

Strategická priorita *Riadenie údajov a „Big data“*

Pracovná verzia
(Verzia 1.2)

Informácia o dokumente

Názov:	Strategická priorita Riadenie údajov a „Big data“
Stav:	Pracivná verzia
Pripravil:	Dagmar Bošanská, Juraj Bárdy, Oleg Lych, Valent Gura , Jiří Mizera,
Verzia:	1.2
Dátum:	1.12.2015
Pripomienkoval:	AKVS
Dátum revízie:	29.1.2016

Distribučný zoznam

Od	Dátum	Kontakt

Pre	Akcia*	Dátum (do)	Kontakt

* Akcia: *Schváliť, Pripomienkovať, Informovať, Realizovať, iné (uvedzte)*

História verzií

Verzia	Dátum verzie	Pripravil/ Zmenil	Pripomienkoval	Kľúčové zmeny
0.1	15.6.2015			Úvodné predstavenie strategickej priority
1.0	11.11.2015			Dopracovanie kapitol a príprava dokumentu na interné pripomienkovanie
1.2	1.12.2015			

Obsah

1	Úvodný pohľad	4
1.1	Skratky a definície	4
1.2	Definícia strategickej priority	6
2	Analýza dát	11
2.1	Ciele realizácie	12
2.1.1	Architektonické ciele	12
2.1.2	Stakeholderi a ich záujmy	14
2.2	Organizácia	15
2.2.1	Zodpovednosť	15
2.2.2	Organizačné zmeny	16
2.3	Stratégia a riešenie	16
2.3.1	Popis súčasného stavu	16
2.3.2	Najlepšie skúsenosti	17
2.3.3	Strategický prístup k riešeniu	17
2.3.4	Biznis vrstva	19
2.3.5	Aplikačná vrstva	21
2.3.6	Technologická vrstva	25
2.3.7	Predpoklady realizácie	25
2.4	Problémy a riziká	26
2.5	Legislatívne požiadavky	26
2.6	Plánovanie a migrácia	27
2.6.1	Centrálny dátový sklad a analytické nástroje	28
2.6.2	Inteligentné regulácie	28
2.6.3	Manažment kvality a výkonu verejnej správy	29
2.6.4	Nové nástroje pre Národný štatistický systém	30
2.6.5	IT nástroje pre test MSP	30
2.6.6	Analýzy sociálnych politík	31
3	Riadenie dát	32
3.1	Ciele realizácie	32
3.1.1	Architektonické ciele	32
3.1.2	Stakeholderi a ich záujmy	33
3.2	Organizácia	34
3.2.1	Zodpovednosť	34
3.2.2	Organizačné zmeny	35
3.3	Stratégia a riešenie	35
3.3.1	Popis súčasného stavu	35
3.3.2	Najlepšie skúsenosti	36
3.3.3	Strategický prístup k riešeniu	36
3.3.4	Biznis vrstva	38
3.3.5	Aplikačná vrstva	43
3.3.6	Technologická vrstva	45
3.4	Problémy a riziká	45
3.5	Legislatívne požiadavky	45

3.6	Plánovanie a migrácia	46
3.6.1	Vybudovanie platformy dátovej integrácie – Rozvoj MUK 2	46
3.6.2	Integrácie všetkých informačných systémov verejnej správy do platformy dátovej integrácie	47

1 Úvodný pohľad

1.1 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. Metaúdaje sú určené najmä na vyhľadávanie, katalogizáciu a využívanie primárnych údajov.
Mikrodáta	V kontexte štatistiky predstavujú mikrodáta informáciu na úrovni individuálnych respondentov, ktorí sa zúčastňujú prieskumov alebo sčítaní ľudu.
Paradáta	Paradáta sú dáta dokumentujúce proces zberu údajov. V prípade vykonania prieskumu môže ísť o informácie zahŕňajúce časy, kedy sa vykonávali pohovory, dĺžku trvania pohovorov, spôsob komunikácie, ochotu respondenta a mnohé ďalšie.
Prístup cez aplikačné rozhranie - API	Ide o aktívny prístup, kde používateľ zadáva serveru dotazy na konkrétne požadované údaje, ktoré server po spracovaní dotazu vyhledá a odošle

Odborný termín	Definícia
Referencovateľný identifikátor	<p>Ide identifikátor dátového zdroja, ktorý referencuje samotné údaje, t.j. po zadaní identifikátora do prehliadača je používateľ presmerovaný na URL vyplnené v metaúdajoch dátového zdroja, napr.:</p> <p>http://data.gov.sk/data/uvo/zmluvy/zmluvy</p> <p>Referencovateľným identifikátorom je teda identifikátor, ktorý:</p> <ol style="list-style-type: none"> 1. má formát Uniformed Resource Identifier (URI), 2. je jednoznačný, 3. je unikátny, 4. je dlhodobu stabilný, 5. je formátovo a štruktúrne konzistentný, 6. je manažovateľný tak, aby umožňoval logicky rozširovať stanovenú štruktúru, 7. je jasný, stručný a krátky, 8. je pre fyzickú osobu jednoducho čitateľný, 9. je bez súborových prípon, 10. neobsahuje programátorské kľúčové slová, 11. neobsahuje reťazec „www“, 12. neobsahuje interpunkciu okrem znakov lomka, pomlčka a bodka, diakritiku a medzery, okrem identifikátora fyzickej osoby podľa osobitného predpisu, kde je možné použiť interpunkciu a diakritiku, 13. obsahuje iba malé písmená, 14. nahrádza špeciálne znaky, napríklad výkričník, úvodzovky, percento, hviezdička, zátvorka, dolár alebo mriežka, pomlčkami a podčiarkovníkmi,
Surové dáta	Surové dáta, tiež nazývané primárne dáta, sú dáta získané priamo z dátového zdroja, ktoré 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)
FIA	Hodnotenie vplyvov na rozpočet (Financial impact assessment)
IFP	Inštitút finančnej politiky

Skratka	Popis
IKT	Informačno-komunikačné technológie
IVP	Inštitút vzdelávacej politiky
MSP	Malí a strední podnikatelia
MUK	Modul úradnej komunikácie
OP EVS	Operačný program Efektívna verejná správa
OP II	Operačný program Integrovaná infraštruktúra
RIA	Hodnotenie vplyvov regulácie (Regulatory impact assessment)
SBA	Slovak business agency
SIA	Hodnotenie sociálnych vplyvov (Social impact assessment)

1.2 Definícia strategickej priority

Strategická priorita

Národná koncepcia informatizácie verejnej správy (2016) 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
- 7 Otvorené údaje
- 8 Vládny cloud
- 9 Komunikačná infraštruktúra
- 10 Informačná a kybernetická bezpečnosť

NKIVS ku každej strategickej priorite informatizácie verejnej správy vysvetľuje jej cieľ, prístup k riešeniu a tiež rámcový architektonický model.

Tento dokument predstavuje prvý návrh a high-level analýzu priority informatizácie verejnej správy **Riadenie údajov a Big data upravenú v kapitole 6.2.6 NKIVS**, pripravenú v rámci Architektonickej kancelárie verejnej správy. Dokument ešte môže byť z pozície architektonickej kancelárie verejnej správy upravovaný a dopĺňaný aj na základe pripomienok a komunikácie s gestormi tejto strategickej priority.

Zodpovednosť za ďalšie detailné riešenie konkrétnej priority, vypracovanie štúdie jej realizovateľnosti a následnú realizáciu formou zabezpečenia implementácie príslušného projektu, resp. projektov, má gestor podľa jemu prislúchajúcej kompetencie.

Strategická priorita Riadenie údajov a „Big data“ ponúka návod na lepšie využívanie dát vo verejnej správe a podporuje prístup k dátam ako k vzácnemu zdroju:

- Podporí sa celý životný cyklus dát s dôrazom na riadenie dátovej kvality a prístupov k dátam na základe rolí a oprávnení.

- Zavedú sa pravidlá a nástroje, pomocou ktorých budú informačné systémy pristupovať k dátovým zdrojom iného informačného systému.
- Dáta budú systematicky využívané na zlepšenie rozhodovania vo verejnej správe vďaka ich analytickému spracovaniu.
- Posilnia sa práva občanov a podnikateľov v súvislosti s používaním dát, ktoré sa ich týkajú.

Vysvetlenie strategickej 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 architektúra 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.

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 získať 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 a ontológii 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 prioritě 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 služby dátovej integrácie Modulu úradnej komunikácie (MUK). Údaje analytického charakteru sa budú replikovať z transakčných a referenčných údajov do Centrálného dátového skladu. V informačnom prostredí verejnej správy je potrebné uplatniť princípy pre pohodlný a kontextovo-orientovaný prístup k dátam uloženým v iných systémoch, ako i definovať pravidlá, na základe ktorých informačné systémy pristupujú 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 podnikateľov. Informačné prostredie verejnej správy musí občanom a podnikateľom zabezpečiť transparentný pohľad na využívanie dát o nich. Na jednom mieste bude pre občana a podnikateľa dostupný prehľad o všetkých dátových objektoch, ktoré verejná správa vo svojich informačných systémoch eviduje. Zároveň bude môcť občan a podnikateľ vidieť, kedy a kto údaje menil alebo k nim pristupoval. 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 podnikateľ získa tak ešte väčšiu kontrolu nad svojim súkromím ako aj pohodlné nástroje na opravu alebo vymazanie údajov a udelenie súhlasu so spracovaním. Pre občanov a podnikateľov sa v budúcnosti stanú ich osobné údaje ekonomickým aktívom.

Súvisiacou strategickou prioritou sú Otvorené dáta, v ktorej sa navrhne, ako sa všetky dáta, ktoré má verejná správa k dispozícii, publikujú verejnosti vo formátoch vhodných na opakované použitie.

