajú typy analýz ako: ex-ante vyhodnocovanie opatrení a variantov, ex-post vyhodnocovanie prijatých opatrení, online vyhodnocovanie, prognózy, simulácie, štatistické zisťovania a podobne. Je vhodné, aby sa všetky typy analýz vykonávali nad spoločnou množinou dát, ktorá má vysokú kvalitu a jednotnú ontológiu a dátový model. Bude tak možné navrhovať a testovať predpoklady jednotlivých modelov a postupne ich zlepšovať. Pre každý úsek verejnej správy tak budeme poznať:
 - aké sú analytické potreby na danom úseku,
 - aké metódy budú použité na danom úseku,
 - aké výstupy (analytické produkty) budú vytvárané (v akých intervaloch),
 - aké sú požiadavky na zdroje údajov,
 - aké sú požiadavky na analytické nástroje.
- **Analytická vrstva:** zámerom je vybudovanie jednotnej analytickej vrstvy na základe najmodernejších technológií pre ukladanie, spravovanie a analyzovanie dát. Pre sekundárne spracovanie budú prístupné údaje z informačných systémov verejnej správy, ale aj z iných zdrojov údajov podľa potreby. Údaje budú konsolidované a normalizované, v súlade s centrálnym dátovým modelom verejnej správy. Analytická vrstva tiež ponúkne nástroje na modelovanie, analýzu dát, spracovanie dát a ich vyhodnocovanie (prediktívne modely, simulácie, vizualizácie a podobne) a funkcie na manažment dátovej vrstvy. Dôležitou úlohou je tiež vyriešenie otázky použitia osobných údajov v analýzach – na základe anonymizácie a pseudoanonymizácie.
- **Zabezpečenie zberu údajov:** v prípade, že potrebné údaje nie sú vo verejnej správe k dispozícii, je potrebné nastaviť jednoduchý a efektívny spôsob, ako takéto údaje získať (spojené s reformou zberu údajov).

Základné rozhodnutia

Prístup k realizácii analytickej vrstvy

Analytická vrstva bude postavená na systéme virtuálnych DB (aby nebolo potrebné budovať centrálnu DWH riešenia). Každý systém bude poskytovať historické údaje pre potreby analýz (čo znamená aj závislosť na schémach pôvodných systémov) a možné problémy z dôvodu komplexity. Práve komplexita však bude adresovaná vďaka spoločnému modelu údajov postavenom na ontológiách.

Prístup k realizácii posudzovania vybraných vplyvov

Posudzovanie vybraných vplyvov patrí k prioritám reformy verejnej správy a je potrebné analyticky vytvárať podklady pre rôzne oblasti vplyvov. Otázkou je, akým spôsobom pripraviť IT nástroje: či pôjde o vzájomne nezávislé riešenia, alebo sa vytvorí spoločná platforma pre všetkých gestorov vybraných vplyvov. Preferujeme práve spoločné riešenie.

4.3.6 Manažment prístupov a oprávnení

Údaje o identitách

Údaje o identitách sú údaje popisujúce fyzické osoby ale aj osoby právnické, organizácie, predmety a objekty. Údaje o identitách sú veľmi rôznorodé, siahajú od jednoznačných identifikátorov, cez mená a adresy až po profily správania. Údaje o identitách predstavujú kritický prvok v každom informačnom systéme:

- V prvom rade ide o údaje, ktoré sa používajú ako prístupové účty do informačného systému. Tieto údaje sú používané na autentifikáciu používateľov informačných systémov, na autorizáciu ich operácií, pre účely auditu a pre podobné bezpečnostné účely.
- V druhom rade ide o informácie ktoré samotný informačný systém spracúva. Údaje o identitách sú v tomto prípade uložené v informačných systémoch ako bežné záznamy. Aj v tomto prípade však platí, že údaje musia byť presné, správne a aktuálne. Toto obzvlášť platí v prostrediach štátnej a verejnej správy, kde podstatná časť agendy sa týka občanov.

Údaje o identite občanov sú pre správnu činnosť systémov kritické. Informačné systémy verejnej správy musia zabezpečovať prístup úradníkov k údajom o občanoch, ale na druhej strane rovnako chrániť tieto údaje pred neoprávneným prístupom – a to najmä pred neoprávneným prístupom z vnútra organizácií štátnej správy. Bezpečnostné analýzy konzistentne označujú interné hrozby ako najväčšie. Najväčšou hrozbou pre ochranu údajov o identite občana nie je teda útočník (hacker). Najväčšou hrozbou je úradník, ktorý zneužije prístup k údajom ktorý by nemal mať.

Problematika údajov o identitách sa teda dá rozdeliť na dve oblasti:

- Konzistencia údajov. Zabezpečenie, aby údaje boli správne, úplné a aktuálne.
- Ochrana prístupu k údajom. Striktné riadenie prístupových práv.

Systémy na správu identít a prístupov

Obe tieto oblasti sú v moderných integrovaných informačných riešeniach riešené pomocou systémov na správu identít a prístupov (Identity and Access Management, IAM). Jednou z úloh systémov riadenia identít a prístupov je poskytnúť používateľovi ilúziu, že namiesto interakcie s množstvom rôznych aplikácií interaguje s uceleným informačným systémom. Cieľom IAM je zjednotiť vlastnosti týkajúce sa správy identít: zjednotenie používateľských profilov, správy prístupových práv, autentifikácie, správy autorizačných údajov, správy hesiel a podobne.

Situácia je ale komplikovaná heterogenitou aplikácií ako aj ich rôznym stupňom vyspelosti a adaptability. Dôležitým aspektom je aj ekonomická realizovateľnosť: IAM riešenie nesmie byť drahšie ako benefity ktoré prináša. Z tohto dôvodu je zriedkakedy možné zásadne meniť existujúce aplikácie aby sa prispôbili IAM riešeniu. Musí to byť práve IAM riešenie ktoré musí byť efektívne

prispôsobiteľné na existujúce aplikačné prostredie. IAM systém sa preto architektonicky skladá z niekoľkých kľúčových prvkov ktoré možno prispôbovať a kombinovať:

- Úložisko údajov o identitách (identity repository), väčšinou realizované adresárovou službou.
- Autentifikačný systém, ktorý môže obsahovať aj mechanizmus jednotného prihlásenia (Single Sign-On, SSO).
- Systém správy identít (identity management) a dozoru nad identitami (identity governance).

Tieto prvky sa v riešení navzájom dopĺňajú a spolu tvoria ucelený IAM systém.

Úložisko údajov o identitách

Ako už samotný názov hovorí, úložisko údajov o identitách (identity repository) slúži na ukladanie údajov o osobách, používateľoch, používateľských účtoch a podobných údajov priamo či nepriamo sa vzťahujúcich k osobám. Okrem toho obsahuje aj údaje vzťahujúce sa na riadenie prístupových práv: skupiny, role, privilégia, organizačnú štruktúru a podobne.

Architektonicky jednoduchším riešením je centralizované úložisko údajov. V tomto prípade sú údaje uložené len v tomto úložisku a všetky aplikácie sa úložiska na údaje dopytujú v momente keď ich potrebujú. Riešenia založené len na centrálnom úložisku údajov sú však v praxi realizovateľné len pre veľmi jednoduché riešenia. Pre komplexné riešenie sa využíva distribuované a replikované úložisko.

Úložiská údajov sa najčastejšie implementujú pomocou adresárových služieb (directory service) pre výhodné inžinierske vlastnosti týchto systémov pri práci s údajmi o identitách (škálovateľnosť, nenáročnosť na zdroje).

Autentifikačný subsystém

Autentifikačný subsystém zabezpečuje autentifikáciu používateľov, to znamená vyžiadanie a overenie „dôkazu“ o ich identite. Autentifikácia sa vykonáva na prednej strane (front-end) jednotlivých komponentov informačného systému.

Autentifikácia pomocou zjednoteného prihlasovania (SSO)

Zjednotené prihlasovanie (Single Sign-On, SSO) predstavuje riešenie zamerané na zjednodušenie prihlasovania sa používateľov do aplikácií pomocou autentifikačného a autorizačného servera. Nevyhnutným predpokladom pre vybudovanie autentifikačného servera a SSO infraštruktúry je vytvorenie centrálného úložiska údajov o identitách. Je dôležité si uvedomiť aj to, že SSO systém nijako nezjednodušuje správu používateľských účtov v aplikáciách, ktoré nepodporujú SSO – účty je aj naďalej potrebné spravovať. Preto je vhodné prepojiť SSO systém s centrálnou správou používateľov (identity management) a spravovať jeho úložisko ako ďalší z koncových systémov.

Správa identít

Systémy na správu identít (identity management) umožňujú centralizovane vytvárať, meniť a rušiť používateľské účty a ich prístupy v rámci organizácie. Ich hlavnou úlohou je zjednodušiť existujúce procesy správy účtov, ktoré sú najčastejšie manuálne, ad-hoc a neriadené a nahradiť ich centralizovaným systémom, ktorý zjednocuje správu účtov a prístupových práv v jednotlivých koncových systémoch do jednotného riešenia. Databázy používateľských účtov na koncových systémoch sa využívajú aj naďalej, natívne nástroje na správu účtov nie sú technologicky nijako obmedzené. Pre koncový systém je centrálna správa používateľov transparentná.

Jednou z najdôležitejších funkcií riešení pre správu identít je ich schopnosť implementovať organizačné a bezpečnostné politiky. Napríklad riadiť prístup k systémom na základe rolí, implementovať nezlučiteľnosť pracovných funkcií (segregation of duties) a podobne. Kritickým prvkom pre dlhodobú udržateľnosť riešenia je dohľad nad identitami (identity governance), tiež označované ako podpora pre dohľad, rizikový manažment a zhodu (governance, risk management and compliance, GRC). Tieto schopnosti zabezpečujú, že implementované politiky budú konzistentné a že bude možné ich priebežne prispôbovať zmenám prostredia, legislatívy a regulácií.

Výber prístupu

V našom koncepte potrebuje vyriešiť najmä manažment prístupu k údajom. Navrhujeme centrálné riešenie, ktoré umožní:

- Centrálné úložisko údajov o identitách (prepojené s organizačným modelom verejnej správy).
- Centrálnu správu oprávnení prístupu (ktorá umožní implementovať príslušnú biznis logiku).

K riešeniu budú mať prístup administrátori na úrovni jednotlivých inštitúcií.

Prehľad najlepších skúseností

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

Smernica INSPIRE ako príklad pre správne nastavenie pravidiel

Smernica INSPIRE sa primárne zameriava na harmonizované zdieľanie priestorových údajov a súvisiacich služieb v rámci celej Európskej únie. Smernica INSPIRE sa uplatňuje na priestorové údaje so zadefinovanou štruktúrou, ktoré uchovávajú orgány verejnej moci alebo sú uchovávané v ich mene, a na využívanie priestorových údajov orgánmi verejnej moci pri plnení ich verejných úloh. Za určitých podmienok by sa však mala uplatňovať aj na priestorové údaje, ktoré uchovávajú fyzické alebo právnické osoby, ktoré nie sú orgánmi verejnej moci za predpokladu, že tieto fyzické alebo právnické osoby o to požiadajú. Motiváciu takýchto osôb je zvyčajne vstup na trh s priestorovými údajmi.

Ontológie

- Ontológie - SEMIC odporúčené 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 odporúčené 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 náklady a zvyšuje kvalitu.

Lepšia regulácia

- e-Government podporujúci lepšie regulácie v Dánsku,
- Spôsob, akým pristúpila k vyhodnocovaniu dopadov regulácii v USA OIRA (Office of Information and Regulatory Affairs) od roku 2011 – regulatory.gov a reginfo.gov

Štatistický systém

- Európsky štatistický systém- Vízia 2020 a Akčný plán a Roadmapa pre Big Data.

Voľný pohyb dát

- Piatou slobodou Jednotného trhu je Voľný pohyb dát. V oblasti voľného cezhraničného pohybu dát sa nadviaže na poznatky v oblasti cezhraničných cloudových služieb získané zapojením Slovenska do projektu Cloud4Europe. Je potrebné vytvárať podmienky, ktoré umožnia rozvoj dátového hospodárstva a podporia lepšie rozhodovanie na základe analýz pri súčasnom aplikovaní princípů



súkromie ako súčasť dizajnu riešení („Privacy by design“). Uvažuje sa tiež o nových certifikáciách podnikov intenzívne využívajúcich osobné údaje ako aj o zavedení spoločných zásad pre tieto podniky.

5 Architektúra riešenia

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

5.1 Princípy

Zoznam architektonických princípov, na ktorých je riešenie postavené.

Všeobecné princípy

- **Údaje sú aktívom**, ktoré má hodnotu a je podľa toho riadené a spravované. Údaje sú hodnotným zdrojom VS, majú reálnu merateľnú hodnotu. V zjednodušenom ponímaní, účelom údajov je pomáhať pri rozhodovaní, podporovať rozhodovanie. Presné a včasné údaje sú rozhodujúce pre presné a včasné rozhodnutia.
- **Údaje sú zdieľané** – Používatelia majú prístup k údajom potrebným na plnenie svojich povinností, údaje sú teda zdieľané naprieč verejnou správou. Včasný prístup k presným údajom je nevyhnutným predpokladom zlepšenia kvality a efektivity rozhodovaní vo verejnej správe.
- **Údaje sú dostupné** – Údaje sú dostupné používateľom na výkon svojich úloh. Široký prístup k údajom vedie k efektívnosti a účinnosti rozhodovania a poskytuje včasnú reakciu na žiadosti o informácie a poskytnutie služieb. Prístup k údajom rovnako šetrí čas a prostriedky žiadateľa rovnako ako aj poskytovateľa služieb.
- **Údaje majú vlastníka a správcu** - Každý údajový prvok má jasne definovaného vlastníka a správcu zodpovedného za jeho kvalitu. Len vlastník údajov môže rozhodovať o ich obsahu. Použitie údajov by malo byť v rámci verejnej správy čo najjednoduchšie, pri zohľadnení legislatívnych limitov. Každý spracovateľ údajov zodpovedá za ich dôvernosť, integritu a dostupnosť.
- Aplikovanie princípu **súkromie ako súčasť dizajnu riešení** („Privacy by design“) - Každý subjekt evidencie (FO/PO) má právo na ochranu súkromia svojich údajov, najmä tým, že jeho údaje nie sú neoprávnene používané, má jednoduchý prístup ku všetkým údajom o ňom vedených, informáciu o každom použití jeho údajov a dôvode použitia, má možnosť rozhodnúť kto a za akým účelom jeho údaje používa (nad rámec legislatívne určenej povinnosti).

Referenčné údaje

- **Referenčné údaje sú kanonické a majú jasný dôvod pre svoju existenciu** - Ide o jediný právne záväzný zoznam pre špecifický typ údajov, respektíve dátového prvku.
- **Referenčné údaje reprezentujú minimálny zmysluplný dataset** - Ak je potrebné, aby boli spolu s referenčnými údajmi zdieľané aj údaje, ktoré sú ako referenčné vedené v inom referenčnom registri, nezdieľajú sa hodnoty týchto údajov ale odkazy.
- **Všetky zdieľané údaje sú postavené na štandardizovaných pojmoch a dátových prvkoch** definovaných v ontológii v súlade s výnosom o štandardoch - Pre každý objekt tak existuje vopred štandardizovaná schéma vyjadrená v XSD (vyriešia sa tiež nevyplnené položky dátového prvku, povinné a nepovinné prvky).
- **Referenčné registre sú živé zoznamy**, nielen jednoducho publikované dáta - K údajom možno pristupovať ručne alebo strojovo prostredníctvom API.
- **Referenčné údaje sú jasne kategorizované** – podľa navrhutej kategorizácie uvedenej v prílohe 8.3 Klasifikácia a kategorizácia údajov.
- **Referenčné údaje obsahujú surové údaje** (nie odvodené údaje).
- **Uniformné API na prístup k údajom** - na prístup konzumentov k referenčným údajom bude zavedené uniformné API, to je rovnaké pre všetky typy údajov. API je read only RESTful API a obsahuje bezstavové, synchrónne funkcie (primárne):

- getObj: vráť konkrétny objekt spolu s jeho referenčným URI.
- getChanged: vráť sadu objektov zmenených za obdobie spolu s ich URI - podľa typu objektu a časového intervalu.
- search: vyhľadanie objektov evidencie na základe určených kritérií, ktoré môžu byť špecifické pre každý typ objektu evidencie.

Dátový model a dátové prvky

- **Identifikácia pomocou Jednotného referencovateľného identifikátora (URI)** - Identita objektu je vyjadrená aj pomocou Jednotného referencovateľného identifikátora. Pre zdieľanie objektov je toto primárny identifikátor, to je. všetky objekty, a aj prepojenia medzi objektami sa vyjadrujú pomocou RefId. Toto pravidlo sa použije aj objektmi jedného registra navzájom. Kľúčovým rysom je centrálna správa týchto URI identifikátorov, ktorá bude zabezpečovaná prostredníctvom MetaIS, ktorý bude mať jedinú kompetenciu vydávať referenčný URI identifikátor.
- **Dátový model:** Údaje sú jednotne popísané prostredníctvom ontológií. Koncepty a vzťahy medzi nimi sú konzistentne definované v celej VS pre danú problematiku a definície sú zrozumiteľné a k dispozícii. Koncepty a vzťahy, ktoré budú reprezentované v informačných systémoch, musia mať spoločnú definíciu v celej verejnej správe v snahe uľahčiť spoluprácu a zdieľanie informácií a vedomostí. Okrem toho je tento princíp nevyhnutný na efektívnu definíciu rozhraní systémov, výmenu a publikáciu údajov a na odstránenie neurčitosti pri interpretácii pojmov.

Bezpečnosť údajov

- **Bezpečnosť údajov** - Údaje sú chránené pred neoprávneným prístupom, manipuláciou, použitím a zverejnením (zachovanie dôvernosti údajov), ich úmyselnou alebo neúmyselnou modifikáciou (zachovanie integrity údajov), pričom sú dostupné v požadovanom čase a v požadovanej kvalite (zachovanie dostupnosti údajov).
- **Pravosť údajov** - Používateľ pracuje len s údajmi, ktorých hodnovernosť a pôvod sú zabezpečené napríklad ich autorizáciou, a ktoré sú z dôveryhodného zdroja s garantovanou identitou.

Priestorové údaje

Princípy pre priestorové údaje sú dané predovšetkým Smernicou INSPIRE pre infraštruktúru priestorových informácií a to nasledovne:

- Infraštruktúra musí byť navrhnutá tak, aby sa zabezpečilo, že priestorové údaje sa uchovávajú, sprístupňujú a udržiavajú **na najvhodnejšej úrovni**.
- **Kombinácie priestorových údajov** - Musí byť možné kombinovať priestorové údaje z rôznych zdrojov v rámci celej Európskej únie.
- **Zdieľanie priestorových údajov** - Priestorové údaje musia byť schopní zdieľať viacerí používatelia a aplikácie.
- Priestorové údaje zhromaždené na jednej úrovni orgánu verejnej moci musí byť možné zdieľať s inými orgánmi verejnej moci (v Slovenskej republike je ekvivalent „orgánu verejnej moci“ pojem „povinná osoba“).
- Nesmú existovať neodôvodnené obmedzenia priestorových údajov na ich široké využitie.
- Musí byť funkčný jednoduchý spôsob vyhľadávania existujúcich priestorových údajov, vyhodnotenia ich účelnosti a podmienok uplatniteľných na ich použitie.

Kľúčovým pri riešení problematiky priestorových údajov bude využívanie RPI ako jednotného konzistentného zdroja údajov. Súčasná podoba RPI však zabezpečuje len čiastočné naplnenie legislatívnych požiadaviek z oblasti INSPIRE (meta-údaje a časť sieťových služieb), preto bude potrebné pre naplnenie vyššie uvedených princípov zabezpečiť zostávajúce legislatívne požiadavky a technické odporúčania. Okrem iného bude RPI v sebe obsahovať nástroje na hodnotenie kvality priestorových údajov a služieb priestorových údajov, nástroje na identifikáciu duplicitnej či multiplicitnej správy priestorových údajov a podobne.

5.2 Biznis vrstva

5.2.1 Prehľad požiadaviek

Požiadavky na inštitúcie verejnej správy

- Každá inštitúcia organizačne zabezpečí oblasť manažmentu údajov – stanoví vlastníkov dát a dátových kurátorov.
- Každá inštitúcia, ktorá spracováva údaje, zabezpečí a zavedie riadenie kvality údajov v rámci manažérskych funkcií.
- Každá inštitúcia, ktorá spracováva údaje má mať definované kritériá dátovej kvality pre svoje spracovávané údaje (na základe požiadaviek centrálnej úrovne).
- Každá inštitúcia pravidelne vyhodnocuje dátovú kvalitu ním spracovaných údajov,
- Každá inštitúcia proaktívne odstraňuje prípadné nedostatky v kvalite údajov.
- Každá inštitúcia proaktívne hľadá príčinu prípadných nedostatkov v dátovej kvalite a eliminuje ich výskyt.

Prístup k údajom pre subjekty

Subjekty môže pristupovať k údajom, ktoré o nej verejná správa eviduje:

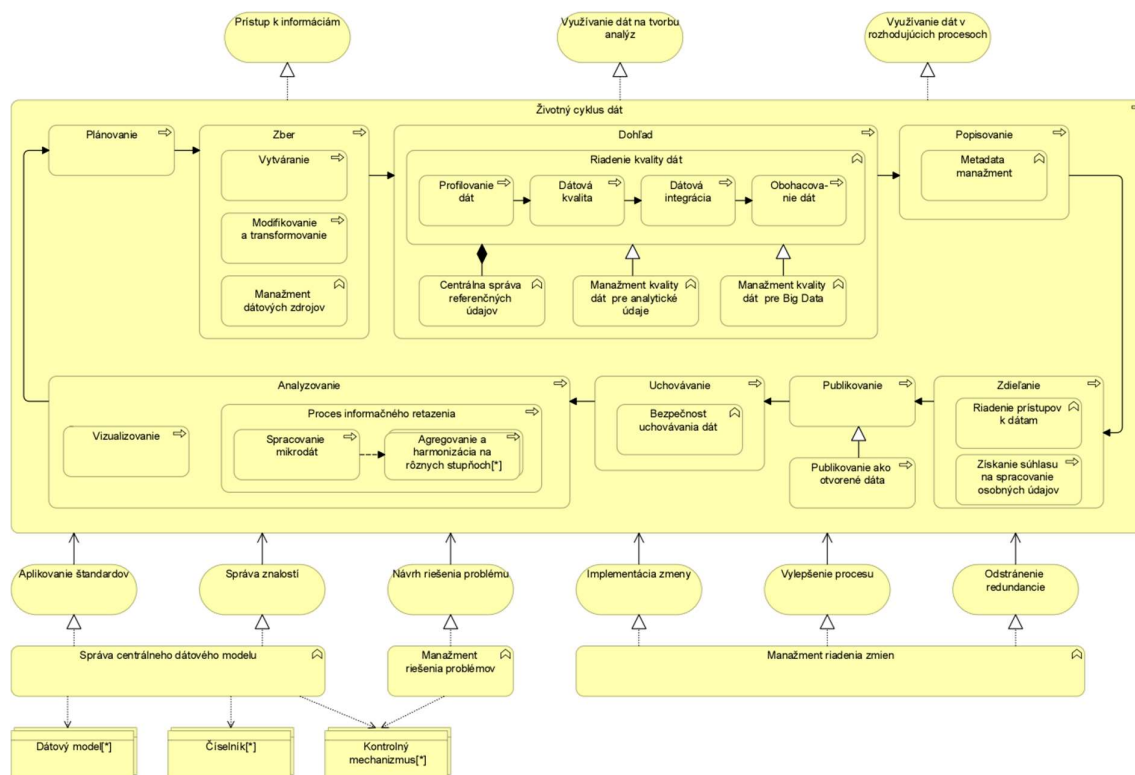
- K dispozícii je prehľad všetkých dátových objektov, ktoré sa týkajú subjektu na jednom mieste a prístup k ich obsahu.
- O každom dátovom objekte je možné zistiť ako s ním bolo manipulované (dátum vzniku, história modifikácií, kto z verejnej správy k objektu pristupoval a prečo).
- Každému dátovému objektu je možné nastaviť úroveň zdieľateľnosti podľa zaradenia daného objektu (z hľadiska možnosti prístupu k údajom a ich ďalšieho použitia pre konkrétnu inštitúciu verejnej správy alebo mimo verejnej správy).
- Subjekt poskytuje nové údaje verejnej správe len raz, nie každej inštitúcii samostatne.

5.2.2 Prehľad procesov

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

5.2.2.1 Riadenie dát

Obrázok 3: Biznis procesy pre manažment údajov



Tabuľka 9: Procesy pre životný cyklus údajov

Proces	Centrálné	Lokálne
Životný cyklus údajov:		
– Plánovanie	Metodika a centrálné plánovanie	Plánovanie na úrovni inštitúcie
– Zber:		
– Identifikovanie dátových zdrojov.	Realizácia	
– Poverenie vlastníkov, ktorí zbierajú a aktualizujú dáta.	Realizácia	
– Zavedie validačných pravidiel pre dáta a biznis pravidiel pre vytvorenie, zmenu alebo zber údajov.	Metodika	Realizácia
– Popisovanie	Metodika	
– Dohľad – Manažment kvality dát	Realizácia	
– Zdieľanie:		



Proces	Centrálne	Lokálne
– Definovanie pravidiel bezpečnosti prístupu a ich dodržiavanie.	Realizácia	
– Vytvorenie štandardného dopytu a reportov.	Metodika	Realizácia
– Poskytovanie funkcionality používateľom.	Realizácia	
– Identifikácia závislostí		Pre každú ISVS/agendu bude existovať popis požiadaviek na externé údaje/potvrdenia/prepojenia. Závislosti sa identifikujú čím skôr aby bolo možné dobre riadiť synchronizáciu. Primárne sa tieto závislosti popíšu už v KRIS a budú uložené v centralizovanej evidencii (MetaIS), vďaka čomu budú tieto evidencie verejne dostupné
– Manažovanie využívania systému.	Realizácia	
– Monitorovanie výstupnej dátovej kvality.	Metodika	Realizácia
– Poskytovanie vhodných metadát.	Metodika	Realizácia
– Publikovanie:		
– Zaevidovanie do katalógov ako Katalóg otvorených dát a do registrov	Realizácia	
– Nastavenie procesov pre vytváranie aplikačných rozhraní (API)	Realizácia	
– Uchovávanie:		
– Vytvorenie a dodržiavanie pravidiel uchovávaní.	Metodika	Realizácia
– Vymazávanie dát v súlade s biznis pravidlami.	Metodika	Realizácia
– Analyzovanie.		