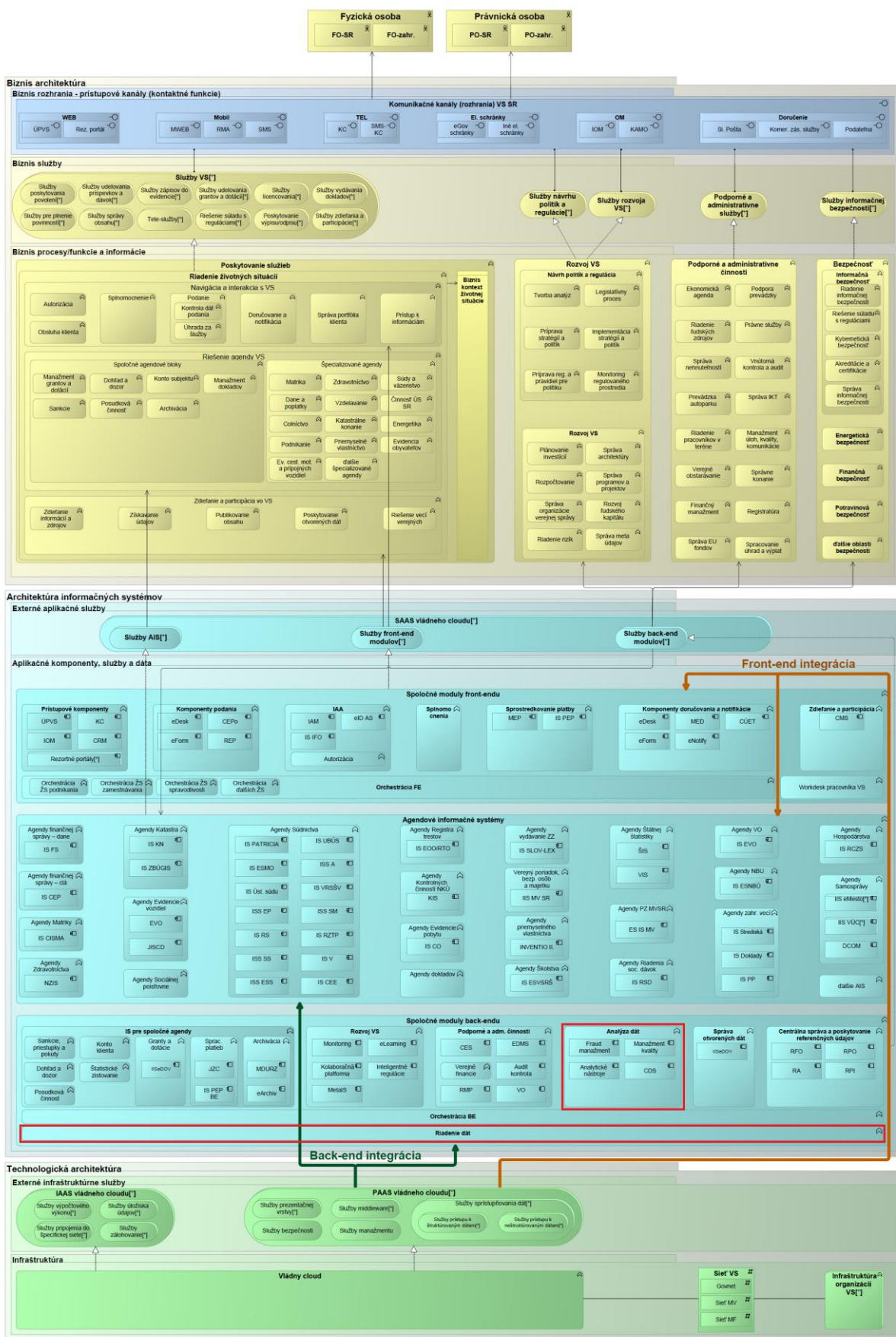
Vzťah k modelu architektonickej vízie

Strategická priorita Riadenie dát a „Big data“ je kľúčovým prvkom Architektonickej vízie verejnej správy 2020. Je úzko spätá so strategickými prioritami Integrácia a orchesterácia, Vlády cloud, Otvorené údaje a Bezpečnosť. Na úrovni aplikačnej vrstvy strategická priorita priamo súvisí:

- S aplikačnou funkciou Analýza dát (v ktorej hrajú význam komponenty CDS (Centrálny dátový sklad), Analytické nástroje, IS pre manažment kvality),
- S Centrálnou správou a poskytovaním referenčných údajov, v rámci ktorej sa implementujú nástroje dátovej integrácie a riadenia prístupu k referenčným údajom prostredníctvom Modulu úradnej komunikácie (MUK).
- S aplikačnou funkciou Riadenie dát.

Na úrovni biznis vrstvy strategická priorita súvisí s:

- S biznis funkciou **Tvorba analýz**, ktorá je súčasťou základnej funkcie Návrh politik a regulácií v rámci Rozvoja verejnej správy. Funkciu Tvorba analýz využívajú funkcie ako Príprava stratégií a politik; Príprava regulácií, Legislatívny proces, Monitoring regulovaného prostredia a Manažment kvality, úloh a komunikácie verejnej správy.



Strategická priorita Riadenie údajov a „Big data“ 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 dát a poskytovať svoje dáta pre analýzu dát. Znamená to, že základné komponenty pre riadenie dát a analýzu dát 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.

Stavebné bloky

Za základné stavebné bloky, na základe ktorých je štruktúrovaný tento dokument, sú považované:

- Riadenie dát,
- Analýza dát.

2 Analýza dát

„Data-driven state“

V rámci funkcie Analýza dát budú nasadené informačné nástroje, ktoré umožnia implementáciu konceptu „Data-driven state“ (teda štátu, ktorý funguje na základe využívania dát) do praxe.

Znamená to, 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.

Oblasti pre systematickú analýzu dát

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

- Strategické priority – príprava stratégií, analýza súčasného stavu a porovnávanie alternatív a následné sledovanie napĺňania 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.
- Investície – vyhodnocovanie dopadov projektov a verejných investícií,
- Výkonnosť a kvalita – vyhodnocovanie výkonu verejnej správy pri poskytovaní služieb v jednotlivých sektoroch,
- Operatívne vylepšovanie chodu verejnej správy a zavádzanie inovácií.

Typy analýz a analytické produkty

V jednotlivých oblastiach budú vykonávané rôzne typy analýz:

- Ex-ante vyhodnocovanie opatrení a alternatív,
- Ex-post vyhodnocovanie prijatých opatrení,
- Online vyhodnocovanie,
- Prognózy,
- Simulácie,
- Štatistické zisťovania.

Je dôležité, aby sa všetky typy analýz vykonávali nad spoločnou množinou dát, ktorá má vysokú kvalitu a ideálne aj jednotnú ontológiu a dátový model. Bude tak možné navrhovať a testovať predpoklady jednotlivých modelov a postupne ich zlepšovať.

Pre jednotlivé sektory verejnej správy a oblasti analýz bude možné definovať množinu analytických produktov. Tieto produkty budú pravidelne vytvárané a publikované pre potreby vlády SR, ministerstiev ale i verejnosti. Súhrny prehľad všetkých produktov by mal byť dostupný na jednom webovom sídle.

Analytické jednotky vo verejnej správe

Analytické produkty je možné chápať ako dopyt verejnej správy a verejnosti po činnosti analytických jednotiek. Analytické jednotky sú špeciálne organizačné útvary v rámci verejnej správy (spravidla súčasť jednotlivých

ministerstiev vlády SR), ktoré sa venujú tvorbe analýz. Zvýšenie kapacít analytických jednotiek a ich ďalšie zriaďovanie patrí k hlavným opatreniam Operačného programu Efektívna verejná správa. Nástroje informačných technológií, potrebné pre svoju prácu a produkovanie analytických produktov sú popísané v tomto dokumente.

Podpora informačnými technológiami

Zámerom tejto strategickej priority je vybudovanie Centrálného dátového skladu (CDS) na základe najmodernejších technológií pre ukladanie, spravovanie a analyzovanie dát. V tomto sklade budú ukladané všetky relevantné dáta pre sekundárne spracovanie nielen z agendových systémov verejnej správy, ale aj z iných dátových zdrojov vrátane dátových zdrojov Big Data. Pri importe dát z agendových systémov cez Modul úradnej komunikácie budú dáta konsolidované a normalizované, aby spĺňali najvyššie kritéria kvality. Zároveň bude vznikať zjednotený dátový model verejnej správy a ďalej sa budú štandardizovať formáty, metadáta, paradáta a ontológie interoperabilné s krajinami EÚ. Nad centrálnym dátovým skladoom bude vytvorená analytická vrstva, ktorá ponúkne nástroje na modelovanie, analýzu dát, spracovanie dát a ich vyhodnocovanie (prediktívne modely, simulácie, vizualizácie a podobne). Centrálny dátový sklad musí obsahovať funkcie na manažment dátovej vrstvy. Vzhľadom na to, že CDS bude obsahovať aj neanonymizované údaje, je nevyhnutné zabezpečiť dostatočnú ochranu osobných údajov v súlade s novovznikajúcou legislatívou EÚ pre Digitálnu ekonomiku. Centrálny dátový sklad bude umiestnený v eGovernment cloude.

Dátová kancelária verejnej správy

Problematika si vyžaduje odbornú koordináciu procesov životného cyklu dát ako zberu, čistenia a spracovania dát, a zároveň technickú prevádzku dátových skladov a odbornú asistenciu pri využívaní moderných analytických nástrojov. Takéto úlohy by bolo vhodné riešiť centrálné a profesionálne. Navrhujeme preto zriadenie Dátovej kancelárie verejnej správy, ktorá bude mať na starosti bázu dát verejnej správy a poskytovanie podpory a asistencie ostatným analytickým jednotkám.

Kancelária strategických vylepšovaní

Po získaní kapacít aj na operatívne vyhodnocovanie procesov takmer v reálnom alebo reálnom čase sa odporúča zaviesť aj Kanceláriu strategických vylepšovaní pre plné využitie potenciálu „Data-driven“ štátu. Takáto kancelária môže napríklad nadrezortne sledovať využívanie služieb verejnej správy a spokojnosť občanov, identifikovať slabé miesta na front-office úradov a v back-office, sledovať postupné zavádzanie reforiem verejnej správy do praxe prostredníctvom takzvaných smerodajných indikátorov, ktoré indikujú úspešnosť reformy ešte pred možnosťou zmerať hodnoty cieľových ukazovateľov. Na základe takéhoto nezávislého nadhľadu môže jednoducho prichádzať s riešeniami problémov alebo s vhodnými inováciami, ktoré môžu byť často pre ľudí hlboko ponorených do svojej problematiky skryté.

2.1 Ciele realizácie

2.1.1 Architektonické ciele

Táto kapitola obsahuje špecifikáciu architektonických cieľov¹, ktoré by mali byť dosiahnuté prostredníctvom realizácie strategickej priority.