Tabuľka 10: Štandardizácia a centrálny dátový model

Proces	Centrálné	Lokálne
Štandardizácia dátových prvkov - údaje sa vymieňajú v štandardizovaných jednotkách - dátových prvkoch. Štandardizácia je dôležitá najmä v procese vyhlasovania referenčných údajov, ale týka sa údajov všeobecne. Pred vyhlasovaním referenčných údajov je potrebné v procese štandardizácie nastaviť a určiť formát a štruktúru pre relevantné dátové prvky (odporúča sa čím skôr).	Realizácia	
Správa centrálného modelu údajov , v rámci ktorej vznikne tiež kompetencia na riešenie otázok interoperability, ako aj na presadzovanie štandardov na pomenovávanie údajov a na obsah jednotlivých polí (číselníky, prípustné hodnoty, kontrolné mechanizmy).	Realizácia	

Tabuľka 11: Procesy pre manažment kvality údajov

Proces	Centrálné	Lokálne
Manažment kvality údajov - definuje role, zodpovednosti, pravidlá a procedúry spojené so získavaním, spravovaním, šírením a usporadúvaním dát. Pre jeho úspech je nevyhnutná spolupráca biznis (určujú biznis pravidlá, ktorými sa riadi kvalita dát, za ktorú zodpovedajú) a technických (vytvárajú a spravujú technologické prostredie, jeho architektúru, systémy a databázy) rolí zamestnancov. Požiadavky na dátovú kvalitu musia byť presne stanovené a za ich dodržiavanie ako aj určenie stupňa súladu musí byť identifikovaná zodpovedná osoba. Stupeň súladu určuje mieru tolerovania chýb. Manažment kvality dát prebieha v štyroch fázach:	Metodika kvality údajov: definovanie kritérií kvality a ukazovateľov kvality	Realizácia
– Profilovanie dát – ide o proces, v ktorom sa získa porozumenie toho, ako existujúce dáta korešpondujú s nastavenými kritériami kvality.	Metodika a dohľad	Realizácia



Proces	Centrálne	Lokálne
<ul style="list-style-type: none">– Riešenie dátovej kvality – na základe poznatkov z profilovania je dôležité prísť na problémy a ich príčiny. Napríklad kvôli nedostatočným opravným mechanizmom alebo zle nastaveným biznis procesom sa môžu dáta duplikovať. Dôležité je poznamenať, že v tomto kroku sa vylepšuje kvalita už existujúcich dát, ale neriešia sa príčiny problémov. Na riešenie príčin je nutné naštartovať projekty, ktoré napríklad vylepšia biznis procesy alebo nasadia vylepšené nástroje pre zber dát. Možné riešenia problémov v dátovej kvalite zahŕňajú:<ul style="list-style-type: none">– Vylúčiť dáta, hlavne pri ich kopírovaní do dátových skladov, ak sú problémy príliš vážne a neopraviteľné;– Prijatť dáta s chybami, ak sú v tolerovateľných rozmedziach;– Opraviť dáta, napríklad vyhľadáním referenčného dátového záznamu a zjednotením ostatných záznamov, alebo aplikovaním ďalších postupov dátovej vedy;– Vložením prednastavenej hodnoty, napríklad „neznámy“ alebo „nedefinovaný“.	Metodika a dohľad	Realizácia
<ul style="list-style-type: none">– Integrácia údajov a postupné zjednocovanie dátovej vrstvy verejnej správy.	Realizácia	
<ul style="list-style-type: none">– Obohacovanie dát - môže ísť o obohacovanie dát o danom subjekte evidencie o informácie z iných databáz v rámci rezortu alebo z iných rezortov alebo z externých systémov. Kľúčové pre tento krok je vedieť nájsť také parametre subjektu, na základe ktorých sa dá s istotou učiť, že ide o informácie o tom istom subjekte = stotožnenie subjektov evidencie. Takémuto kombinovaniu dát extrémne napomáha aj riadenie metadát, ako je popísané nižšie.	Metodika a dohľad	Realizácia
Manažment kvality dát pre analytické údaje	Metodika a dohľad	Realizácia
Manažment kvality dát pre Big Data (zo senzorov, internet of things, industry 4.0)	Metodika a dohľad	Realizácia

Tabuľka 12: Procesy MDM

Proces	Centrálne	Lokálne
Metadata manažment - umožňuje riadenie popisných údajov o dátach. Možno aplikovať rôzne štandardy a procesy podľa typu dát, napríklad pre štatistické dáta a dáta z oblasti sociálnych vied - Data Documentation Initiative (DDI) ¹³ , ISO-TS 17369 Statistical Data and Metadata Exchange (SDMX) ¹⁴ , Dublin Core Metadata Initiative (DCMI) ¹⁵ , ISO/IEC 11179 ¹⁶ . Jeden štandard pre metadáta však nepokryje celý životný cyklus dát. Dôležité je vybudovať distribuovaný systém na tvorbu a správu metadát, pretože informácie o dátach vznikajú pri rôznych situáciách v rôznom čase a priestore. Metadáta musia byť referencovateľné a musí byť možné opakovane ich použiť cez referenciu. Harmonizácia aktivít smerom k tvorbe vysoko kvalitných metadát si vyžaduje najlepšie praktiky, metodiky a disciplínu. Hoci sa väčšina práce dá vykonať automaticky alebo poloaautomaticky, často je tiež potrebné zasiahnuť manuálne a dohľadať chýbajúci kúsok znalosti o dátach.	Metodika a dohľad	Realizácia
Riadenie identifikátorov - Správa a automatizované prideľovanie URI - slúži na registráciu, schvaľovanie a publikáciu všetkých referenčných URI identifikátorov dátových prvkov (ontológie, objektové a dátové vlastnosti), katalógov a datasetov (číselníky, hierarchie), a vo forme URI šablón aj individuí a definícií s podporou publikovaných služieb MetaIS.	Realizácia	
Správa referenčných údajov a správa číselníkov – zavedenie procesov a pravidiel pre zber, stotožňovanie, konsolidáciu, ukladanie a distribúciu kritických a referenčných údajov, umožní zefektívnenie a zvýšenie kvality poskytovaných služieb, odstráni redundanciu a nekonzistentnosť údajov.	Realizácia	

¹³ <http://www.ddialliance.org>¹⁴ <http://sdmx.org/>¹⁵ <http://dublincore.org/>¹⁶ <http://metadata-standards.org/>

Proces	Centrálne	Lokálne
Plánovanie a skoré vyhlasovanie referenčných registrov - pre správcov ISVS je mimoriadne dôležité vedieť, kedy ktoré údaje "začnú" byť referenčné. Vytvorí sa dlhodobý plán vyhlasovania referenčných údajov: ktoré objekty (a ich atribúty), z ktorých evidencií budú vyhlásené za referenčné, kedy je plánované schválenie týchto údajov za referenčné. plán bude priebežne aktualizovaný tak, aby odrážal realitu.	Metodika a dohľad	Realizácia
Manažment riešenia problémov – nastavenie procesov, pravidiel a roli umožní urýchliť vyriešenie incidentov.	Metodika a dohľad	Realizácia
Manažment riadenia zmien – realizáciou procesov riadenia zmien bude zabezpečené minimalizovanie implementácie biznis procesov zbierajúcich duplicitné dáta a spravovanie redundancie dát. Poskytne konkrétny postup a nástroje pre hlásenie a riešenie chýb v referenčných údajoch.	Metodika a dohľad	Realizácia

Tabuľka 13: Procesy pre manažment prístupov a oprávnení

Proces	Centrálne	Lokálne
Riadenie prístupov k dátam - jedná sa o nastavenie procesov pre ochranu osobných údajov, obchodného tajomstva a utajovaných skutočností, ako aj o udeľovanie oprávnenia k prístupu rôznym roliam.	Metodika a dohľad	Realizácia
Bezpečnosť uchovávaní dát – jedná sa o nastavenie procesov pre zálohu a obnovu dát ako ochranu pred stratou alebo poškodením, archivovanie dát a ich prípadné odstraňovanie na základe presne daných pravidiel.	Metodika a dohľad	Realizácia

5.2.2.2 Informačné reťazenie údajov

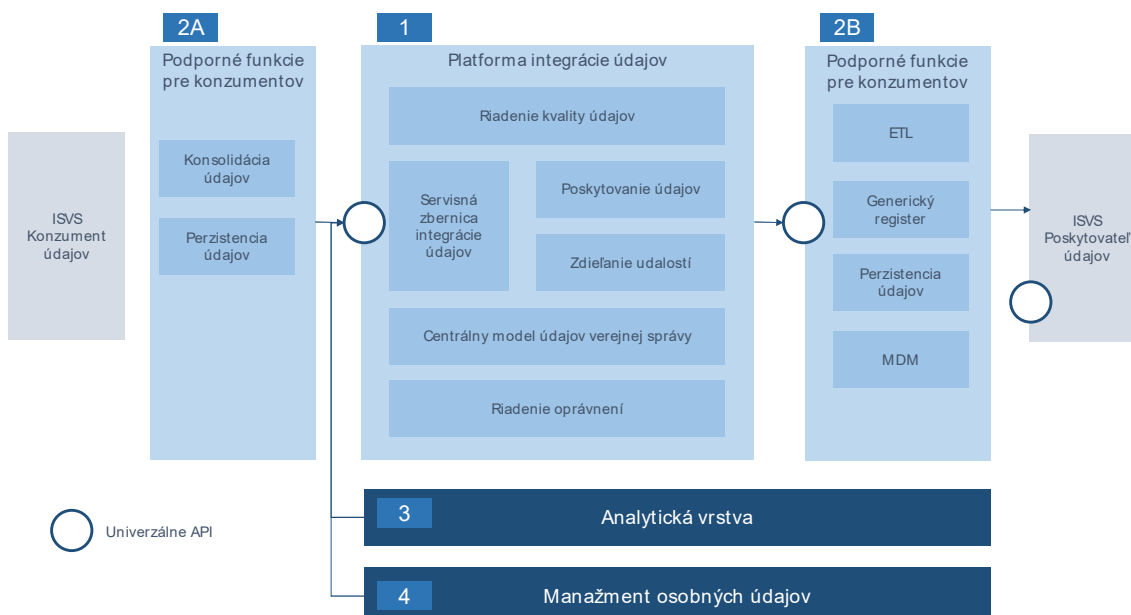
Proces	Centrálne	Lokálne
Proces informačného reťazenia je dôležitý predovšetkým pre procesy štatistického zisťovania a pre tvorbu analýz. Popisuje tok dát od mikrodát (získaných napríklad z transakčných databáz, senzorov a prieskumov) cez rôzne stupne procesov agregácie a harmonizácie v hierarchii primárnych a sekundárnych používateľov. Tieto procesy agregácie musia byť presne zdokumentované pre sekundárnych používateľov. Je veľmi dôležité presne zadať role zamestnancov, ktorí majú prístup k mikrodátam a surovým dátam, aby nedochádzalo k obmedzovaniu znalostí, ktoré môžu z dát získať (keďže proces agregácie a harmonizácie je spravidla nevratný, každý takýto stupeň zakrýva detail pôvodne zozbieranej informácie a znemožňuje tak aplikovať ľubovoľné dostupné analytické procesy a nástroje).	Metodika a dohľad	Realizácia

5.3 Aplikačná vrstva

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

5.3.1 Koncept architektúry integračnej platformy

Obrázok 4: Koncept pre platformu zdieľania údajov a integrácie údajov



Platforma integrácie údajov vytvára priestor na zdieľanie údajov medzi konzumentami údajov a poskytovateľmi (producentmi údajov). Konzument získava objekty evidencie v štruktúre podľa kontextu procesu, v ktorom sa nachádza.

Platforma sa skladá z dvoch základných vrstiev:

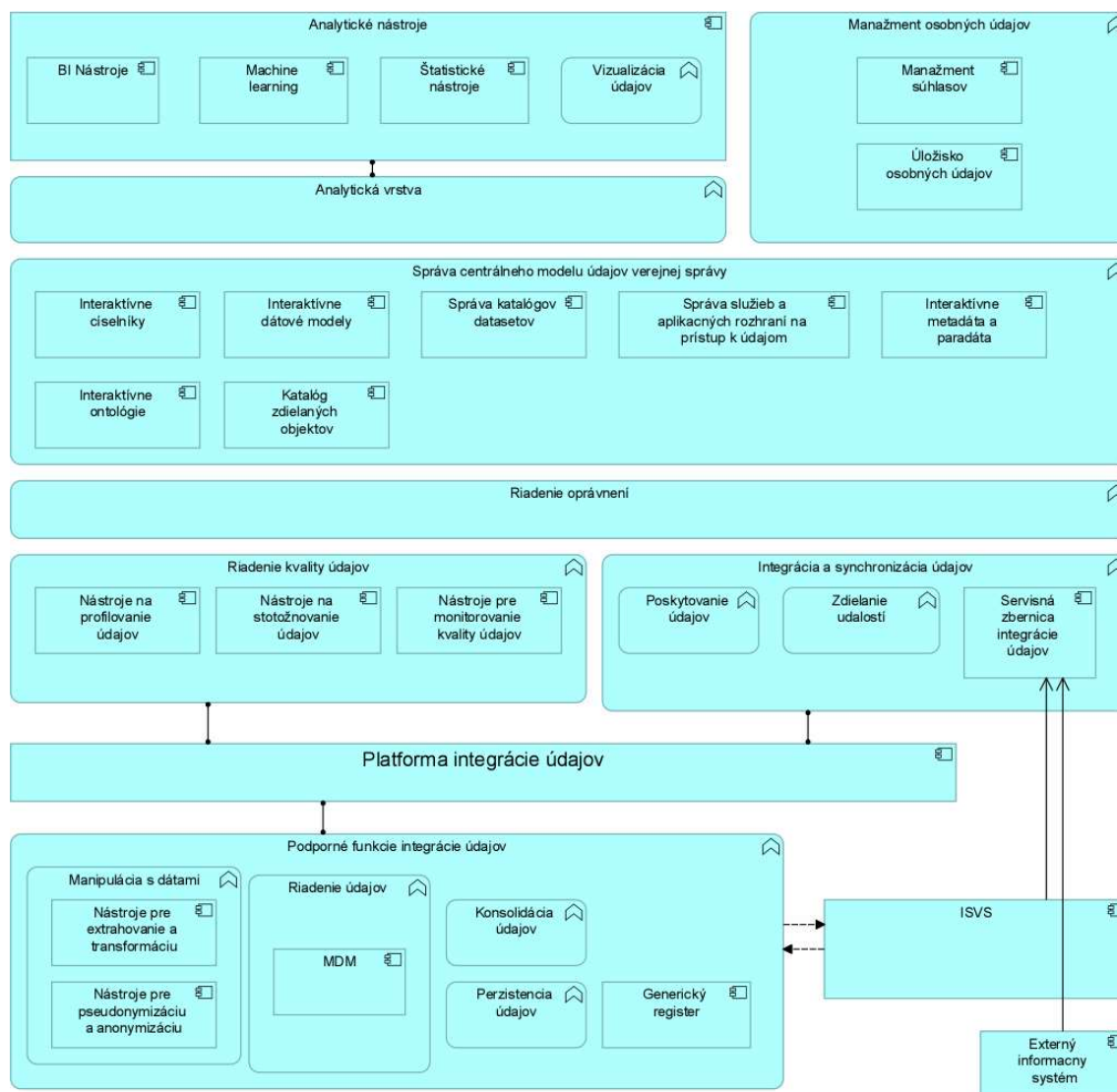


- **Centrálne funkcie pre zdieľanie a integráciu údajov (dostupné ako PaaS, prostredníctvom univerzálneho API):**
 - Manažment prístupu k údajom (cez modul riadenia oprávnení a katalóg zdieľaných objektov, dátové schémy sú evidované v centrálnom modeli údajov VS),
 - Získavanie údajov (postupné skladanie údajov pre objekt evidencie z jednotlivých informačných systémov, zabezpečenie smerovania požiadaviek zo spoločného prístupového bodu na správny systém, v ktorom sú údaje poskytované).
- **Podporné funkcie pre integráciu údajov (dostupné ako SaaS):**
 - Pre poskytovateľov údajov (pripojenie k dátovým službám zdrojového informačného systému)
 - zabezpečujú najmä synchronizáciu údajov, ETL, perzistenciu údajov, či realizáciu rozhraní. Špeciálnou podpornou službou je Generický register.
 - Pre konzumentov údajov – konsolidáciu údajov a perzistenciu údajov.

Platforma je otvorená a umožňuje vytvárať (pridávať) ďalšie služby a rozšírenia, napríklad pre kvalitu údajov, čistenie údajov, synchrónnu komunikáciu, manažment osobných údajov (služba Moje dáta), analytické spracovanie údajov a podobne – pozri aplikačné komponenty a ich služby.

5.3.2 Prehľad aplikačných funkcií a komponentov

Obrázok 5: Prehľad aplikačných funkcií a komponentov



Platforma integrácie údajov – centrálné služby

Je integračným prvkom na úrovni vrstvy údajov. Umožňuje automatizovanú komunikáciu a zdieľanie údajov medzi jednotlivými informačnými systémami verejnej správy. Ide o centrálny komponent povinne používaný na prístup k údajom. Za jeho prevádzku zodpovedá správca MUK dátová časť.

V rámci riešenia budú využívané centrálné funkcie:

- Servisná zbernica integrácie údajov – sprístupňuje funkcie API pre konzumentov údajov, vykonáva autorizáciu požiadavky, na základe typu objektu, identity producenta a konzumenta a identity údajov z komponentu správy prístupových oprávnení. Nijako nemanipuluje obsah ani formát požiadavky ani odpovede.
- Poskytovanie údajov - zabezpečenia smerovanie požiadaviek zo spoločnej zbernice (spoločného prístupového bodu) na správny systém, v ktorom sú údaje poskytované.

- Riadenie kvality údajov, sledovanie kvality údajov a sledovanie riešenie identifikovaných problémov, plánovanie, nasadenie, monitoring, tvorbu opatrení.

Podporné funkcie pre konzumentov údajov

- Konsolidácia údajov - umožňuje z dátových objektov dostupných prostredníctvom platformy zdieľania údajov „vyskladať“ dátové objekty špecifického formátu a štruktúry.
- Perzistencia údajov - zabezpečuje uchovanie údajov. Na strane konzumenta údajov je týmto spôsobom možné zabezpečiť uchovávanie údajov pre neskoršie použitie, čím sa aj zníži záťaž platformy zdieľania údajov.

Podporné funkcie pre poskytovateľov údajov

- Nástroje pre manipuláciu s údajmi a transformácie:
 - ETL – Extract, Transform, Load - umožňuje pripojenie na existujúce aplikačné alebo databázové rozhrania producenta údajov vytvorí potrebné štruktúry údajov a API pre zapojenie poskytovateľa do platformy zdieľania údajov.
 - Nástroje pre anonymizáciu a pseudoanonymizáciu.
- Perzistencia údajov - zabezpečuje uchovanie údajov. Na strane producenta údajov je týmto spôsobom možné zabezpečiť riešenie kapacitných nárokov na zdieľanie údajov alebo umožniť on-line prístup k údajom platných k určitému času v minulosti tam, kde sú aktuálne údaje dostupné iba asynchrónne.
- Profilovanie dát a stotožňovanie dát, nástroje MDM.
- Generický register - umožňuje poskytovateľovi údajov viesť určitú evidenciu/register pomocou centrálne poskytovanej služby prispôbenej jeho požiadavkám. Obsah a funkcie generického registra presahujú oblasť zdieľania údajov, tento komponent je potrebné detailnejšie špecifikovať v príslušnej strategickej prioritě (napr. spoločné moduly).

Centrálny model údajov verejnej správy (súčasť MetaIS)

Modul umožní automatizovane pridelovať referenčné URI pre všetky údaje, informačných systémov verejnej správy, vrátane nástrojov pre vyhľadávanie a lokalizovanie dát a dátových zdrojov. Obsahuje katalogizované údaje o poskytovateľovi, evidencii/registri a type poskytovaných údajov, vrátane definície dátového objektu vo forme prepojených údajov. Položky modelu sú schvaľované štandardizovaným procesom. V rámci modelu sú tiež definované číselníky, ontológie, datasety a aplikačné rozhrania služieb.

Pre jednu agendu môže existovať viacero typov poskytovaných objektov, či už delením záznamu na samostatné časti, alebo aj vzájomne sa typovo prekrývajúce objekty (napríklad v prípade odlišenia autorizácie na prístup k údajom). Model tiež obsahuje informáciu o poskytovaných typoch notifikácií pre určitý typ objektu a prípustné kritériá (parametre volania funkcie) vyhľadávacích služieb.

Modul riadenia oprávnení

Modul riadenia oprávnení umožňuje nastavenie prístupov užívateľov podľa navrhutej klasifikácie dát (na základe ich citlivosti) a používateľov (na základe ich zaradenia, role).

Báza záznamov o oprávneniach na prístup k údajom. Štandardne sa oprávnenia definujú pre určitého konzumenta a typ poskytovaných údajov, štandardizovaným procesom v gescii správcu MÚK, na základe súhlasu producenta údajov. Ak niektorý poskytovateľ údajov potrebuje použiť detailnejšie riadenie prístupu, môže si ho implementovať interne.

Unifikované aplikačné rozhranie

Obsahuje sadu funkcií používaných na prístup ku všetkým typom objektov. Rovnaká sada funkcií môže byť implementovaná na strane poskytovateľa údajov (respektíve v zastúpení podpornými službami) a spoločným prístupovým bodom.

Rozhranie obsahuje nasledovné typy funkcií:

- Prístupové – na základe referencovateľného identifikátora objektu poskytne jeho obsah v štruktúre podľa katalógu zdieľaných objektov.
- Synchronizačné – poskytne zoznam identifikátorov zmenených objektov od zadaného timestamp.
- Vyhľadávacie – poskytne zoznam identifikátorov kandidátov spĺňajúcich zadané kritériá.
- Notifikačné – „push“ notifikácie o zmene v určenom type objektov.

Mechanizmus pre zdieľanie rozhodnutí

Mechanizmus pre zdieľanie rozhodnutí je modul, ktorý slúži na evidenciu meta-údajov, ktoré boli vykonané vo verejnej správe. Poskytuje služby na evidenciu rozhodnutia (zápis meta-údajov), vyhľadávanie rozhodnutí, rozhodnutia týkajúce sa subjektu (vrátane vyhľadania v zdrojovom systéme respektíve elektronickom spise).

Manažment osobných údajov

Modul umožňuje získať prístup k osobným údajom pre subjekt. Využíva pritom nástroje pre integráciu údajov. Modul implementuje službu „Moje dáta“ a pre svoju funkčnosť využíva funkcie:

- Manažment súhlasov – umožňuje nastaviť preferencie používateľa pre použitie, zdieľanie a publikáciu jednotlivých objektov.
- Úložisko osobných údajov – dočasné úložisko osobných údajov prístupné pre potreby používateľa.

Analytická vrstva

Analytická vrstva (ktorá pristupuje k dátam z agendových informačných systémov - umožňuje ukladanie, manipulovanie a riadenie dát a Big Data v rôznych typoch databáz: Relačné databázy (SQL), NoSQL databázy (dokumentovo-orientované, sklady s „key-value“, inšpirované veľkými tabuľkami alebo grafovo orientované), New SQL.

Riešenie umožní čítať údaje priamo zo zdrojových systémov. Analytická vrstva obsahuje ďalšie komponenty pre prácu s dátami napríklad ETL podobne.

Analytická vrstva poskytuje rôzne úrovne skladovania dát, ktoré možno poskytovať ako klasické data marty podľa doménových členení, v ktorých budú prebiehať analýzy dát špecifické pre jednotlivé domény. Údaje možno prehľadávať vďaka katalógom metadát a vyhľadávacím nástrojom. Analytická vrstva Bezpečnosť sa stará o súkromie údajov (vďaka anonymizácii a pseudonymizácii, kryptografii a úrovniam prístupu s rôznou granularitou), správu dát (transakčné logy, audity na najnižšej úrovni granularity, data provenance) a ich integritu vďaka monitoringu v reálnom čase, validácii koncového bodu a filtrovaniu.

Analytická vrstva bude tiež slúžiť pre spracovanie veľkého množstva neštruktúrovaných dát napríklad zo sociálnych sietí, alebo IoT zariadení.

Analytické nástroje

Analytické nástroje umožňujú vykonávať analýzy nad dávkovými údajmi aj nad dátovými tokmi pre operatívne analýzy. Poskytuje:

- tradičné BI nástroje, ktoré podporujú analytické funkcie nad klasickými štruktúrovanými dátovými zdrojmi (faktov),
- machine learning – sada nástrojov, ktoré podporujú spracovanie údajov metódami umelej inteligencie,
- štatistické nástroje, ktoré umožňujú tvorbu štatistických dátových modelov, testovanie hypotéz, faktorové analýzy, korelácie, regresie a podobne,
- nástroje pre simulácie, ktoré umožňujú simulovať historický aj budúci priebeh modelovaných udalostí nad vybranou množinou historických a aktuálnych údajov.

Kolaboračné nástroje (platforma pre posudzovanie vplyvov)

Kolaboračné nástroje zabezpečia najmä:

- služby zdieľania, na základe ktorých je možné zdieľať výsledky analýz alebo analytický model nad definovanou množinou dát z analytickej vrstvy.
- služby kolaborácie, pomocou ktorých môžu analytici spolupracovať nielen interne na tvorbe analytického modelu, ale aj s externými špecialistami.
- služby vizualizácie výsledkov, ktoré umožňujú vidieť výsledky analýz v interaktívnych grafoch vrátane ich interpretácie a publikácie,

5.4 Technologická vrstva

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

6 Realizácia

Kapitola sa venuje návrhu praktických krokov, ktoré zabezpečia realizáciu navrhutej 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).

6.1 Riziká a otvorené otázky

Počas realizácie programu manažmentu údajov bude potrebné adresovať nasledujúce riziká a otvorené otázky.

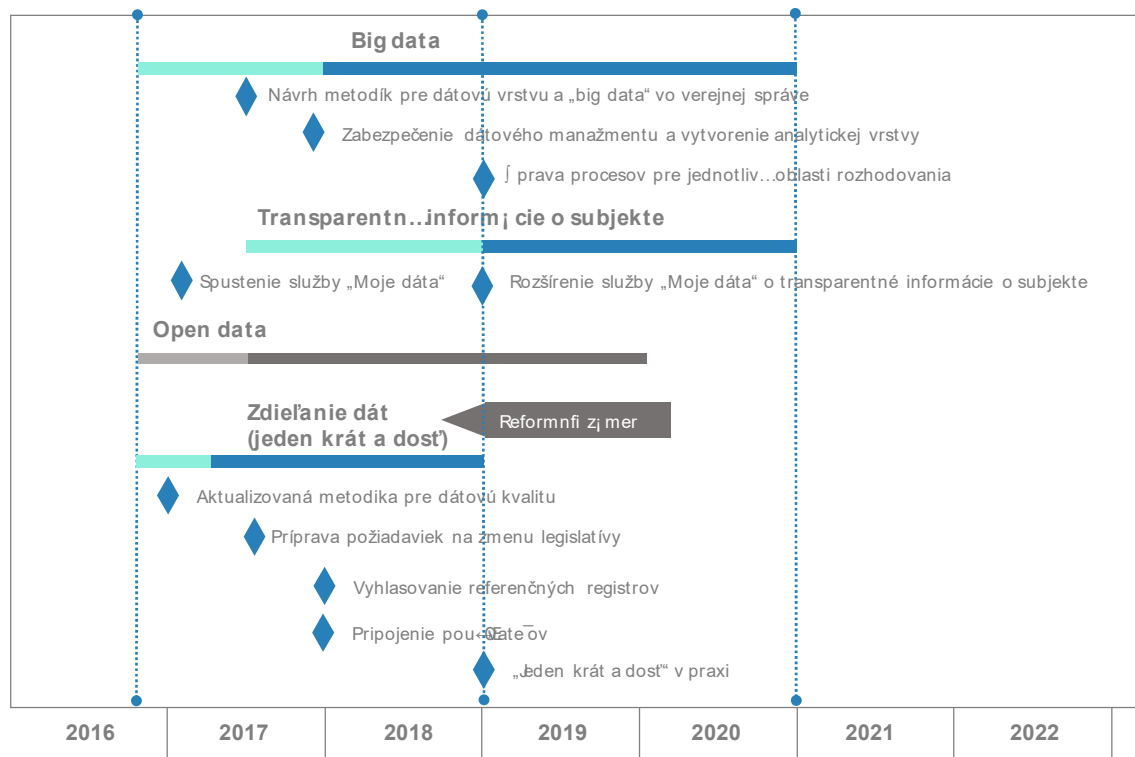
Tabuľka 14: Prehľad rizík/otvorených otázok a spôsobov ich zmiernenia

Riziko/otvorená otázka	Spôsob zmiernenia
Ako zabezpečiť otvorený, participatívny a inkluzívny proces manažmentu údajov	Fungovanie programu Manažment údajov bude nastavené tak, aby každý kto má záujem (zo strany inštitúcií verejnej správy, neziskového sektora, či záujmových združení) mohol prispieť.
Verejná správa nebude mať dostatok energie a zdrojov plošne zvyšovať kvalitu údajov	<ul style="list-style-type: none">– Získanie politickej podpory pre program Manažment údajov.– Obsadenie kľúčových pozícií motivovanými osobami.– Zabezpečenie pozitívnej motivácie (zdroje, prestíž).– Zabezpečenie negatívnej motivácie (dôsledne pripravená legislatíva, ktorá nastaví systém dohľadu a dozoru).– Dôsledné poznanie a monitorovanie stavu a postupu v dosahovaní cieľov.
Zabezpečenie dôslednej legislatívnej prípravy - Pripravenosť legislatívy poskytovateľa údajov na zdieľanie jeho dát, modifikácie údajov v registri v zmysle referencovania a podobne	Príprava legislatívy sa stane integrálnou súčasťou programu Manažment údajov. Detailné posudzovanie zmien v legislatíve z pohľadu manažmentu údajov vo verejnej správe (vznik, kvalita, štruktúra, G2G väzby, prístup/poskytovanie/zverejňovanie údajov).
Navrhnuť správny spôsob financovania a realizácie - realizačný a finančný model pre zavedenie princípu 1x a dosť, ktorý vyžaduje koordináciu a súčinnosť ÚPPVII, poskytovateľov a konzumentov dát	Navrhovaný mechanizmus umožní financovať: <ul style="list-style-type: none">– Zmenu procesov a fungovania.– Posilnenie interných kapacít.– Riešenie integrácie jednotlivých inštitúcií verejnej správy.– Vybudovanie centrálnych komponentov.
Zabezpečenie dôslednej ochrany údajov	Na bezpečnosť musí byť kladený vysoký dôraz.