Zoznam cieľov

Cieľ	Ukazovateľ cieľa	Čas dosiahnutia cieľa
Zlepšenie kvality politik v jednotlivých sektoroch verejnej správy	■ Ukazovatele budú identifikované pre konkrétne politiky	■ 2020

¹ Viď dokument - Biznis kontext a motivačný aspekt EA

Cieľ	Ukazovateľ cieľa	Čas dosiahnutia cieľa
Zlepšenie kvality regulačného prostredia	<ul style="list-style-type: none"> ■ Počet ovplyvnených rozhodnutí na základe nepriaznivého hodnotenia (ex-post) (Očakáva sa, že to bude 70 % z rozhodnutí s nepriaznivých hodnotením) ■ Počet upravených predkladaných materiálov na základe nesúhlasného stanoviska Stálej pracovnej komisie pre posudzovanie vybraných vplyvov (ex-ante) (Očakáva sa, že to bude 90 % z materiálov s nesúhlasným stanoviskom) ■ Zníženie regulačných nákladov podnikateľov (očakáva sa o 30 %) ■ Zníženie administratívneho zaťaženia občanov (očakáva sa o 40 %) 	<ul style="list-style-type: none"> ■ 2020
Zlepšenie kvality verejných investícií	<ul style="list-style-type: none"> ■ Zníženie prevádzkových nákladov IKT o 10 % napriek zvýšeným investíciám ■ Rozhodovanie na základe prepracovanej CBA, ktorej vypočítané prínosy a náklady sa budú v praxi striktne vyhodnocovať 	<ul style="list-style-type: none"> ■ 2020
Zvýšenie výkonnosti v jednotlivých sektoroch pri poskytovaní služieb občanom a podnikateľom	<ul style="list-style-type: none"> ■ Skrátenie dĺžky správneho konania ■ Skrátenie lehôt ■ Zníženie počtu odvolaní a rozkladov 	<ul style="list-style-type: none"> ■ 2020

Naplnenie cieľov OP II

Cieľ OP II	Ukazovateľ cieľa
7.5: Zlepšenie celkovej dostupnosti dát verejnej správy vo forme otvorených dát	<ul style="list-style-type: none"> ■ Počet nových datasetov publikovaných vo formáte s vysokým potenciálom na znovu použitie
7.7 Umožnenie modernizácie a racionalizácie verejnej správy IKT prostriedkami	<ul style="list-style-type: none"> ■ Dodatočný počet úsekov verejnej správy, v ktorých je rozhodovanie podporované analytickými systémami (cieľový stav sa očakáva 77) ■ Počet nových optimalizovaných úsekov verejnej správy

Naplnenie cieľov OP EVS

Cieľ OP EVS	Opatrenia a ukazovatele cieľa
Modernizované RLZ a zvýšené kompetencie zamestnancov	<p>Podpora analytických kapacít a tvorby verejných politík (vrátane vytvorenia analytickej jednotky na vykonávanie RIA):</p> <ul style="list-style-type: none"> ■ Počet ústredných orgánov štátnej správy s legislatívnou alebo právnou právomocou, ktoré získali podporu na zavedenie procesu RIA (cieľový stav sa očakáva 13). ■ Počet zamestnancov v analytických jednotkách v orgánoch štátnej správy na začiatku podpory. ■ Počet ústredných orgánov štátnej správy, ktoré získali podporu na zavedenie a/alebo posilnenie analytických jednotiek (cieľový stav sa očakáva 9)
Skvalitnené systémy a optimalizované procesy VS	<p>Opatrenia zamerané na tvorbu analýz, vstupov do reformných politík, hodnotenie dopadov politík a legislatívy:</p> <ul style="list-style-type: none"> ■ Zavedený funkčný systém RIA.

2.1.2 Stakeholderi a ich záujmy

V tejto časti budú identifikovaní stakeholderi, ktorých sa daná strategická priorita (alebo jej časť) týka a ich relevantné (záujmy) hnacie prvky¹.

Ministerstvo financií SR

- Zavedenie princípov štátu, fungujúcom na základe využívania dát, do praxe,
- Vykonávať dohľad nad efektívnosťou politík,
- Vykonávať dohľad nad efektívnosťou investícií,
- Mať prehľad o dopadoch opatrení na verejné financie a štátny rozpočet.

Ministerstvo vnútra SR

- Mať prehľad o výkone a kvalite služieb miestnej štátnej správy,
- Mať 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 vlády SR

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

Ministerstvo hospodárstva SR

- Posun k službám zameraným na nárast konkurencieschopnosti podnikateľského prostredia,
- Vytvorenie efektívneho regulačného rámca pre podporu podnikateľského prostredia,
- Koordinácia procesu hodnotenia vybraných vplyvov a vplyvov na podnikateľské prostredie, podľa kompetencie.

Analytické jednotky verejnej správy

- Mať 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, politík a regulácií (relevancia analýz v procese, nie tvorba „do zásuvky“),
- Získať rešpekt v odbornej komunite a uznanie verejnosti,
- Mať k dispozícii užitočné nástroje a IT infraštruktúru, kvalitné dáta a modely, aby bolo možné pohodlne vytvárať analytické produkty.

Verejnosť

- Mať prehľad o vykonávaných analýzach,
- Mať prehľad o predpokladaných dopadoch politík a regulácií,
- Mať prehľad o zdôvodnení investícií vo verejnom záujme,
- Mať prehľad o výkone verejnej správy v jednotlivých sektoroch (až na úroveň jednotlivých inštitúcií) – vrátane ratingov,
- Možnosť participatívnej účasti na návrhu politík, regulácií a inovácii služieb.

Podnikateľské prostredie

- Mať 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.

2.2 Organizácia

2.2.1 Zodpovednosť

Zodpovedný gestor

- Gestorom strategickej priority Analýza dát je **Ministerstvo financií SR** – zodpovedá za vytvorenie rámca pre lepšie využívanie dát, riadenie dát a vytvorenie Centrálného dátového skladu
- 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.

Zapojené organizácie

Zoznam organizácií, ktoré budú dotknuté realizáciou priority:

- Analytické jednotky, ktoré budú vytvorené v rámci jednotlivých sektorov (ministerstiev):
 - Inštitút finančnej politiky (MF SR),
 - Inštitút vzdelávacej politiky (MŠVVŠ SR),
 - Analyticko-metodická jednotka (MV SR),
 - Analytické centrum (MPSVR SR),
 - Odbor ekonomických nástrojov a analýz (MŽP SR),
 - Inštitút stratégie (MDVRR SR),
 - Inštitút zdravotnej politiky (MZ SR),
 - Odbor hlavného ekonóma (PMU),
 - Kancelária Rady pre rozpočtovú zodpovednosť,
 - Slovak Business Agency (SBA),
 - Analytická jednotka MH SR,
 - ďalšie vznikajúce analytické jednotky verejnej správy.

2.2.2 Organizačné zmeny

Aby bolo možné realizovať strategickú prioritu Riadenie údajov a „Big data“, sú navrhované nasledujúce organizačné zmeny:

- Zriadenie Dátovej kancelárie verejnej správy, ktorá bude mať na starosti správu dátovej analytickej vrstvy a prípravu nástrojov pre analytické jednotky, ako aj správu Centrálného dátového skladu a jeho informačnú a dátovú architektúru. Dátová kancelária verejnej správy by mala byť zriadená v rámci Ministerstva financií SR. Dátová kancelária bude mať na starosti centrálnu koordináciu analytických jednotiek vo vzťahu k transformácií verejnej správy na štát fungujúci na základe využívania dát.
- 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.
- Vytvorenie tímu zodpovedného za Riadenie dát (Data governance) teda za celý životný cyklus dát, v ktorom budú zástupcovia vlastníkov dát, dátoví stewardi 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.
- Zriadenie Kancelárie strategických vylepšovaní, ktorá sa bude venovať stratégií operatívnych vylepšovaní, sledovaniu kvality procesov a pokroku implementácie reformy VS, analýze procesov obsluhy v takmer reálnom čase, zavádzaniu inovácií a nasadzovaniu vylepšovaní.

2.3 Stratégia a riešenie

2.3.1 Popis súčasného stavu

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 (pôjde o jednu z priorit slovenského predsedníctva Európskej rady v 2016).

Dáta vo verejnej správe sa 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.

Z hľadiska analytického využívania dát je najďalej Ministerstvo financií SR. Prevádzkuje Centrálny konsolidačný systém (CKS), ktorý v určitej miere implementuje jednotlivé kroky procesu centrálného zberu, vyhodnocovania a publikácie dát. Inštitút finančnej politiky ako analytická jednotka Ministerstva financií SR má najväčšie skúsenosti s ekonomickým vyhodnocovaním dát.

Riziká súčasného stavu

Zoznam rizík, ktoré vyplývajú zo skutočnosti, ak by sa súčasný stav nezmenil:

- Verejná správa nevyužije potenciál, ktorý prinášajú nové technológie spracovania údajov a Big Data – z dát nevznikne vzácny zdroj a Slovenska republika nevytvorí jedinečné inovačné prostredie v oblasti dátových analýz,
- Zámery reformy verejnej správy v oblasti lepšieho návrhu politík a regulácií (RIA) sa nepodarí naplniť – regulačné prostredie v Slovenskej republike ostane suboptimálne,
- Nedostatok dát a ich nekonzistentnosť a nízka kvalita znemožnia správnu realizáciu a implementáciu reformy a nezvýši sa výkonnosť verejnej správy ani miera transparentnosti,
- Rozhodnutia vo verejnej správe nebudú podporované kvalitnou analýzou a založené na dôkazoch.

2.3.2 Najlepšie skúsenosti

Zoznam najlepších skúseností (best-practice) zo zahraničia, ktoré sú relevantné pre strategickú prioritu, respektíve sa z nich vychádza:

- E-government podporujúci lepšie regulácie v Dánsku,
- Spôsob, akým pristúpila k vyhodnocovaniu dopadov regulácií v USA OIRA (Office of Information and Regulatory Affairs) od roku 2011 – regulatory.gov a reginfo.gov,
- Európsky štatistický systém- Vízia 2020 a Akčný plán a Roadmapa pre Big Data.

Slovensko bude ďalej sledovať výsledky a podporovať zapojenie inštitúcií a podnikov do európskych aktivít výskumu a inovácií v oblasti Prípravy politík a regulácií v ére Big Data (spoločenská výzva 6 - BigDataEurope (BDE)) ako aj v jednotlivých tzv. Lighthouse projektoch, ktoré budú prichádzať s novým prístupom k Big Data a dátovej vede v jednotlivých sektoroch ako doprava a zdravotníctvo.

Slovensko bude tiež nasledovať nové regulačné opatrenia spojené s piatou slobodou Jednotného trhu, a tou 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 princípu *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.

2.3.3 Strategický prístup k riešeniu

Prístup k systematickému riešeniu Analýzy dát je možné zhrnúť ako:

- Ako súčasť reformy verejnej správy sa vytvorí kapacita (formou Dátovej kancelárie verejnej správy) schopná centralizovane a systematicky riadiť a koordinovať používanie dát vo verejnej správe:
 - Manažment dátového modelu verejnej správy (vrátane ontológie a definície prepojení),
 - Spolupráca pri návrhu analytických produktov, definícia požiadaviek na dátové zdroje,
 - Riadenie zberu dát z informačného prostredia verejnej správy,

- Riadenie získavania dát z externého prostredia, jeho princípov a pravidiel,
- Definovanie preferovaných platforiem pre prácu s dátami.
- Vytvorí sa stratégia pre využívanie Big Data vo verejnej správe spolu s taxonómiou (schémou klasifikácie) pre Big Data, ktorá zadefinuje v ontológii ako OWL:
 - Procesy (zber, kurátorstvo, reprezentácia, ukladanie, riadenie, analýza, vizualizácia dát a podobne),
 - Role (napríklad dátový architekt, dátový vedec, vlastník),
 - Charakteristiku dát (rôznorodosť, rýchlosť, objem, hodnovernosť a podobne),
 - Informačné artefakty (ako graf, tabuľka, dátový model, databáza, flat súbor),
 - Informačné nosiče (ako cluster, DVD,),
 - Vzťahy medzi elementmi Big Data.
- **Presne sa definujú a kategorizujú osobné a citlivé údaje** ako aj údaje pre verejné blaho v ére senzorov, internetu vecí („internet of Things“), inteligentných internetových prehliadačov a softvérových riešení,
- **Zavedú sa najlepšie praktiky pre technické riešenia ochrany osobných údajov** ako pre riadenie prístupov k dátam a pre anonymizáciu a pseudonymizáciu vrátane identifikovania presných scenárov, kedy musia byť použité,
- Na jednotlivých ministerstvách budú zriadené analytické jednotky zodpovedné za konkrétne oblasti verejnej správy (politiky, regulácie, výkonnosť, investície). Jednotlivé analytické jednotky dostanú na starosti zoznam konkrétnych analytických produktov, ktoré budú vypracovávať. Činnosť analytickej jednotky bude maximálne transparentná a jej výsledky dostupné prostredníctvom webového sídla pre publikáciu analýz dát. Dá sa očakávať konkurencia analytických jednotiek a systém riadenia podporujúci inováciu.
- Analytické jednotky vytvoria a implementujú vlastné analytické modely a nástroje, pričom budú využívať dostupné dáta najlepším spôsobom. Pravidlom bude, že každá analytická jednotka bude mať prístup k relevantným dátam v dátovom sklade (anonymizovaným alebo pseudoanonymizovaným, v prípade dostatočne odôvodenej potreby aj k neanonymizovaným), ako i k analytickým výstupom ostatných jednotiek, ktoré si budú môcť aj zdieľať na „peer review“.
- Vznikne aj Kancelária strategických zlepšení, ktorá na základe analýzy dát v reálnom alebo takmer reálnom čase dokáže prichádzať s operatívnymi vylepšeniami a inováciami naprieč celou verejnou správou.
- Implementuje sa informačný systém Centrálny dátový sklad, ktorý bude zhromažďovať dáta z verejnej správy a umožní ich analytické využitie. CDS bude umiestnený v eGovernment cloude a založený na najnovších technológiách podporujúcich aj Big Data (teda SQL aj noSQL databázy).
- Jednotlivé informačné systémy verejnej správy budú pripájané k Centrálnemu dátovému skladu cez Modul úradnej komunikácie (MUK), ktorý implementuje dátovú integráciu a ETL procesy.

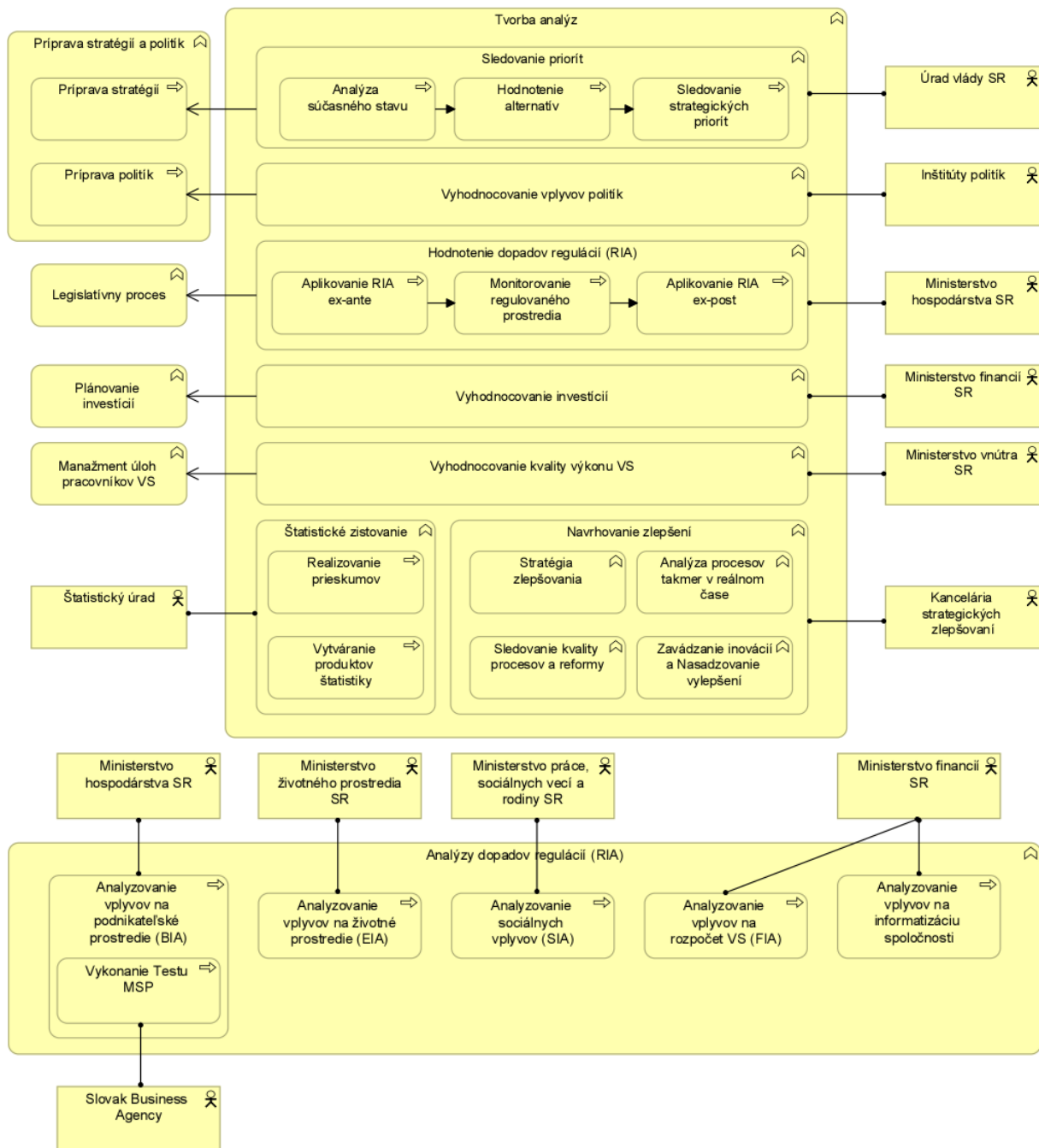
Princípy

Zoznam architektonických princípov¹, na ktorých je riešenie postavené:

- Údaje sú aktívne,
- Údaje sú zdieľané,
- Údaje sú kvalitné,
- Údaje majú vlastníka a správcu,
- Spoločné údajové ontológie a definícia konceptov,
- Súkromie ako súčasť dizajnu riešení („Privacy by design“).

2.3.4 Biznis vrstva

Obrázok 1: Pohľad na biznis vrstvu v oblasti Analýz dát



Požiadavky

Zoznam požiadaviek:

- Riešenie podporuje workflow,
- Riešenie obsahuje prednastavené modely,
- Riešenie obsahuje mechanizmy pre zabezpečenie kvality dát,
- Riešenie podporuje dokumentáciu a vyhľadávanie v metaúdajoch,

- Analýzy a modely je možné vykonávať s dátami z ľubovoľných zdrojov vrátane Big Data,
- Analýzy je možné vykonávať online nad navrhnutým modelom, a je možné spolupracovať aj s externými partnermi,
- Výstupy analýz je možné referencovať, zdieľať, vyhľadávať, komentovať a sú dostupné cez jednotný prístup.

Služby

Zoznam služieb:

- Zaevidovanie pripravovaného materiálu predkladateľom, ktorý bude mať k dispozícii nástroje na modelovanie plánovaných opatrení a kontextu regulovaného prostredia,
- Identifikácia dopadov regulácie a ich rozsahu,
- Modelovanie dopadov regulácie alebo politiky,
- Vykonalie konzultácie,
- Kolaborácia nad analýzami,
- Zdieľanie výstupov a pripomienkovanie,

Procesy a funkcie

Základné funkcie pre Tvorbu analýz:

- Legislatívny proces
- Modelovanie legislatívneho prostredia
- Sledovanie priorít:
 - Analýza súčasného stavu
 - Hodnotenie alternatív
 - Sledovanie strategických priorít
- Vyhodnocovanie politík
- Hodnotenie dopadov regulácií (RIA):
 - Aplikovanie RIA ex-ante
 - Monitorovanie regulovaného prostredia
 - Aplikovanie RIA ex-post
- Vyhodnocovanie investícií
- Vyhodnocovanie výkonu verejnej správy
- Navrhovanie zlepšení:
 - Stratégia zlepšovania
 - Sledovanie kvality procesov a reformy
 - Analýza procesov v takmer reálnom čase
 - Zavádzanie inovácie
 - Nasadzovanie vylepšení.