Riziko/otvorená otázka	Spôsob zmiernenia
Chýbajúce referenčné údaje pre orgány verejnej moci	<p>ÚPPVII očakáva skoré vyhlásenie referenčných údajov orgánov verejnej moci (OVM).</p> <p>Keďže pred rokom Štatistický úrad SR avizoval, že predmetné RÚ budú vyhlásené do konca roka 2015, neskôr oznámil, že do marca 2016.</p> <p>Centrálny metainformačný systém verejnej správy je integrovaný prostredníctvom CSRÚ na RPO tak, že pri zakladaní povinnej osoby v MetaIS po zapísaní IČO sa používateľovi dotiahnu do MetaIS všetky ostatné údaje od OVM, a tiež to isté platí pri evidovaní ich podriadených organizácií.</p> <p>Nielen MetaIS ale aj celá verejná správa očakáva skoré vyhlásenie RPO - referenčných údajov orgánov verejnej moci.</p>

6.2 Plánovanie a migrácia (v súlade s akčným plánom)

Obrázok 6: Akčný plán podľa NKIVS



Kľúčové míľniky

- Vo všetkých konaniach zavedený princíp “jedenkrát a dost” (do konca roka 2018).
- Monitoring výkonnosti poskytovania elektronických služieb (už povinné pre všetky orgány verejnej moci).
- Otvorené údaje (do konca roka 2019).

- Transparentnosť využívania údajov subjektu (koniec roka 2020).

6.2.1 Aktivity a úlohy pre centrálnu úroveň

Tabuľka 15: Aktivity a úlohy pre centrálnu úroveň (ktorú reprezentuje ÚPPVII)

Aktivita	Dátum
Vypracovať podrobný návod pre aktualizáciu KRIS (po vytvorení metodiky bude podpora)	Marec 2017
Zriadenie Dátovej kancelárie verejnej správy	Apríl 2017
Pripraviť realizačný plán pre zavedenie princípu jedenkrát a dost' a jeho spustenie	Apríl 2017
Plán vyhlasovania referenčných údajov na rok 2017	Apríl 2017
Vypracovať zrozumiteľný a praktický návod a/alebo metodické pokyny pre inštitúcie verejnej správy v oblasti: <ul style="list-style-type: none">– Mapovanie dátových tokov,– Identifikáciu referenčných údajov,– Analýzu dátovej kvality a riešenia jej nedostatočnosti,– Meranie dátovej kvality,– Praktických krokov pre výkon referencovania.	Máj 2017
Vypracovať úvodný logický model údajov vo verejnej správe	Júl 2017
Vypracovať zrozumiteľný a praktický koncept pre službu Moje dáta a harmonogram krokov jej realizácie	Júl 2017
Vypracovať zrozumiteľný a praktický koncept pre službu analytické využitie údajov a harmonogram krokov realizácie	September 2017
Pripraviť program vzdelávania o správe a využívaní údajov pre inštitúcie verejnej správy	December 2017
Návrh legislatívnych zmien	Február 2018 Účinnosť zmien je žiadúca od Januára 2019
Gescia platformy integrácie údajov a jej úprava na dosiahnutie súladu s konceptami a architektúrou uvedenou v tomto dokumente	Júl 2018
Systematicky vyhodnotiť kvalitu údajov a publikovať výsledky	Každý rok (prvý krát Február 2018)
Pripraviť technologické riešenie pre správu údajov o subjekte (službu „moje dáta“)	September 2018
Pripraviť technologické riešenie pre analytickú platformu	September 2018



Aktivita	Dátum
Navrhnuť a štandardizovať centrálny dátový model postavený na ontológiách	Apríl 2018
Aktualizovať a schváliť štandard pre metódu tvorby a použitia referenčného URI identifikátora	Marec 2017
Pripraviť centrálnu riešenie pre správu prístupových oprávnení	August 2018

6.2.2 Aktivity a úlohy pre inštitúcie verejnej správy

Nasledujúce kroky

Tabuľka 16: Aktivity a úlohy pre inštitúcie verejnej správy

Aktivita	Dátum
Aktualizácia MetaIS v súlade s požiadavkami na Manažment údajov – vznikne tak plán na lepšie využívanie údajov pre danú inštitúciu	Júl 2017
Realizácia navrhnutých organizačných zmien (zabezpečenie dostatočných kapacít a vymenovanie rolí podľa kapitoly 3.3 organizačné zmeny)	August 2017
Zavedenie systematického manažmentu kvality údajov na úrovni inštitúcie	December 2017
Pripojenie sa na ako konzument do platformy integrácie údajov	September 2017
Integrácia s platformou integrácie údajov v roli poskytovateľa	November 2017
Vyhlásenie referenčných údajov, ktoré má inštitúcia v kompetencii	Január 2018
Stotožnenie vlastných údajov s referenčnými údajmi	Január 2019 (respektíve rok od vyhlásenia konkrétnych referenčných údajov)
Vypracovanie konceptu využívania analytických údajov na úsekoch verejnej správy danej inštitúcie	Január 2018
Nastaviť model prístupových práv k údajom v gescii inštitúcie	December 2018

Príprava koncepcie rozvoja informačných systémov

Do procesu vyhodnotenia súčasného stavu ISVS (v zmysle NKIVS je súčasťou aktualizácie KRIS) bude postupne zahrnuté aj získavanie informácií v súvislosti s manažmentom údajov. Bude potrebné upraviť MetaIS, aby takéto údaje evidoval.

Tabuľka 17: Rozšírenie MetaIS (a KRIS) o náležitosti v súvislosti s manažmentom údajov

Položka	Popis
Aké údaje OVM spracováva	<p>Cieľom je vytvoriť katalóg údajov za všetky OVM. Pre každú evidenciu by mal obsahovať položky:</p> <ul style="list-style-type: none">– názov evidencie/registra,– datasety v tejto evidencii,– príslušná legislatíva,– počet evidovaných údajov, pre každý dataset,– typizované výstupy z evidencie - napr. potvrdenia / formuláre / rozhodnutia. <p>Sústrediť sa treba na tie údaje, ktorých je tento OVM zdroj. Následne ako súčasť dlhodobej koncepcie rozvoja IS organizácia uvedie, ktoré údaje z jej evidencie je plán vyhlásiť za referenčné a kedy (orientačne) a ktoré dátové prvky majú byť vytvárané.</p>
Vyhodnotenie úrovne vyspelosti spracovania údajov	<p>Vyhodnotiť formát spracovania údajov podľa centrálne určených kritérií:</p> <ul style="list-style-type: none">– papierové spracovanie – elektronické textové (napríklad word doc) - elektronické štruktúrované (napríklad excel tabuľky) - databáza– bez štruktúry - ad hoc štruktúra <p>identifikácia údajov evidencie jednotným referencovateľným identifikátorom. Vyhodnotiť samostatne pre každý typ údajov z katalógu (viď. predošlý bod). Následne ako súčasť dlhodobej evidencie údajov v MetaIS organizácia uvedie, ktoré jednotlivé typy údajov budú automatizovane sprístupnené a kedy (odhad).</p>
Vyhodnotenie kvality spracúvaných údajov	<p>Vyhodnotiť kvalitu údajov podľa vybraných kritérií.</p>
Vstupy od iných OVM	<p>Aké údaje od iných subjektov sú požívané pri činnosti / konaniach / v evidencii tohto OVM. Popísať:</p> <ul style="list-style-type: none">– Aké údaje sú priamo používané,– alebo potvrdenia / rozhodnutia / výpisy / formuláre,– názov registra,– na aký účel a podľa akej legislatívy. <p>Budú sa popisovať iba tie údaje, ktoré sú systematicky spracúvané.</p>

Položka	Popis
Legislatíva upravujúca sprístupňovanie údajov	<p>Vyhodnotenie v súčasnosti platných legislatívnych pravidiel upravujúcich sprístupňovanie údajov iným OVM</p> <ul style="list-style-type: none"> – ide o podchytenie špecifických zákonných úprav, mimo režimu zákona o e-Governmente – najmä identifikácia legislatívnych obmedzení, aj potenciálnych <ul style="list-style-type: none"> – pre rozsah údajov – to je ktoré údaje nemôžu byť poskytované – proces prístupu k údajom - špecifické požiadavky na zriadenie prístupu k údajom – schvaľovanie prístupu k údajom, autorizácia údajov

6.2.3 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 18

Projekt	Možný zdroj financovania	Aktivity	Garant
Zlepšenie využívania dát - Vytvorenie dátovej kancelárie verejnej správy (Kompetenčné centrum referenčných údajov)	OPEVS	<ul style="list-style-type: none"> – Zavedenie manažmentu kvality údajov vo verejnej správe – Vybudovanie a správa centrálného modelu údajov – Úprava MetaIS na evidovanie manažmentu údajov – Príprava legislatívnych zmien (zákon o údajoch) – Príprava metodických pokynov pre jednotlivé oblasti manažmentu údajov – Program vzdelávania v oblasti využívania údajov – Centrálny manažment potrieb, prioritizácia a plánovanie (vo vzťahu jedenkrát a dosť vo verejnej správe) 	ÚPPVII



Projekt	Možný zdroj financovania	Aktivity	Garant
Projekty zavedenia princípu jedenkrát a dostať	OPII	<ul style="list-style-type: none"> – Pripájanie konzumentov do platformy integrácie údajov – Pripájanie nových poskytovateľov údajov – Participácia inštitúcii mimo verejnej správy (napríklad banky, utilitné spoločnosti) – Implementácia opatrení pre zvýšenie kvality údajov, najmä kvality dát referenčných registrov, zavádzanie referenčnej integrity 	*bude potrebné určiť spôsob, ako aktivity realizovať (predpokladá sa dopytová výzva)
Rozvoj platformy integrácie údajov (nové centrálné služby)	OPII	<ul style="list-style-type: none"> – Prispôsobenie platformy integrácie údajov riešeniu podľa tohto dokumentu – Služby pre virtualizáciu dát – Zápisová služba pre komunikáciu zdrojového a referenčného registra – Priamy „push model“ distribúcie údajov – Mechanizmus komunikácie o zmenách údajov – Služba vyhľadávania, – Zabezpečenie údajov pre službu „Moje dáta“ – Nástroje pre správu oprávnení – Vytvorenie GUI pre prístup k registrom a referenčným údajom – Vytvorenie platformy generického registra – Služby manipulácie a transformácie údajov 	ÚPPVII
Sprístupnenie prepojených otvorených dát na úrovni 5*	OPII	<ul style="list-style-type: none"> – Zabezpečenie publikovania otvorených údajov 	Inštitúcie VS
Služba „moje dáta“	OPII	<ul style="list-style-type: none"> – Technické riešenie pre správu osobných údajov – Riešenie pre „transparentné logovanie“ prístupov k údajom – Realizácia vizualizácie služby (portfólio klienta) 	ÚPPVII



Projekt	Možný zdroj financovania	Aktivity	Garant
Realizácia zdieľania právoplatných rozhodnutí	OPEVS	– Návrh procesných a legislatívnych zmien v súvislosti so zdieľaním právoplatných rozhodnutí v procesoch verejnej správy	ÚPPVII v spolupráci s MV SR
	OPII	– Analýza a návrh mechanizmu pre zdieľanie rozhodnutí (uvažuje sa o novom spoločnom module) – Implementácia mechanizmu a jeho zapojenie do procesov vo verejnej správe	ÚPPVII v spolupráci s NASES a MV SR
Elektronizácia registrov	OPII	– Ide o súbor projektových aktivít pre registre, ktoré nie sú doteraz elektronizované, očakáva sa implementácia registrov so štandardizovaným API	* Inštitúcie VS, ktoré vedú príslušné registre
Elektronický register Orgánov verejnej moci (OVM)	OPII	– Implementácia registra orgánov verejnej moci a jeho nasadenie v procesoch verejnej správy	MV SR
Elektronický register mimovládnych neziskových organizácií (MNO)	OPII	– Vyvinutý software pre Elektronický register MNO a jeho nainštalovanie, vrátane Otvoreného informačného systému MNO	MV SR
Konsolidovaná analytická vrstva	OPII	– Nasadenie analytických nástrojov a vytvorenie analytickej vrstvy podľa potrieb analytických jednotiek	ÚPPVII
Platforma pre posudzovanie vplyvov	OPII	– Vytvorenie systému evidujúceho platné regulácie, ktorý bude pravidelne aktualizovaný na základe vyhodnotenia efektivity platných regulácií. – Interaktívna analýza prínosov a nákladov, ktorá bude využívať analytické nástroje, umožňujúce simulácie budúcich dopadov. – Implementácia kolaboračných nástrojov.	MH SR

Projekt	Možný zdroj financovania	Aktivity	Garant
Zabezpečenie harmonizovaných referenčných priestorových údajov a súvisiacich služieb	OPII	<ul style="list-style-type: none"> Pre potreby verejnej správy, privátneho sektora a širokej verejnosti budú centrálné získané kvalitné, harmonizované priestorové údaje a súvisiace služby v zmysle požiadaviek INSPIRE, Otvorených údajov a e-Government Tieto priestorové zdroje bude možné využiť aj pre tvorbu mapových výstupov a podkladov 	ÚGKK SR v spolupráci s MZP SR
Reforma zberu štatistických údajov	OPEVS	<ul style="list-style-type: none"> Nové pravidlá pre zber a spôsob získavania údajov pre potreby štatistiky a analytickú podporu rozhodovania vo verejnej správe 	ŠÚ SR
Jednotný informačný systém pre získavanie a vyhodnocovanie štatistických údajov	OPII	<ul style="list-style-type: none"> Implementácia SaaS služby pre zber štatistických údajov vo verejnej správe 	ŠÚ SR
Sémantický model údajov a rozšírenie správy URI v MetaIS	OPII	<ul style="list-style-type: none"> Systematické zavedenie sémantického dátového modelu Podpora elektronických služieb pre automatizovanú registráciu URI identifikátorov využiteľné rôznymi ISVS (MetaIS ako referenčný register URI) Rozšírenie správy záznamov žiadostí o URI identifikátory Rozšírenie URL lokalizovateľnosti jednotlivých entít MetaIS Poskytovanie služieb súvisiacich s URI Publikácia záväzných URI Automatizovaná publikácia datasetov súvisiacich s URI z MetaIS Dereferenciácia URI 	ÚPPVII (ako správca MetaIS)
Sémantické vyhľadávanie nad normatívnymi a vzdelávacími dokumentami	OPII	<ul style="list-style-type: none"> pre IKT a kybernetickú bezpečnosť vo verejnej správe (v štandarde W3C SKOS pre Web 3.0 služby) 	ÚPPVII v spolupráci s NBÚ