Súvisiace funkcie, ktoré používajú služby tvorby analýz:

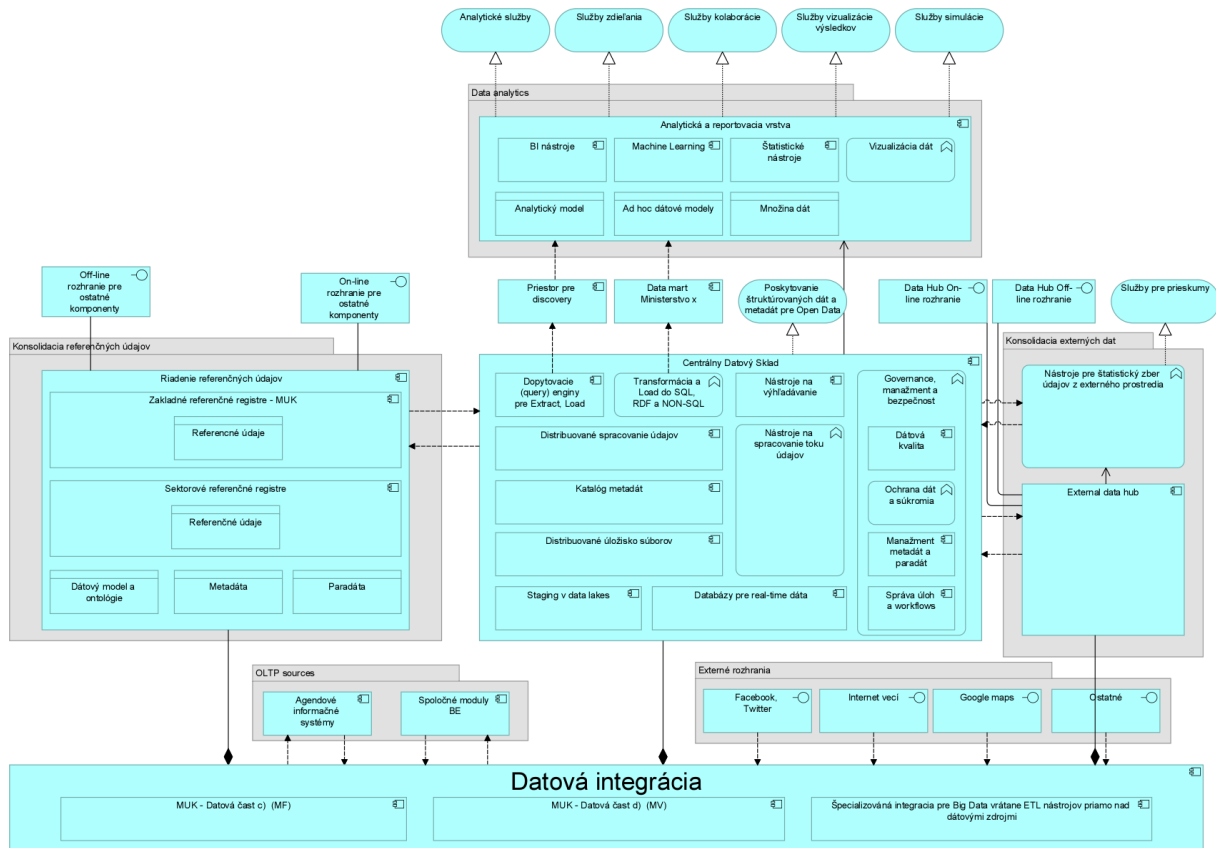
- Príprava stratégií (Sledovanie priorít)
- Príprava politík (Vyhodnocovanie politík)
- Legislatívny proces -> Príprava regulácií (Analýzy RIA)
- Plánovanie investícií (Vyhodnocovanie investícií)
- Manažment kvality verejnej správy (Vyhodnocovanie výkonu verejnej správy)

Súvisiace funkcie, ktoré dopĺňajú komplexný systém:

- Monitoring dát
- Prístup k informáciám

2.3.5 Aplikačná vrstva

Obrázok 2: Pohľad na aplikačnú vrstvu pre Analýzu dát



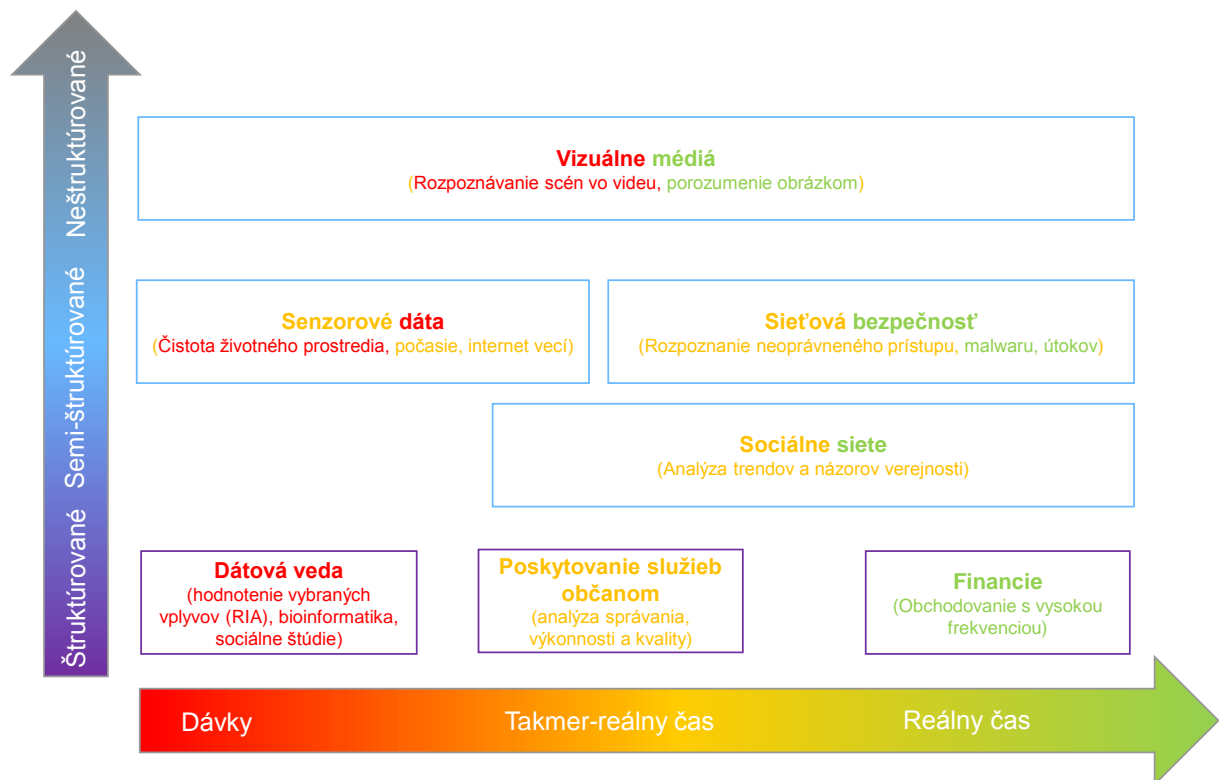
Požiadavky,

Zoznam požiadaviek

- Je možné nastaviť spôsob aktualizácie dát, a tým pádom oneskorenie medzi vznikom udalosti a potrebou na danú udalosť reagovať: dávkovo (oneskorenie je vyššie ako 200 ms), online (v reálnom čase (oneskorenie je nižšie ako 50 ms) alebo takmer v reálnom čase (oneskorenie je medzi 100 a 200ms)) a na základe udalosti,
- Riešenie podporuje ontológie a dátový model je možné spravovať centrálnne, pričom je možné spravovať aj ad hoc dátové modely a nové dátové prvky je možné navrhovať aj dynamicky,
- Riešenie umožňuje vykonávanie a logovanie kontrol dátovej kvality počas celého procesu spracovania údajov,
- Riešenie umožňuje vhodne ukladať a spravovať:
 - Štruktúrované dáta, spravidla ukladané v relačných databázach alebo tabuľkových procesoroch podľa databázového modelu, ktorý určuje, do ktorých polí dáta patria, aký majú dátový typ, aké sú obmedzenia pre ich hodnotu ako aj pre celú množinu elementov za účelom dosiahnutia konzistencie dát (napríklad žiadne duplikáty).
 - Neštruktúrované dáta nemajú preddefinovaný dátový model alebo nie sú organizované daným spôsobom. Ide väčšinou o rozsiahle texty, ale neštruktúrované informácie môžu obsahovať aj dáta ako dátumy, čísla a fakty. Do tejto kategórie spadajú tiež surové dáta (bez značiek - tagov) ako fotky a grafické obrázky,

videá, dáta zo senzorov, webové stránky, pdf súbory, prezentácie v PowerPointe, e-maily, blogy, wiki a dokumenty spracované v textovom procesore.

- Semi-štruktúrované dáta predstavujú typ štruktúrovaných dát, avšak nemajú striktnú štruktúru danú dátovým modelom. Značky – tagy sa používajú na označenie istých elementov v dátach, ale dáta nemajú rigidnú štruktúru, z ktorej by sa dal jednoducho extrahovať kompletný sémantický význam bez potreby ďalšieho rozsiahleho spracovania. Príkladom je pridávanie metadát k dokumentom z textového procesora ako autor, dátum vytvorenia, obsah, kapitoly a sekcie. E-maily majú odosielateľa, prijímateľa, dátum, čas a ďalšie fixné polia pridané k neštruktúrovanému textu a prílohám. Podobne možno pristúpiť k fotkám a obrázkom. XML alebo iné značkovacie jazyky sa používajú na spravovanie semi-štruktúrovaných dát.
- Riešenie ďalej určí relevantné domény pre verejnú správu spolu s presnými dátovými zdrojmi a ich možnosťami aktualizácie a štruktúry ako ukazuje nasledujúci obrázok:



Prevzaté z dokumentu: „Cloud Security Alliance: Bug Data Taxonomy“

- Súvisiace riadenie dát obsahuje holistický súbor štandardizovaných procesov a usmernení, ktoré adresujú to, ako sa majú dáta získavať, spravovať, ukladať a ako sa k nim má pristupovať, aby používatelia mali vždy istotu, že majú správnu verziu dát zo spoľahlivého zdroja. Tieto procesy sa musia tiež odkazovať aj na patričné technológie, ktoré sú využívané na tieto účely. Tieto procesy a technológie musia adresovať aspoň nasledujúce:
 - Získavanie dát zo zdrojov,
 - Čistenie dát a integrácia – ide o proces opravy alebo odstraňovania dát, ktoré sú nesprávne, nekompletné, zle naformátované ale duplicitné,
 - Transformácia,
 - Anonymizácia a pseudonymizácia,
 - Kontrola kvality,
 - Kontrola duplicit,
 - Verzionovanie,
 - Manažment úložísk,

- Archivovanie,
 - Prepojenie na Centrálnu správu referenčných údajov, aby spracovávané referenčné údaje boli vždy aktuálne,
 - Bezpečnosť a súlad s pravidlami a štandardmi,
 - Aplikovanie rôznych politik na ochranu citlivých údajov.
- Riešenie podporuje spracovanie dát:
 - V dávkach (paralelné spracovanie veľkých datasetov na komoditných výpočtových zdrojoch):
 - MapReduce;
 - Bulk Synchronous Parallel;
 - V reálnom alebo takmer-reálnom čase v prúde dát.
 - Riešenie umožňuje tvorbu takých analytických prediktívnych modelov, ktoré môžu byť vždy aktualizované na základe aktuálnych dát (to znamená, že model je možné v každom okamihu v čase prepočítať na základe aktuálnych dát, čím je možné overiť jeho prediktívne schopnosti a prípadne ho vylepšiť, systém sa tak stará o neustálu kompatibilitu historických dát s aktuálnymi a o prepojenia modelov s aktualizovanými množinami dát),
 - Riešenie umožňuje vyberať si do modelov predpripravené číselné rady ako napríklad historický vývoj a prognózu HDP, zamestnanosti a podobne,
 - Riešenie obsahuje základné štatistické a analytické nástroje,
 - Riešenie podporuje distribuovanú prácu s dátami,
 - Riešenie podporuje proces tvorby analytického reportu,
 - Riešenie umožňuje návrh interaktívnych vizualizácií,
 - Riešenie podporuje vytváranie API na zdieľanie a prístup k rôznym datasetom vrátane datasetov Open Data,
 - Riešenie podporuje architektúru webových služieb pre jednoduché lokalizovanie dát, dátových zdrojov a metadát v distribuovanej sieťovej architektúre.

Aplikačné služby

Zoznam služieb

- Analytické služby, ktoré používajú predovšetkým dátoví analytici venujúci sa napríklad Hodnoteniu dopadu vybraných vplyvov alebo pracujúci v analytických jednotkách. Tieto služby poskytujú prístup k celému stacku Centrálného dátového skladu – na základe oprávnenia k anonymizovaným, pseudonymizovaným alebo neanonymizovaným údajom pri dodržiavaní najlepších praktík ochrany súkromia v ére Big Data v súlade s najnovšou európskou reguláciou pre všeobecnú ochranu dát (2012/0011 (COD) General Data Protection Regulation)
- Služby zdieľania, na základe ktorých je možné zdieľať výsledky analýz alebo analytický model nad definovanou množinou dát z Centrálného dátového skladu. **Základným pravidlom tvorby modelov je, že ich je vždy možné prepočítať na základe aktuálnych dát** (to znamená, že predpovedané dáta sa zmenia na aktuálne, a ich porovnaním možno ďalej vylepšovať a kalibrovať modely).
- Služby kolaborácie, pomocou ktorých môžu zamestnanci analytických jednotiek spolupracovať nielen interne na tvorbe analytického modelu, ale aj s externými špecialistami,
- Služby vizualizácie výsledkov, ktoré umožňujú pre zamestnancov alebo pre verejnosť vidieť výsledky analýz v interaktívnych grafoch,
- Služby simulácie, ktoré umožňujú simulovať historický aj budúci priebeh modelovaných udalostí nad vybranou množinou historických a aktuálnych dát,
- Služby pre prieskumy budú poskytovať občanom a podnikateľom čo najefektívnejšie nástroje pre účasť na štatistickom zisťovaní, väčšinou napojené na API informačných systémov, ktoré využívajú. Budú k dispozícii

tiež nástroje pre návrh štúdie a potrebného prieskumu, vytvorenie dotazníka, podporu zberu dát a ich spracovania, agregáciu dát a spravovanie databázy otázok a konceptov,

- Poskytovanie štruktúrovaných dát a metadát pre Open Data.

Aplikačné komponenty

Zoznam základných komponentov:

- Centrálny dátový sklad (ktorý pripája agendové informačné systémy cez MUK a externé zdroje cez špecializovanú integráciu a externý dátový hub), ktorý umožňuje ukladanie, manipulovanie a riadenie dát a Big Data v rôznych typoch databáz:
 - Relačné databázy (SQL) – sú jednoduché, avšak majú výrazné obmedzenie ich veľkosti,
 - NoSQL databázy (dokumentovo-orientované, sklady s „key-value“, inšpirované veľkými tabuľkami alebo grafovo orientované),
 - New SQL.

Vhodnosť jednotlivých databáz sa musí vyhodnotiť pre rôzne prípady použitia na základe aspoň nasledujúcich metrik:

- Atomicnosť – ak jedna časť transakcie sa nevykoná úspešne, nevykoná sa žiadna časť transakcie a stav databázy ostane nezmenený,
- Konzistentnosť – to znamená, že každá transakcia prenesie databázu z jedného platného stavu do druhého platného stavu,
- Izolácia – zabezpečuje, že vykonanie transakcií paralelne prinesie databázu do toho istého stavu, ako ich sériové vykonanie,
- Odolnosť – ak sa raz transakcia ukončí úspešne, jej výsledok ostane zapísaný aj v prípade chyby, výpadku zdroja napätia alebo inej poruchy,
- Vhodnosť pre štruktúrované dáta alebo neštruktúrované dáta,
- Vhodnosť pre dávkové spracovanie alebo spracovanie v reálnom čase.

Centrálny Dátový Sklad neduplikuje konsolidačnú logiku CSRÚ, ale v prípadoch, keď je to možné, načíta konsolidované dáta priamo z CSRÚ. V prípadoch, kedy to možné, nie je (napr. história), načíta dáta priamo zo zdrojových systémov. Centrálny Dátový sklad obsahuje ďalšie komponenty pre prácu s dátami napr. nástroje pre dátovú kvalitu, ETL nástroje a podobne. V detailnej analýze je možné posudzovať varianty, v ktorých tieto platformy môžu byť zdieľané s komponentami MUK, poprípade s inými komponentami.

CDS poskytuje rôzne úrovne skladovania dát, od data lakes až po relačné databázy, ktoré možno poskytovať ako klasické dáta podľa doménových členení, v ktorých budú prebiehať analýzy dát špecifické pre jednotlivé domény. Rovnako je možné robiť analýzy dát priamo v Centrálnom Dátovom Sklade. Dáta možno prehľadávať vďaka sofistikovaným katalógom metadát a vyhľadávacím nástrojom. Nad celým skladom bdie governance, manažment a bezpečnosť. Bezpečnosť sa stará o súkromie údajov (vďaka anonymizácii a pseudonymizácii, kryptografii a úrovniam prístupu s rôznou granularitou), správu dát (transakčné logy, auditu na najnižšej úrovni granularity, data provenance) a ich integritu vďaka monitoringu v reálnom čase, validácii koncového bodu a filtrovaniu.

- Pre požiadavky práce s externými informáciami vznikne samostatný komponent "Data hub", ktorý pomocou on-line alebo off-line rozhraní bude načítavať štruktúrované alebo neštruktúrované dáta z externých zdrojov. Tieto dáta bude následne konsolidovať a potom poskytovať rozhranie pre ich ďalšiu publikáciu do vnútornej IT architektúry.
- Analytická a reportovacia vrstva, ktorá umožňuje vykonávať analýzy nad dávkovými údajmi aj nad dátovými tokmi pre operatívne analýzy. Poskytuje:
 - Tradičné BI nástroje:
 - Machine Learning:
 - Štatistické nástroje:

- Vrstva dátové integrácie sa skladá z Modulu úradnej komunikácie (MUK), ktorý zaisťuje prepojenie na systémy verejnej správy, a modulu špecializované integrácie pre Big Data, ktorý umožňuje získavať dáta napríklad zo sociálnych sietí, IoT zariadení, Google maps a podobne.
- Security nástroje umožňujú riadenie prístupov užívateľov podľa klasifikácie dát (na základe ich citlivosti) a používateľov (na základe ich zaradenia, role, ...). Súčasťou týchto funkčností je aj možnosť auditovanie udalostí.
- V rámci riešenia budú implementované aj nástroje na sledovanie dátovej kvality a sledovanie riešenie identifikovaných problémov.
- Za účelom správy biznis i technických metadát budú implementované špecializované nástroje, ktoré zároveň podpora i Governance celého BI riešeni.
- Big Data platform bude slúžiť pre spracovanie veľkého množstva neštruktúrovaných dát napríklad zo sociálnych sietí, alebo IoT zariadení.
- Rule Engine/BPM
- ECM / DMS.
- MDM/RDM
- Užívatelia budú môcť ku konsolidovaným dátam, reportom a dashboardům pristupovať cez jednotný portál.
- Za účelom dokumentácie riešenia, väzieb medzi systémami a dáta budú použité nástroje pre dátové modelovanie.
- Komunikační/notifikační nástroje

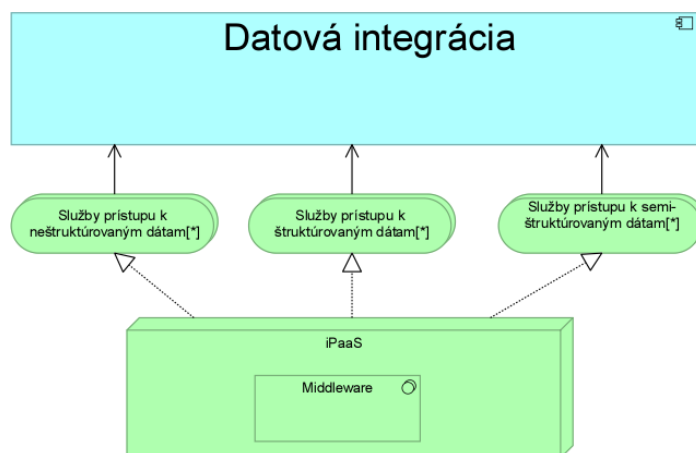
Zoznam spolupracujúcich komponentov (pracujú s výsledkami):

- IS pre inteligentné regulácie
- IS pre manažment kvality

2.3.6 Technologická vrstva

Informačný systém pre Centrálny dátový sklad bude umiestnený v eGovernment cloude a v čo najvyššej možnej miere bude využívať integračnú platformu ako služby pre prístup k neštruktúrovaným, štruktúrovaným a semi-štruktúrovaným dátam.

Obrázok 3: Pohľad na technologickú vrstvu pre Analýzu dát



2.3.7 Predpoklady realizácie

- Stanovenie priorit a stratégie Slovenska v oblasti Dátového hospodárstva.

- Prijatie novej General Data Protection Directive začiatkom roku 2016.

2.4 Problémy a riziká

Dôsledky

Zoznam dôsledkov, ktoré vyplývajú z realizácie strategickej priority pre jednotlivé subjekty:

- Informačné systémy verejnej správy je potrebné integrovať s CDS (cez MUK),
- Zmení sa proces pri posudzovaní vplyvov regulácií, návrhu stratégií a politík, pri štatistickom zisťovaní ako aj pri optimalizovaní chodu verejnej správy predovšetkým v oblasti poskytovania služieb občanom a podnikateľom,
- Vo verejnej správe bude potrebné zamestnať väčší počet dátových analytikov a ekonómov.