6.2.4 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
Zákon č. 305/2013 Z. z. o elektronickej podobe výkonu pôsobnosti orgánov verejnej moci a o zmene a doplnení niektorých zákonov (zákon o e-Governmente)	Úprava zákona o e-Governmente: <ul style="list-style-type: none"> aby každý zdieľaný dataset prešiel dátovou kontrolou (presnejšie aby bolo pre údaje vidno ich kategorizáciu z hľadiska kvality), Znížiť súčasný počet častí MÚK súvisiacich so zdieľaním údajov, ideálne do jednej, Iniciatívu vyhlásenia referenčných údajov nenechať iba na správcovi príslušného registra, ale poveriť ňou aj ÚPVIAI. Vymenovať taxatívne a prísne definované dôvody odmietnutia správcu vymenovať údaj za referenčný. Mechanizmus zdieľania rozhodnutí 	ÚPPVII	Január 2018
GDPR - Všeobecné nariadenie EU (2016) 679 o ochrane údajov	Zohľadniť pri technických realizáciách (najmä služby Moje dáta)	ÚPPVII	Máj 2018
Nový zákon o údajoch vo verejnej správe	Nový zákon o údajoch vo verejnej správe: <ul style="list-style-type: none"> špecifikuje publikovanie otvorených údajov, modernizuje spôsob používania údajov vo verejnej správe – údaje bude možné používať na analytické účely a pre podporu rozhodovania ak sa zabezpečí ochrana osobných údajov, stanoví rámec pre anonymizáciu dát, nastaví práva a povinnosti vlastníkov dát. 	ÚPPVII	Február 2018 Účinnosť zmien je žiaduca od Januára 2019
Vykonávajúci predpis	Zriadenie Dátovej kancelárie verejnej správy	ÚPPVII	Marec 2017
	Definovanie nových rolí a zodpovedností v oblasti Riadenia dát	ÚPPVII	Apríl 2017
Aktualizovaná Jednotná metodika na posudzovanie vybraných vplyvov.	Transformácia procesov posudzovania vybraných vplyvov tak, aby podporovali využívanie analýz vytvorených v systéme.	MH SR	Január 2019

Legislatívny predpis	Legislatívna požiadavka	Garant	Termín ukončenia
Kategorizácia a klasifikácia údajov	Nová kategorizácia a klasifikácia údajov	ÚPPVII	Február 2018
Metodické usmernenie	Návrh transformácie rozhodovacích procesov vo verejnej správe (v rámci reformy verejnej správy)	MV SR	Január 2019
Predpis č. 540/2001 Z. z. Zákon o štátnej štatistike	Aktualizovať o nové metodiky v súlade s Európskym štatistickým systémom a Big Data, centralizovať procesy	ŠÚ SR	Január 2019
Výnos o štandardoch	Metodika tvorby konkrétneho URI identifikátora pre danú entitu (syntax a sémantika referenčného URI)	ÚPPVII	Júl 2017
	Metóda a proces centrálného pridelovania referenčných URI identifikátorov pre údaje ISVS pre všetky referenčné údaje	ÚPPVII	Júl 2017
	Návrh ontologického dátového modelu ako centrálného modelu údajov verejnej správy	ÚPPVII	November 2017
	Návrh štandardov pre analytickú vrstvu vo verejnej správe, internet vecí a Big Data	ÚPPVII	Apríl 2018

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

- Návrh a nastavenie nových procesov pre manažment dát vo verejnej správe pre analytické účely vrátane návrhu analytických produktov pre jednotlivé sektory.
- Podpora zriadenia a vybudovanie kapacít Dátovej kancelárie verejnej správy.
- Vybudovanie analytických jednotiek v relevantných inštitúciách.
- Reforma zberu a spracovania štatistických údajov vo verejnej správe.
- Optimalizácia rozhodovacích procesov vo verejnej správe.
- Vytvoriť katalóg dát (pojmov a definícií), dátových modelov existujúcich agendových informačných systémov.
- Vytvoriť mapu dátových tokov medzi jednotlivými agendovými informačnými systémami. Logické a funkčné prepojenie životných situácií s orchestračnou vrstvou a MUK pre proaktívne fungovanie verejnej správy.

7 Udržateľnosť a financovanie

Požiadavky na udržateľný prístup k projektom

Pre zmysluplné nastavenie projektov v rámci programu Manažment údajov je potrebné :

- V každom projekte na strane inštitúcie verejnej správy, ktorý predstavuje nový register alebo novú agendu, či ich rozšírenie, musí byť vyčlenená aktivita určená pre zapojenie sa do platformy integrácie údajov tak z pohľadu konzumenta údajov, ako aj poskytovateľa údajov, pokrytie vytvárania referenčných údajov z riešenej agendy/registra, zaistenie potrebnej kvality údajov, riešenie organizačného zabezpečenia riadenia údajov, realizácia konceptu „jedenkrát a dost“ pre riešenú agendu/register, publikovanie otvorených údajov. Neexistencia týchto aktivít musí byť explicitne zdôvodnená.
- V prípade realizácie projektu z OPII pre realizáciu „jedenkrát a dost“ a dost zabezpečiť mechanizmus financovania pre hlavného prijímateľa (ÚPPVII) a ostatných prijímateľov (ostatné inštitúcie verejnej správy realizujúce zdieľanie dát).
- Nastaviť podmienky pre inštitúcie verejnej správy pre financovanie zapojenia sa na MUK zo štátneho rozpočtu tak, aby mohli byť tieto zdroje refundovateľné z OPII, vrátane tých inštitúcií verejnej správy, ako sú napríklad stavovské organizácie (komory regulovaných profesií) v zdravotníctve, ktoré nie sú zatiaľ ohľadom financovania preneseného výkonu verejnej správy napojené na štátny rozpočet.

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.

Prevádzka týchto systémov, ako i udržanie potrebných odborných kapacít po skončení projektov si vyžiada krytie zo štátneho rozpočtu. Od roku 2021 (po skončení programu a zároveň programového obdobia) ide odhadom o 1 milión EUR na centrálnej úrovni ročne a kumulatívne 2 milióny EUR ročne na úrovni jednotlivých inštitúcií.

Nové fungovanie verejnej správy (vďaka realizácii týchto zmien) bude mať zásadné dopady na:

- Potenciál verejnej správy lepšie regulovať a vymáhať regulácie (predpoklady pre implementáciu konceptu Regulácie 2.0, zníženie regulačnej záťaže)
- Kvalitu služieb verejnej správy (zníženie administratívnej záťaže, jednoduchšie služby, predpoklady pre personalizáciu, povýšenie manažmentu osobných údajov)
- Efektivitu operácií vo verejnej správe.

Vznikne tiež vysoký potenciál na zníženie potreby zamestnávať pracovníkov na formálne, repetatívne a administratívne činnosti (až o 30 percent). Ušetrené náklady na ľudské zdroje by mali byť použité práve na udržanie vysoko kvalifikovaných pracovných miest, ktoré si systematický manažment údajov vyžaduje.

Dôležitým dopadom je tiež celková stimulácia digitálnej ekonomiky a dátového hospodárstva vďaka navrhovaným opatreniam. Očakávame významné zvýšenie kapacity digitálnej ekonomiky tvoriť inovatívne riešenia (postavené na využití údajov) a „know-how“ v dátovej vede.

Po realizácii navrhovaných opatrení bude verejná správa iná: modernejšia, flexibilnejšia, schopná efektívne fungovať v 21. storočí. Takéto systémové dopady odôvodňujú realizáciu navrhnutého zámeru, aj financovanie jeho výsledkov po skončení programu Manažment údajov v roku 2020.

Odporúčania z ohľadom na udržateľnosť

Aby riešenie dlhodobo fungovalo, odporúčame:

- Zabezpečiť dostatočné zdroje potrebné na manažment údajov (do 4 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).

8 Prílohy

Ako prílohy sú v dokumente zahrnuté:

- Skratky a definície.
- Súčasný stav pre Modul úradnej komunikácie – dátová časť ostatných úsekov verejnej správy.
- Návrh klasifikácie a kategorizácie údajov pre podmienky SR.
- Posudzovanie možností pri výbere strategického smerovania.

8.1 Skratky a definície

Definície

Odborný termín	Definícia
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.

Odborný termín	Definícia
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

Skratka	Popis
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

8.2 Kandidáti na referenčné údaje.

Doterajšia prax ukazuje, že vyhlásenie údajov za referenčný v zmysle zákona 305/2013 predstavuje zatiaľ najúčinnjší nástroj na donútenie orgánov verejnej moci pristúpiť k systémovému riešeniu zdieľania údajov elektronickou formou a naplňať princíp 1x a dost'. Dôležitou skutočnosťou v tomto procese je legislatívna istota pre orgán verejnej moci, ktorý v danom konaní používa referenčné údaje.

Potenciálne referenčné údaje uvedené v tejto prílohe neboli identifikované na základe dôslednej legislatívnej analýzy jednotlivých poskytovateľov alebo konzumentov údajov. Zoznam predstavuje súhrn poznatkov získaných z komunikácie s OVM počas implementácie projektu IS CSRÚ a súvisiacich aktivít.

Jednotliví kandidáti na referenčné údaje sú identifikovaní a definovaní na logickej, resp. biznis úrovni a nie na úrovni konkrétneho dátového prvku. Detailná špecifikácia a dátový typ referenčného údajov je možná len v komunikácii so správcom údajov.

Finančné vzťahy subjektu a štátneho rozpočtu

Potenciálne referenčné údaje pre oblasť finančných vzťahov subjektu a štátneho rozpočtu, respektíve subjektu a poisťovních fondov:

- Nedoplatok subjektu evidovaný na colnom úrade,
- Nedoplatok subjektu evidovaný v Justičnej pokladnici,



- Nedoplatok subjektu evidovaný v Centrálnom registri pohľadávok štátu,
- Vybrané údaje o platiteľovi zdravotného poistenia,
- Vybrané údaje o platiteľovi sociálneho poistenia,
- Plnenie odvodových povinností na zdravotnom poistení,
- Plnenie odvodových povinností na sociálnom poistení,
- Nedoplatok evidovaný zdravotnou poisťovňou,
- Nedoplatok evidovaný Sociálnou poisťovňou,
- Vybrané údaje o prijatej štátnej pomoci pre subjekt,
- Daňové identifikačné číslo.

Sociálny status subjektu

Potenciálne referenčné údaje pre oblasť posudzovania sociálneho statusu subjektu a pracovnoprávných vzťahov:

- Dávky a príspevky v hmotnej núdzi poskytované subjektu,
- Peňažné príspevky na kompenzáciu dôsledkov ZŤP,
- Dávky štátnej sociálnej podpory,
- Náhradné výživné,
- Materská dávka a sociálne dávky poskytované subjektu,
- Potvrdenia ZŤP,
- Údaje (kontroly) o nelegálnej práci a zamestnávaní,
- Údaje o práci v zahraničí.

Vzdelávanie

Potenciálne referenčné údaje v oblasti vzdelávania:

- Údaje o dosiahnutom stupni vzdelania a kvalifikácií,
- Údaje z dokladov o vzdelaní alebo štátnych skúškach,
- Údaje o zápise/návšteve základnej alebo strednej školy,
- Údaje o zápise a štúdiu študenta na vysokej škole,
- Údaje o plnení školskej dochádzky,
- Údaje o získanej kvalifikácii v procese ďalšieho vzdelávania,
- Register škôl a školských zariadení,
- Register akreditovaných vzdelávacích inštitúcií poskytujúcich získanie kvalifikácie v procese ďalšieho vzdelávania.

Právny status subjektu

Potenciálne referenčné údaje o právnom statuse subjektu:

- Údaje o zápise osoby v Registri partnerov verejnej správy,
- Údaje o právnej skutočnosti v Registri úpadcov,
- Údaje o právnej skutočnosti v Registri exekúcií,
- Údaje z rodného, sobášneho a úmrtného listu.

Údaje o nehnuteľnostiach a vlastníckych vzťahoch

Údaje, ktoré sú vedené v rámci katastra nehnuteľností:

- Parcela, stavba, budova, byt a ich atribúty



- Geografické hranice KÚ, okres, obec
- Vlastníci, nájomcovia, držiteľia, ich atribúty a vzťah k nehnuteľnosti
- Údaj o konaní a jeho stave vo väzbe na nehnuteľnosť

Geografické údaje o objektoch pevne spojených so zemou

- Množstvo typov objektov od železničnej trate až po vodné toky
- Ich geografické zobrazenie
- Záväzný názov objektu

Adresy a ich vzťahy

Údaje vedené v rámci Registra adries:

- Štát, VÚC, okres, obec, časť obce, ulica, orientačné číslo, popisné číslo
- Záväzný názov objektu
- Geografická poloha adresného bodu
- Geografický priebeh ulice

Ďalšie referenčné údaje

Ostatné potenciálne referenčné údaje a základné číselníky:

- Údaje o vydaných licenciách (napríklad pre taxi služby a podobne),
- Číselníky Úradu geodézie, kartografie a katastra SR,
- Číselníky ISCO (klasifikácií zamestnaní),
- Číselníky rozpočtovej klasifikácie.

URI Identifikátory

Kľúčovým atribútom implementácie princípov Linked Data je stanovenie takzvaného referenčného URI, ktorý môže vydávať výhradne Centrálny metainformačný systém. Keďže sa môže stať, napr. z historických alebo iných dôvodov, že nejaká entita má vo rôznych informačných systémoch vedené rôzne URI identifikátory, tak referenčným je ten, ktorý bol pridelený proces pridelovania URI v MetaIS. MetaIS sa tak stáva kľúčovým centrálnym prvkom informačnej architektúry verejných dát Slovenskej republiky, to je MetaIS je referenčným registrom pre URI:

- Jednotné referencovateľné identifikátory URI (MetaIS).