Problémy

Zoznam problémov, ktoré bude potrebné vyriešiť počas realizácie strategickej priority:

- Ako navrhnuť prioritizáciu vzniku dátových skladov, pripájanie informačných systémov do dátového skladu a následné nasadzovanie analytických produktov,
- Ako zabezpečiť ochranu osobných a citlivých údajov a údajov týkajúcich sa národnej bezpečnosti,
- Ako potlačiť rezortné silové uvažovania a neochotu zdieľať údaje,
- Ako získať dostatočné kapacity, schopné spravovať dáta a vytvárať analytické produkty – nástroj European Data Market Monitoring mapuje stav dopytu a ponuky na trhu dátového hospodárstva,
- Ako konsolidovať rôzne aktivity jednotlivých inštitúcií verejnej správy týkajúcich sa analytického spracovania (vrátane konsolidácie informačného prostredia).

Riziká

Zoznam rizík, ktoré je dôležité mitigovať pri realizácii strategickej priority:

- Návrh aplikačnej architektúry a dátového modelu bude príliš komplexný a nepoužiteľný v praxi,
- Nepodarí sa získať dostatočný počet inštitúcií, ktoré sa zapoja do centrálného riešenia (budú sa uprednostňovať lokálne riešenia),
- O analytické produkty nebude záujem (verejnosti, vedenia ministerstva),
- Kvalita dát bude nízka,
- Nepodarí sa transformovať rozhodovacie procesy vo verejnej správe,
- Nepodarí sa implementovať ambicióznú víziu Big Data.

2.5 Legislatívne požiadavky

Legislatívne požiadavky

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

Legislatívny predpis	Navrhované opatrenie
Nový zákon o dátach vo verejnej správe	Nový zákon o dátach vo verejnej správe: <ul style="list-style-type: none"> – špecifikuje publikovanie otvorených dát, – modernizuje spôsob používania dát vo verejnej správe – všetky dáta bude možné používať na analytické účely a pre podporu rozhodovania ak sa zabezpečí ochrana osobných údajov, – stanoví rámec pre anonymizáciu dát, – nastaví práva vlastníkov dát.
Aktualizovaná Jednotná metodika na posudzovanie vybraných vplyvov.	Transformácia procesov posudzovania vybraných vplyvov tak, aby podporovali využívanie analýz vytvorených v systéme.
	Nová kategorizácia a klasifikácia údajov
	Zriadenie Dátovej kancelárie verejnej správy
	Zriadenie kancelárie strategických zlepšení
	Zriadenie štandardov pre analytickú vrstvu vo verejnej správe, internet vecí a Big Data
	Návrh transformácie rozhodovacích procesov vo verejnej správe
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
Všeobecné nariadenie COM(2012)11 o ochrane údajov	Zohľadniť pri technických realizáciách

Požiadavky na Operačný program 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.

2.6 Plánovanie a migrácia

Identifikácia projektov, ktoré je vhodné realizovať:

- Centrálny dátový sklad a analytické nástroje,
- Inteligentné regulácie,
- Manažment kvality a výkonu verejnej správy,
- Nové nástroje pre Národný štatistický systém,

- IT nástroje pre test MSP,
- Analýzy sociálnych politík
- *Je možné počítať s ďalšími projektami pre jednotlivé analytické jednotky.*

2.6.1 Centrálny dátový sklad a analytické nástroje

Popis a cieľ

Cieľom projektu je vybudovanie technologickej infraštruktúry, integrácia zdrojov dát, naplnenie relevantnými dátami návrh analytických a dátových modelov a zavedenie používania analytických nástrojov do reálnej praxe.

Aktivity

- Analýza a návrh riešenia,
- Implementácia CDS,
- Implementácia analytických nástrojov,
- Vytvorenie modelov,
- Integrácia zdrojov dát,
- Testovanie.

Navrhovaný realizátor

- Ministerstvo financií SR

Navrhovaný partneri

- IFP,
- Inštitúcie verejnej správy, ktoré budú integrované.

Odhadovaná dĺžka realizácie projektu (projektov)

- 3 roky (viacero fáz, výstupy postupne)

Závislosti na iných projektoch a prioritách (OPII a OPIS)

- Centrálny konsolidačný systém

Závislosti na iných projektoch a prioritách (OPEVS)

- Vybudovanie dátovej kancelárie verejnej správy,
- Vznik analytických jednotiek.

2.6.2 Inteligentné regulácie

Popis a cieľ

Projekt zameraný na vytvorenie systému evidujúceho platné regulácie, ktorý bude pravidelne aktualizovaný na základe vyhodnotenia efektivity platných regulácií. Ku každej regulácii sa bude v ňom nachádzať interaktívna analýza prínosov a nákladov, ktorá bude napojená na analytické nástroje, umožňujúce simulácie budúcich dopadov.

Aktivity

- Analýza a návrh riešenia,
- Implementácia IR,
- Konfigurácia registra regulácií,
- Vytvorenie modelov RIA,
- Integrácia zdrojov dát,
- Testovanie.

Navrhovaný realizátor

- Ministerstvo hospodárstva SR

Navrhovaný partneri

- Slovak business agency

Odhadovaná dĺžka realizácie projektu (projektov)

- 3 roky (viacero fáz, výsledky budú dodávané približne)

Závislosti na iných projektoch a prioritách (OPII a OPIS)

- Centrálny dátový sklad,

Závislosti na iných projektoch a prioritách (OPEVS)

- Zriadenie Centra lepšej regulácie,
- Podpora koordinácie procesov RIA a agendy lepšej regulácie.

2.6.3 Manažment kvality a výkonu verejnej správy

Popis a cieľ

V rámci projektu sa navrhne a implementuje systém na manažment kvality výstupov a výsledkov inštitúcií verejnej správy. Pôjde o systém, ktorý umožní definovať a sledovať KPI podľa výberu a plánovanie optimalizačných stratégií.

Aktivity

- Analýza a návrh riešenia,
- Implementácia manažmentu a kvality výkonu VS,
- Implementácia analytických nástrojov,
- Vytvorenie modelov,
- Integrácia zdrojov dát,
- Testovanie.

Navrhovaný realizátor

- Ministerstvo vnútra SR

Navrhovaný partneri

- Inštitúcie verejnej správy

Odhadovaná dĺžka realizácie projektu (projektov)

- 4 roky

Závislosti na iných projektoch a prioritách (OPII a OPIS)

- Centrálny dátový sklad

Závislosti na iných projektoch a prioritách (OPEVS)

- Meranie efektívnosti poskytovania služieb,
- Vytvorenie Kancelárie strategických zlepšování,
- Zavedenie manažmentu kvality.

2.6.4 Nové nástroje pre Národný štatistický systém

Popis a cieľ

V súčasnosti majú štatistické systémy problém vymieňať si údaje s externými systémami ako aj popisovať kontext, v ktorom boli dáta vyprodukované, pomocou paradát. Cieľom je preto posunúť štátnu štatistiku do novej éry Big Data a centralizovať služby štatistického zisťovania pre celú verejnú správu. Inovuje sa zber údajov, zníži sa administratívna záťaž podnikateľov a zvýši sa kvalita dát a analýz.

Aktivity

- Analýza a návrh riešenia,
- Implementácia - Pilotné riešenie centrálného komponentu pre zber, vykonanie pilotných zberov a analýz,
- Integrácia zdrojov dát,
- Testovanie - Otestované pilotné riešenie centrálného komponentu pre zber, implementované zmeny z pilotných iterácií,
- Nasadenie - Produkčné prostredie centrálného komponentu pre zber,
- Školenia.

Navrhovaný realizátor

- Štatistický úrad Slovenskej republiky

Navrhovaný partneri

- Vybrané segmenty VS minimálne v rozsahu organizácií tvoriacich Národný štatistický systém

Odhadovaná dĺžka realizácie projektu (projektov)

- 3 roky (viacero fáz, výsledky budú dodávané pribežne)

Závislosti na iných projektoch a prioritách (OPII a OPIS)

- Centrálny dátový sklad,

Závislosti na iných projektoch a prioritách (OPEVS)

- Reforma zberu a spracovania štatistických údajov

2.6.5 IT nástroje pre test MSP

Popis a cieľ

Projekt poskytne nástroje pre zapojenie malých a stredných podnikateľov do prípravy regulácií a zabezpečenie analýz dopadov na MSP (realizácia MSP testu).

Aktivity

- Analýza a návrh riešenia,
- Implementácia a integrácia modulov MSP test,
- Integrácia zdrojov dát,
- Testovanie.

Navrhovaný realizátor

- SBA

Navrhovaný partneri

- N/A

Odhadovaná dĺžka realizácie projektu (projektov)

- 1 rok

Závislosti na iných projektoch a prioritách (OPII a OPIS)

- Inteligentné regulácie

Závislosti na iných projektoch a prioritách (OPEVS)

- Projekt vytvorenia analytickej jednotky pre RIA a MSP test – Centrum lepšej regulácie

2.6.6 Analýzy sociálnych politík

Popis a cieľ

Projekt bude zameraný na nové metódy zberu a spracovania dát v sociálnej oblasti a vytvorenie dátového modelu sociálnych interakcií.

Aktivity

- Analýza a návrh riešenia,
- Implementácia dátových modelov,
- Implementácia nástrojov pre zber sociálnych dát,
- Integrácia zdrojov dát,
- Testovanie.

Navrhovaný realizátor

- Ministerstvo práce, sociálnych vecí a rodiny SR

Navrhovaný partneri

- N/A

Odhadovaná dĺžka realizácie projektu (projektov)

- 2 roky (viacero fáz, výsledky budú dodávané pribežne)

Závislosti na iných projektoch a prioritách (OPII a OPIS)

- Centrálny dátový sklad,

Závislosti na iných projektoch a prioritách (OPEVS)

- N/A

3 Riadenie dát

Dáta sú dôležitým aktívom každej inštitúcie, preto je potrebné ako každé iné aktívum ich riadiť v životnom cykle. Dáta v špecifickom, zmysluplnom kontexte vytvárajú informácie a rozširujú individuálne aj kolektívne znalosti, ktoré následne umožňujú realizovať aktivity vedúce k efektívnemu napĺňaniu cieľov verejnej správ. Riadenie dát je centralizovaný program naprieč celou verejnou správou, je to súhrn pravidiel, štandardov, procesov, ľudí a technológií, ktoré slúžia k dosiahnutiu stanovených cieľov. Základnými oblasťami riadenia dát sú: dátová kvalita, dátová bezpečnosť, štandardizácia dát, konsolidácia dát a informačná využiteľnosť. Aby bolo možné riadiť dátovú vrstvu efektívne, v súlade s navrhnutými cieľmi, je potrebné sa cielene venovať údajom počas všetkých etáp ich životného cyklu (pri vzniku, distribúcií, pri používaní, pri ich aktualizácii a archivácii).