8.3 Klasifikácia a kategorizácia údajov

Metodológia a procesy pre riadenie budú nastavené tak, aby vychádzali z navrhutej klasifikácie:

- podľa charakteru údajov,
- podľa vlastníctva údajov,
- podľa pôvodu údajov,
- podľa zdroja,
- podľa typu a formy,
- podľa úrovne strojového spracovania,
- podľa typu interoperability zverejnených dát.

Charakter údajov

1. **Normatívne** (upravujú, regulujú správanie či činnosti subjektov - zvyčajne vo forme dokumentov, čiastočne vnútorne štruktúrovaných)



- a. Legislatívne (napríklad medzinárodné zmluvy, ústava, zákony, podzákonné normy),
 - b. Riadiace (napríklad stratégie, metodiky, návody, interné normy OVM)
 - c. Súdne (napríklad rozhodnutia, nálezy, judikáty),
 - d. Zmluvné (napríklad zmluvy, dohody práce, záručné podmienky),
 - e. Agendové (rozhodnutia v rámci agend verejnej správy určené osobe),
 - f. Štandardizačné (normy a štandardy, napríklad STN, ISO),
 - g. Terminologické (napríklad terminologický slovník krízového manažmentu),
 - h. Špecializované (napríklad pravidiel slovenského pravopisu).
2. **Kategorizačné** (umožňujú zaradenie do jednej z konečnej množiny kategórií)
- a. Číselníky,
 - b. Zoznamy,
 - c. Jednoznačné referencovateľné identifikátory,
 - d. Kritériá.
3. **Evidenčné** (údaje o subjektoch, objektoch, veciach, financiách či iných entitách vedené v IS verejnej správy)
- a. Registre (napríklad RFO, OR, národné zdravotné registre, register konečných užívateľov výhod)
 - b. Údaje v systémoch evidencií (napríklad v matrike, zozname kategorizovaných liekov)
 - c. Údaje v ďalších IS VS
4. **Agendové**
- a. Súvisiace s výkonom agend verejnej správy (napr. konania, žiadosti, potvrdenia, agendové doklady, stanoviská, vyjadrenia, námietky, schválenia, prekladané materiály na rokovanie vlády, a podobne)
 - b. Interné (agenda sa týka OVM alebo výkonu jeho činnosti)
 - c. Externé (agenda sa týka osoby alebo inej entity mimo verejnej správy).
5. **Informatívne** (info-data)
- a. Súvisiace s výkonom verejnej moci OVM (organizačná štruktúra a poriadok, úradné hodiny, miesto úradu, telefónne čísla ...)
 - b. O relevantných skutočnostiach a dianí v pôsobnosti OVM (napr. vestníky, varovania o nebezpečných výrobkoch, o chrípke, o novej legislatíve, ...)
6. **Dokumentárne** (dokumentujú, zaznamenávajú obraz reality a dianie v reálnom svete).
- a. Podľa oblasti:
 - i. Spravodajstvo (TV, rozhlas, spravodajstvo, youtube), životné prostredie (napríklad údaje SHMÚ)
 - ii. Bezpečnosť a ochrana (napr. monitorovanie verejných priestorov a perimetrov objektov OVM kamerami, web kamery)
 - iii. Monitoring cestnej dopravy
 - b. Podľa formy:
 - i. Video: napr. záznamy videokamier vo verejnom priestore, veľká časť videí na youtube, TV spravodajstvo RTVS, internetové video kanály
 - ii. Zvuk: zvukové nahrávky – napr. z jednaní súdov, rozhlasové spravodajstvo, záznamy hudobných vystúpení, záznamy mikrofónov web kamier
 - iii. Foto: napríklad ako súčasť dokumentácie k udalosti, objektu, javu.
 - iv. Textové: napríklad zápisnice, výpovede svedkov, dodaná dokumentácia k projektom, systémom a produktom vo VS, dokumentácia k aktu verejnej moci (napr. k stavebnému rozhodnutiu)

- v. Numerické: napríklad z technických zariadení (napr. vlhkosť, teplota, tlak, v rámci IoT)
- 7. **Prevádzkové** (súvisiace s prevádzkou danej inštitúcie, IKT, kontrolných a riadiacich systémov, elektroniky)
 - a. Údaje z HW zariadení (napr. senzorov)
 - b. Súbory OS a aplikácií:
 - c. Inštalačné súbory
 - d. Zdrojový kód
 - e. Konfiguračné údaje a súbory
 - f. Zálohy (súbory so záložnými údajmi)
 - g. Archivačné súbory
 - h. Identifikačné a autentizačné údaje
 - i. Logy
 - j. Ďalšie transakčné údaje
 - k. URL
 - l. Meta-dáta
- 8. **Odvodené údaje** (sú výsledkom spracovania údajov z ostatných domén)
 - a. Anonymizované, pseudonymizované údaje (prestanú byť osobnými údajmi v procese agregácie a oddelenia)
 - b. Štatistické (anonymizované údaje po spracovaní, vznikajú spracovaním najmä evidenčných údajov)
 - c. Analytické (sú výsledkom analýz nad primárnymi údajmi, štatistickými údajmi a big data)
 - d. Big data
 - e. Linked data (sémantické prepojenie údajov)
- 9. **Doménové:**
 - f. Finančné (údaje vo finančných inštitúciách súvisiace s financiami)
 - i. Bankové (účty a ich vlastníci, finančné transakcie nad účtami - komerčné banky, Štátna pokladnica)
 - ii. Súvisiace s finančnými produktami a službami (napr. hypotéky, úvery, pôžičky, poistenie, burza, ...)
 - g. Priestorové údaje
 - i. Geodetické údaje
 - ii. Priestorová priama, alebo nepriama referencia
 - iii. Údaje v geografických informačných systémoch
 - iv. Údaje ktorých súčasťou sú súradnice
 - v. Základné referenčné a tematické geodetické údaje o objektoch a javoch
 - h. Vzdelávacie (slúžia na vzdelávacie účely, sú produktom autora alebo vzdelávacej ustanovizne, týkajú sa najmä regionálneho školstva a VŠ)
 - i. Učebnice, učebné pomôcky a vzdelávacie materiály v elektronickej forme
 - ii. Digitalizované (napríklad scanované) pôvodne papierové učebnice
 - iii. ďalšie
 - i. Vedecko-technické (výsledok vedecko-technickej činnosti)
 - j. Zdravotné (health, healthcare)
 - i. O konkrétnej osobe (napr. v elektronickej zdravotnej knižke, v IS Univerzitnej nemocnice)

- ii. Údaje verejného zdravotníctva (napr. výsledky rozborov vzoriek vody, tkanív a biologických vzoriek pri podozrení na vtáčiu chrípku)
 - iii. ďalšie
 - k. Kultúrne (Culture, Entertainment)
 - i. Autorské diela v papierovej podobe (napr. knihy, obrazy, partitúry)
 - ii. Digitalizovaná forma autorského diela (napr. scanovaná publikácia, dokument obraz v rámci kultúrneho dedičstva)
 - iii. Autorské diela zachytené v elektronickej podobe (napr. balet, divadelné predstavenie, koncert)
 - iv. Autorské diela vytvárané v elektronickej podobe (napr. literárne dielo, videoklip, pesnička, elektronický obraz)
 - l. Biznisové
 - i. Súvisiace s činnosťou právnických osôb a fyzických osôb – podnikateľov mimo VS.
 - ii. Súvisiace s produktami a službami právnických osôb a fyzických osôb – podnikateľov mimo VS.
 - iii. ďalšie.
9. **Ostatné** (údaje nezaraďené do žiadnej z horeuvedených skupín)
- a. Nekategorizované
 - b. Iné

Vlastníctvo údajov

(dôležité z hľadiska ochrany údajov pri sprístupňovaní a zdieľaní, napríklad v duchu GDPR)

1. Údaje verejnej správy Slovenskej republiky.
2. Údaje verejnej správy inej členskej krajiny EÚ (napríklad DE, FR, CZ).
3. Údaje verejnej správy krajiny mimo EÚ, s ktorou má SR alebo EÚ dohodu o cezhraničnej výmene údajov (napríklad USA).
4. Údaje verejnej správy krajiny mimo EÚ, s ktorou nemá SR alebo EÚ dohodu o cezhraničnej výmene údajov (napríklad Nigéria).
5. Údaje medzinárodnej organizácie (napríklad OSN, Interpol), ktorej je SR členom.
6. Údaje medzinárodnej organizácie (napríklad OSN, Interpol), ktorej nie je SR členom.
7. Údaje fyzických a právnických osôb (mimo verejnej správy) Slovenskej republiky.
8. Údaje fyzických a právnických osôb (mimo verejnej správy) inej členskej krajiny EÚ.
9. Údaje fyzických a právnických osôb (mimo verejnej správy) krajiny mimo EÚ, s ktorou má SR alebo EÚ dohodu o cezhraničnej výmene a ochrane údajov (napr. USA).
10. Údaje fyzických a právnických osôb (mimo verejnej správy) krajiny mimo EÚ, s ktorou nemá SR alebo EÚ dohodu o cezhraničnej výmene a ochrane údajov (napríklad Papua Nová Guinea).
11. Údaje nezistených alebo neznámych vlastníkov.
12. Údaje bez vlastníckeho práva.

Pôvod údajov

- Primárny zdroj údajov (IS, do ktorého sú dáta prvotne zapísané – napr. matrika, zvyčajne IS je aj legislatívne vymedzený, napr. pre zákony e-Zbierka na slov-lex)
- Sekundárny zdroj (IS, ktorého dáta sú preberané z primárneho zdroja odpisom, výpisom, referenciou, ...)
- Terciárny zdroj (IS, ktorého dáta sú preberané zo sekundárneho zdroja)
- Referenčný zdroj (IS, z ktorého sú ďalšie IS VS oprávnené čerpať údaje)

Údaje úrovne strojového spracovania

- **Úroveň 0*:** Údaje nie sú poskytované v elektronickej forme.
- **Úroveň 1*:** Údaje sú dostupné vo webovom prostredí, avšak formát dát nemá štruktúrovaný obsah a nie je poskytovaný v otvorenom formáte. (napr. PDF)
- **Úroveň 2*:** Údaje sú dostupné vo webovom prostredí; obsah je štruktúrovaný tak, že umožňuje automatizované spracovanie, avšak nie je poskytovaný v otvorenom formáte. (napr. XLS)
- **Úroveň 3*:** Minimálna úroveň pre publikovanie podľa štandardov pre ISVS: údaje sú dostupné vo webovom prostredí; obsah je štruktúrovaný tak, že umožňuje automatizované spracovanie v otvorenom formáte, avšak nie s ohľadom na význam (napr. CSV, XML).
- **Úroveň 4*:** Na identifikáciu entít sa používajú Jednotné referencovateľné identifikátory v podobe URI. Jednotlivé údaje sú popísané prostredníctvom národných ontológií (data.gov.sk). Tým sa dáta vedia strojovo spracovať s ohľadom na význam (napr. RDF, OWL).
- **Úroveň 5*:** Národné entity sú prepojené na externé globálne štandardizované entity (napríklad odporúčené ontológie EÚ na popis dát).

Údaje podľa typu interoperability zverejnených dát

- **Katalógy** - sú kolekcie spravovaných údajov o datasetoch
- **Datasety** - sú kolekcie dát, publikovaných a spracovaných definovaných gestom, prístupné na prezeranie alebo na stiahnutie v jednom alebo viacerých formátoch
- **Distribúcia datasetu** - Reprezentuje špecifickú formu dostupného datasetu. Každý dataset môže byť dostupný v rôznych formách, pričom tieto formy môžu reprezentovať rôzne súborového formáty (XML, CSV, RDF, ...) rozličné prístupové aplikačné miesta (API, RSS) alebo verzie v čase.

8.4 Súčasný stav pre Modul úradnej komunikácie – dátová časť ostatných úsekov verejnej správy

Dátová časť ostatných úsekov Modulu úradnej komunikácie bola uvedená do produkcie v 11/2015 v zmysle zákonom predpísaných podmienok. Modul úradnej komunikácie predstavuje spoločný modul e-Governmentu a je prevádzkovaný na vládnom cloude v Datacentre Ministerstva financií SR. Funkčnosti tohto spoločného modulu boli vytvorené v projekte implementácie IS Centrálnej správy referenčných údajov.

V súčasnosti poskytuje modul nasledovné skupiny funkčností:

- Jednotný prístup k údajom – predstavuje vrstvu nad API poskytovateľov dát, ktorá poskytuje dáta pre konzumentov jednotným spôsobom. Z pohľadu konzumenta to znamená použitie jednej, resp. dvoch aplikačných služieb, ktorými vie získať ľubovoľné dáta dostupné v spoločnom module.
- Správa kmeňových dát a číselníkov – ak OVM potrebuje spravovať kmeňové/referenčné dáta alebo číselníky, ktorých charakter nevyžaduje vytvorenie osobitného agendového systému, môže na túto správu použiť spoločný modul.
- Dátová kvalita – nástroje pre čistenie údajov, data profiling a stotožňovanie subjektov evidencie. V prípade, že osobitný charakter dát nevyžaduje zo strany OVM nasadenie lokálne implementovaných nástrojov, môže použiť nástroje dostupné v cloude formou služby.

IS CSRÚ je zložený z nasledovných modulov:

- LDAP pre správu používateľov (dočasné riešenie kvôli chýbajúcej prístupovej časti MUK)
- CSRÚ Portál – vyvíjaný modul pre správu konfiguračných dát, ktoré riadia prístup k objektom evidencie; taktiež obsahuje GUI pre riadenie konkrétneho DQ procesu

- Talend MDM HUB – je dátové jadro systému, ktoré zabezpečuje správu dátových modelov, meta a riadiacich dát a samotných dát (vyššie v dokumente spomínaný „generický register“). MDM Hub slúži ako dátové úložisko pre zabezpečovanie ďalších funkcií systému - aktualizácia, konsolidácia, propagácia, stotožňovanie a dátová kvalita
- Talend DQ – zabezpečuje funkcie súvisiace s riadením dátovej kvality objektov evidencie pomocou dátového profilovania, identifikácie a stotožňovania, označovania záznamov a poskytovania výstupných súborov „vyčistených“ dát
- Talend MDM GUI – je používateľské WEB rozhranie pre používateľov prístupujúcich na MDM Hub v roli Data Stewarda podľa pridelených prístupových práv
- Talend DQ Portál – slúži na vykonávanie aktivít súvisiacich s riadením kvality dát a prezentáciu výsledkov formou grafov a reportov – napr. sledovanie trendu dátovej kvality apod.
- Talend Data Integration / ESB – integračné jadro (platforma) systému, zabezpečuje realizáciu integračných scenárov
- Talend Administration Console – slúži na centrálny monitoring a správu konfigurácie systému, nastavenie behu jobov a sledovanie ich priebehu a stavu
- Talend Studio – slúži na nastavenie dátovej, aplikačnej a integračnej logiky, nastavenia a konfigurácie dátového modelu, integračných scenárov a jobov.

IS CSRÚ pracuje na logickej úrovni s objektami evidencie. Objekt evidencie je dátová štruktúra popísaná príslušnou xsd schémou. Objekty evidencie sú definované pre poskytovateľov aj pre konzumentov údajov. Systém pripravuje objekt evidencie konzumenta podľa jeho požiadaviek. Objekt evidencie konzumenta môže mať dátovú štruktúru totožnú s objektom evidencie poskytovateľa alebo môže byť vytváraný z viacerých objektov evidencie poskytovateľov.

Objekty evidencie sú pre konzumentov dostupné unifikovaným spôsobom (web služba), pričom konzument v parametroch služby špecifikuje konkrétne ID objektu evidencie, ku ktorému má v riadiacich údajoch povolený prístup. Služba má synchrónny aj asynchrónny variant.

Spolu s uvedenou funkčnosťou boli v rámci projektu vypracované metodické postupy pre Riadenie kvality údajov, pre Použitie služieb a pre Vyhlasovanie referenčných údajov.

Zoznam referenčných údajov a základných číselníkov:

Metodika pre vyhlasovanie referenčných údajov bola v praxi overená na doteraz vyhlásených referenčných údajoch a základných číselníkoch, ktoré sú:

1. IČ DPH, register: Register daňových subjektov, správca: FR SR,
2. Dátum zaradenia uchádzača o zamestnanie do evidencie uchádzačov o zamestnanie, register: Register uchádzačov o zamestnanie, správca: UPSVaR,
3. Dátum vyradenia uchádzača o zamestnanie z evidencie uchádzačov o zamestnanie, register: Register uchádzačov o zamestnanie, správca: UPSVaR,
4. Cca 38 položiek o právnickej osobe, register: RPO, správca: Štatistický úrad SR,
5. Položky o fyzickej osobe, register: RFO, správca: MV SR,
6. 27 základných číselníkov, register: MetaIS, správca ÚPPVII,
7. Základný číselník úsekov verejnej správy a agend verejnej správy, register: MetaIS, správca ÚPPVII,
8. Číselník životných situácií a okruhov životných situácií, register: MetaIS, správca ÚPPVII.

Konzumenti referenčných údajov

Jednotný prístup k údajom a funkčnosti dátovej kvality boli overené na pilotnej vzorke orgánov verejnej moci, zapojených v rámci implementačného projektu.

Orgány verejnej moci používajúce spoločný Modul úradnej komunikácie, dátová časť ostatných úsekov:

1. Finančná správa SR,
2. Štatistický úrad SR,



3. Ústredie práce sociálnych vecí a rodiny,
4. Úrad pre dohľad nad zdravotnou starostlivosťou,
5. Úrad podpredsedu vlády pre investície a informatiku,
6. Sociálna poisťovňa,
7. Zdravotné poisťovne (okrem Union ZP),
8. Datacentrum obcí a miest (aktuálne v testovaní),
9. Ministerstvo financií SR,
10. Datacentrum MF,
11. Mesto Bratislava,
12. Mesto Košice.

8.5 Posudzovanie variantov a výber strategického prístupu

V nasledujúcej kapitole sú zhrnuté varianty, ktoré boli zvážené počas tvorbu dokumentu Stratetickej priority Manažment údajov.

8.5.1 Referenčné údaje a „jedenkrát a dost“

Vymedzenie referenčných údajov

Variant	Výber	Zdôvodnenie a použitie
Referenčné údaje sú všetky údaje, ktoré vyjadrujú jedinečný fakt (kmeňové (evidenčné) a agendové údaje)	X	Mechanizmus pre zdieľanie údajov bude využitý univerzálnejšie.
Referenčné údaje sú iba kmeňové (evidenčné) údaje		

„Jedenkrát a dost“

Variant	Výber	Zdôvodnenie a použitie
Využitie metód zdieľania údajov (preferovane referenčných)	X	Preferované, keďže je k dispozícii technologické riešenie a potrebná legislatíva.
Zasielanie údajov na základe splnomocnenia	X	V obmedzených prípadoch (ak údaje nie sú k dispozícii ako referenčné údaje)
Ad-hoc výmena údajov	X	Do vzniku registra rozhodnutí.
Výmena údajov cez synchronizáciu procesov		
Spoločná databáza		

Identifikácia fyzických osôb

Variant	Výber	Zdôvodnenie a použitie
Zmiešané použitie RČ a BIFO (Pričom tieto identifikátory môžu byť aj vo formáte URI) - V jednom sete objektov môžu takto súčasne byť prítomné záznamy s RČ aj BIFO (máme na mysli stav viditeľný na rozhraní platformy zdieľaných údajov, nie internú realizáciu).	X	
Použitie sektorových identifikátorov FO	X	Kompromisné riešenie: systém segmentov sa určí základe zoznamu segmentov, ktorý sa pripraví na základe (cost-benefit analýzy a analýzy rizík)
Zrušiť plán sektorových identifikátorov FO		V praxi nie sú nasadené. Nie je jasný dôvod, prečo majú byť používané, náklady na prechod na ich používanie, ani prevádzkové náklady v rámci celej VS - najmä dopad na zníženie efektívnosti práce údajmi. V súčasnosti považujeme koncept striktne oddelených sektorov za prekonaný. Napríklad v rámci použitia cloudu by jeho nasadenie znamenalo že potenciálne na jednom fyzickom serveri bežiace rôzne inštancie OS/DB by museli spolu komunikovať cez IS IFO, čo je extrémne zvýšenie nákladov. Rovnako opačné uvažovanie je vidieť v koncepte spoločných komponentov, alebo v NKIVS uvažovaný integrovaný ISVS.

Akú formu budú mať rozhodnutia

Variant	Výber	Zdôvodnenie a použitie
Rozhodnutia sú strojovo spracovateľné	X	60 percent údajov z rozhodnutí bude možné strojovo spracovávať.

Variant	Výber	Zdôvodnenie a použitie
o rozhodnutiach sú strojovo spracovateľné	X	100 percent meta-dát o rozhodnutiach bude strojovo spracovateľných.

Ako zdieľať rozhodnutia (v rámci konania verejnej správy)

Variant	Výber	Zdôvodnenie a použitie
Prostredníctvom zdieľania elektronického spisu	X	
Prostredníctvom registra právoplatných rozhodnutí – celé rozhodnutia		
Prostredníctvom registra právoplatných rozhodnutí – meta-údaje	X	Evidencia meta-údajov na jednom mieste môže výrazne pomôcť pri orientácii sa v rozhodnutiach iných inštitúcií (po vzore registra zmlúv).
Na základe ad-hoc výmeny údajov – distribuovaný prístup		

8.5.2 Kvalita údajov

Prístup k riadeniu kvality údajov

Variant	Výber	Zdôvodnenie
Centralizované		
Decentralizované		
Hierarchické - centrálna metodika, parametre, procesy a monitoring, decentralizovaná zodpovednosť za kvalitu údajov	X	Takéto riešenie umožní zároveň aplikovať aj centrálnu úlohu a zodpovednosť pri manažmente údajov a zároveň princíp, že za údaje zodpovedá správca príslušného informačného systému.

Nástroje pre riadenie kvality údajov

Variant	Výber	Zdôvodnenie
Decentralizované riešenie kvality údajov		

Variant	Výber	Zdôvodnenie
Použitie lokálnych nástrojov pre kvalitu údajov na základe centrálnej metodiky	Vo vybraných prípadoch	Pre dôležité údaje.
Použitie centrálnych nástrojov a služieb pre zvyšovanie kvality údajov (formou štandardného katalógu služieb pre kvalitu údajov)	Prioritne	Použitie spoločných metód a nástrojov je ekonomické v kontexte vládneho cloudu.

Monitoring kvality údajov

Variant	Výber	Zdôvodnenie
Centrálny zber KPI kvality údajov	X	Potrebné, aby bolo zabezpečené porovnateľné spracovanie výsledkov a tiež dostatočná pozornosť pre aktivitu.
Zverejňovanie kvality údajov pre jednotlivé inštitúcie	X	Potrebné, aby bol vytvorený motivačný tlak pre jednotlivé inštitúcie (vďaka transparentnému zverejňovaniu výsledkov).

8.5.3 Integrácia údajov

Koncept integrácie údajov

Variant	Výber	Zdôvodnenie a použitie
Využitie jednotnej platformy integrácie údajov- zlúčiť všetky dátové časti modulu úradnej komunikácie od jednej, ktorá bude preferované v kompetencii ÚPVII - prostredníctvom tohto modulu bude realizované sprístupňovanie všetkých údajov	X	Cieľom je vytvorenie spoločného mechanizmu integrácie údajov.
Rozdelenie na 2 platformy integrácie údajov (v každom DC vládneho cloudu)		
Distribuovaný model s podpornými centrálnymi riešeniami (ak niektorý poskytovateľ údajov má problém (kapacitný, technický) implementovať štandardizované API : mať možnosť prijímať od neho údaje pomocou maximálne flexibilných rozhraní, ktoré údaje normalizujú a ďalej voči konzumentom sa používa štandardizované rozhranie)		

Logika integrácie údajov

Variant	Výber	Zdôvodnenie a použitie
Jednotný prístup k údajom - ontológia	X	Predpoklad pre efektívne a interoperabilné zdieľanie údajov naprieč verejnou správou (a „data-driven state“).
Rešpektovanie biznis logiky poskytovateľa – lokálny metamodel		

Spôsob riešenia komunikácie pri integrácii údajov

Variant	Výber	Zdôvodnenie a použitie
Dávkové spracovanie – asynchrónna výmena veľkého množstva údajov, platných k určitému času	X	Len v nevyhnutných prípadoch (pre veľmi veľké objemy údajov z dôvodu časovej efektívnosti)
Asynchrónna komunikácia	X	Preferované z pohľadu rýchlej optimalizácie procesov a ďalšieho nastavenia riešenia
Real-time (Synchrónna komunikácia)	X	Ideálne z pohľadu konzumenta – dá sa očakávať, že bude využitá po nasadení interaktívneho riešenia životných situácií
Virtualizácia dát	X	Pre komplexnú integráciu údajov z viacerých zdrojov a typov

Koncept správy kmeňových dát - MDM

Variant	Výber	Zdôvodnenie
Federácia údajov – jednotný virtuálny pohľad na dáta dostupný pre konzumentov údajov, ktorý si zabezpečia stotožnenie	X	Model vhodný pre heterogénne prostredie, akým je verejná správa. Výhodou je možnosť jednotného virtuálneho pohľadu.
Konsolidácia údajov – kmeňové dát z rôznych systémov sú konsolidované v jednom úložisku (operational data store) – odkiaľ sú následne replikované do jednotlivých databáz		



Variant	Výber	Zdôvodnenie
Propagácia údajov – výmena a kopírovanie kmeňových údajov medzi systémami - prijímatelia si zabezpečujú stotožnenie		

Ukladanie údajov (v jednotnej platforme integrácie údajov)

Variant	Výber	Zdôvodnenie
Centrálne		
Dočasné	X	Obmedzené pre konkrétne prípady použitia
Distribúované	X	Údaje budú ukladané v zdrojových systémoch

Technológia

Variant	Výber	Zdôvodnenie
ETL		
ESB	X	
Centrálne stanovené uniformné manažované API		
Linked data	X	
JSON Centric IpaaS		

Zaradenie integrácie údajov – komunikácia z registrom

Variant	Výber	Zdôvodnenie
Oddelenie funkcií na zdieľanie údajov od zápisových funkcií - funkcie zdieľania údajov riešiť oddelene od ostatných integrácií.	X	Funkcie na sprístupňovanie údajov majú podstatne odlišný charakter od zápisových funkcií, najmä z pohľadu riadenia prístupu, rozsahu subjektov ktoré majú služby využívať, špecifickosti pre doménu, závislosti na procesoch. Preto je správne vytvoriť pre zdieľanie údajov špecifické pravidlá (vrátane aplikačnej podpory), oddelene od zápisových funkcií.
Funkcie zdieľania údajov a funkcie zápisu sú riadené spoločnými pravidlami		

8.5.4 Moje dáta a ochrana osobných údajov

Spôsob skladania údajov pre moje dáta

Variant	Výber	Zdôvodnenie
Centrálné ukladanie		
Dočasné vyskladanie	X	Zdrojom údajov sú iné informačné systémy, kde dochádza k ich dynamickej zmene.

8.5.5 Analytické spracovanie údajov vo verejnej správe

Prístup k realizácii analytickej vrstvy

Variant	Výber	Zdôvodnenie
Distribúované riešenia		
Klasický DWH		
Systém virtuálnych DB	X	Nie je potrebné duplikovať údaje. Vhodné ako „štartovacie riešenie“ pre analytické jednotky.
„Data lake“		
„Data hub“		

Prístup k realizácii posudzovania vybraných vplyvov

Variant	Výber	Zdôvodnenie
Distribúované riešenia		
Spoločná platforma	X	Obdobnosť procesov pri posudzovaní vplyvov a potreba zdieľať skúseností pri návrhu dátových modelov a analýz uprednostňuje myšlienku spoločnej platformy, ktorá umožní posudzovanie vplyvov pre všetky relevantné inštitúcie.



Prístup k spracovaniu osobných údajov v analýzach

Variant	Výber	Zdôvodnenie
Osobné údaje nie sú spracovávané		
Pseudonymizácia	X	Osobné údaje a údaje, ktoré súvisia s daným subjektom sú predstavujú veľmi dôležitý vstup do analytických modelov. Aby mali predpovede zmysel a reálne použitie, je potrebné aby údaje mali čo najvyššiu informačnú hodnotu. Odporúčame preto pseudonymizovať osobné údaje, ide o vhodný kompromis medzi ochranou osobných údajov a ich využitím pre potreby lepšieho rozhodovania.