Riadenie dát prináša nové procesy, role a zodpovednosti. Základnými rolami sú vlastníci dát, dátový steward a dátový špecialista.

Zavedenie a nastavenie procesov riadenia dát umožní zabezpečiť, že v každom kontexte riešenia problému budú k dispozícii všetky potrebné dáta, v požadovanej kvalite a čase, ktoré má verejná správa k dispozícii a ku ktorým má príslušný zamestnanec verejnej správy respektíve automatizovaný informačný systém oprávnenie. Znamená to potrebu rozšíriť Modul úradnej komunikácie tak, aby podporil komplexné riadenie dát. Táto integrácia na úrovni dátovej vrstvy zabezpečí, že jednotlivé povinné osoby cez informačné systémy verejnej správy budú musieť:

- Publikovať svoje primárne dáta (cez štandardizované, ale aj cez proprietárne rozhrania),
- Pripájať sa na sekundárne dáta (cez štandardizované rozhrania).

K základným funkciám platformy dátovej integrácie patrí:

- **Manažment prístupu k dátam** - v rámci riešenia bude zabezpečený spoločný a jednotný pohľad na všetky dáta vo verejnej správe. Ochrana osobných a citlivých údajov si vyžiada prepracovaný manažment prístupu k dátam, ktorý zabezpečí, aby k dátam mohli pristupovať a manipulovať s nimi len oprávnené osoby. Občanom sa tiež umožní (v súlade so Všeobecným nariadením COM(2012)11 o ochrane údajov) získať elektronický prístup k všetkým osobným údajom, ktoré sa o nich vedú, pričom budú informovaný o situáciách, v ktorých dochádza k narábaniu s ich dátami. Súčasťou riešenia budú definície pravidiel, na základe ktorých budú môcť agendové aplikácie a systémy funkčných blokov pristupovať k dátovým zdrojom iných systémov, čím sa zabezpečí zjednodušenie a automatizácia procesov. Povinnosťou vlastníka údajov je preto špecifikovať pravidlá a prístup k jeho agendovým údajom. Jednotlivé dátové zdroje budú kategorizované podľa obsahu na: neštruktúrované dáta; znalosti; agendové dáta, referenčné údaje v registroch a geografické dáta.
- **Dátová integrácia** – prepojenie agendových systémov a ich sekundárnych registrov na dátovej úrovni. Cieľom dátovej integrácie je v čo najväčšej miere uľahčiť tok dát a informácií. K tomu je nevyhnutné zosúladiť definície údajov, vytvoriť štandardy, určiť kritéria dátovej kvality, identifikovať dátové toky. Riešenie bude predstavovať samostatný komponent oddelený od vrstvy „procesnej integrácie“, ktorej sa venuje strategická priorita Integrácia FE/ BE.

3.1 Ciele realizácie

3.1.1 Architektonické ciele

Táto kapitola obsahuje špecifikáciu architektonických cieľov², ktoré by mali byť dosiahnuté prostredníctvom realizácie strategickkej priority.

² Vid' dokument - Biznis kontext a motivačný aspekt EA

Zoznam cieľov

Cieľ	Ukazovateľ cieľa	Čas dosiahnutia cieľa
Zvýšenie efektivity verejnej správy	<ul style="list-style-type: none"> ■ Náklady na integrovateľnosť dátových zdrojov agendových systémov ■ Pokles chybovosti dát 	■
Zvýšenie dôveryhodnosti a transparentnosti verejnej správy	<ul style="list-style-type: none"> ■ Percento dát o občanoch a podnikateľoch publikovaných do osobných úložísk 	■

Naplnenie cieľov OP II

Cieľ OP II	Ukazovateľ cieľa
7.2: Zvýšenie inovačnej kapacity najmä malých a stredných podnikateľov v digitálnej ekonomike	<ul style="list-style-type: none"> ■ Počet nových inovatívnych aplikácií nasadených MSP (open data, language resources, personal information management systems)
7.3: Zvýšenie kvality, štandardu a dostupnosti eGovernment služieb pre podnikateľov	<ul style="list-style-type: none"> ■ Celková spokojnosť podnikateľov so službami eGovernmentu ■ Celkové používanie služieb eGovernmentu podnikateľmi
7.4: Zvýšenie kvality, štandardu a dostupnosti eGovernment služieb pre občanov	<ul style="list-style-type: none"> ■ Celková spokojnosť občanov so službami eGovernmentu ■ Celkové používanie služieb eGovernmentu občanmi
7.7: Umožnenie modernizácie a racionalizácie verejnej správy IKT prostriedkami	<ul style="list-style-type: none"> ■ Počet nových optimalizovaných úsekov verejnej správy ■ Počet dodatočných centrálné využitých podporných systémov vnútornej správy v rámci ISVS (ako služieb v cloude SaaS)

Naplnenie cieľov OP EVS

Cieľ OP EVS	Ukazovateľ cieľa
Modernizované RLZ a zvýšené kompetencie zamestnancov	<ul style="list-style-type: none"> ■ Počet zamestnancov v analytických jednotkách v orgánoch štátnej správy na začiatku podpory. ■ Počet ústredných orgánov štátnej správy, ktoré získali podporu na zavedenie a/alebo posilnenie analytických jednotiek (cieľový stav sa očakáva 9)

3.1.2 Stakeholderi a ich záujmy

V tejto časti budú identifikovaní stakeholderi, ktorých sa daná strategická priorita (alebo jej časť) týka a ich relevantné (záujmy) hnacie prvky.

Strategická priorita „Riadenie údajov“ sa týka orgánov naprieč celou verejnou správou, všetkých agendových informačných systémov a v neposlednom rade aj verejnosti. Jej význam nespočíva len v dátach ako takých, ale v kvalite a rýchlosti rozhodnutí, ktoré kvalitné, spoľahlivé, dostupné, aktuálne, referencovateľné údaje umožňujú realizovať.

Ministerstvo financií

- Zavedenie riadenia dát ako aktív (vzácnych zdrojov) do praxe,
- Mať k dispozícii účinné nástroje, ktoré umožnia naplnenie požiadaviek eGovernmentu, princíp „1x a dost“.
- Zvýšiť efektívnosť využívania finančných prostriedkov verejnej správy (racionálna prevádzka eGovernmentu):
 - odstránením vytvárania peer to peer rozhraní,
 - odstránením vytvárania duplicitných agend,
 - odstránením manuálneho zadávania referencovateľných údajov,
 - zvýšením kvality údajov.
- Mať spoľahlivé dáta, informácie slúžiace k analýzám súčasného stavu spoločnosti a verejnej správy, k vyhodnocovaniu dopadu existujúcich opatrení, politik a regulácií, k vytváraniu prediktívnych modelov.

Inštitúcie verejnej správy

- Byť motivovaný k racionálnej správe dát,
- Mať k dispozícii kvalitné, aktuálne dáta potrebné k výkonu verejnej správy,
- Mať k dispozícii prostriedky a nástroje na účinnú správu dát, za ktoré inštitúcia verejnej správy zodpovedá,
- Znížiť administratívnu záťaž svojich zamestnancov.

Verejnosť

- Mať informácie o vlastných osobných údajoch vedených v informačných systémoch verejnej správy,
- Mať informácie o stave riešenia životných situácií,
- Mať prístup ku kvalitným službám verejnej správy, ktoré sú jednoduché, rýchle a pohodlné.

3.2 Organizácia

3.2.1 Zodpovednosť

Zodpovedný gestor

Zodpovedným gestorom za oblasť Riadenia dát je **Ministerstvo financií SR** – zodpovedá za architektúru informačných verejnej správy, eGovernment, kde dátová vrstva predstavuje kľúčový komponent.

Zapojené organizácie

Zoznam organizácií, ktoré budú dotknuté realizáciou priority.

- Ministerstvo vnútra SR – prevádzkuje časť dátového centra eGovernmentu a má v kompetencii kľúčové informačné systémy verejnej správy,
- NASES – prevádzkuje Ústredný portál verejnej správy, v rámci ktorého je potrebné sprístupniť verejnosti pohľad na ich údaje v informačných systémoch verejnej správy.
- Všetky orgány verejnej správy - Realizáciou priority riadenie dát budú dotknuté všetky orgány verejnej správy, ktoré poskytujú svoje údaje iným orgánom verejnej správy alebo verejnosti, minimálne v rozsahu zavedenia manažmentu dátovej kvality a metadata manažmentu.

3.2.2 Organizačné zmeny

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

- Vytvorenie nadrezortného tímu zodpovedného za Riadenie dát (Data governance) teda za celý životný cyklus dát, v ktorom budú minimálne zástupcovia vlastníkov dát, dátoví stewardi 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 bizns pravidiel, dátových definícií a obmedzení na dátovú integritu definovaných v dátovom modeli a ontológiách – čím sa zabezpečí vysoká kvalita a integrita dát.
- Programový manažér a projektový líder sú zodpovední 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 vlastníkmi dát za účelom zariadenia požiadaviek na dátovú kvalitu.
- Manažér 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 dát – má zodpovednosť a právomoc na rozhodovanie o prístupe k dátam, k ich distribúcii, definíciám ako aj o nastavení správnych kritérií kvality dát.
- Dátový analytik pomáha vlastníkovi dát pretaviť jeho požiadavky do dátového modelu a do procesov zberu a distribúcie dát.
- Dátový steward – riadi a stará sa o dáta ako o vzácny zdroj, **avšak dáta 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 .
- Dátový kurátor – sa venuje 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, zabraňuje poškodeniu dát, neoprávnenému prístupu k dátam.

Pre koordináciu činnosti týchto rolí je vhodné zriadiť Dátovú kanceláriu verejnej správy. Centrálna Dátová kancelária bude mať na starosti koordináciu segmentových dátových kancelárií, ktoré predstavujú dátové kancelárie na úrovni orgánov verejnej správy. Metodicky riadi dátových stewardov, dátových špecialistov, vlastníkov dát a dátových analytikov jednotlivých orgánov verejnej správy.

3.3 Stratégia a riešenie

3.3.1 Popis súčasného stavu

V súčasnosti každý orgán verejnej správy zbiera a spracováva svoje vlastné údaje. Dáta sú ukladané v rôznych IS, databázach s rôznymi požiadavkami na kvalitu, dostupnosť, bezpečnosť. Dáta, veľa krát redundantné, líšia sa formátom, štruktúrou, definíciou. Neexistuje komplexný dátový model verejnej správy, nie sú informácie, aké dáta, akým spôsobom a na aké účely sú využívané. Chýbajú interaktívne katalógy číselníkov, metadát a paradát.

Výmena potrebných dát medzi jednotlivými orgánmi verejnej správy je spôsobom „peer to peer“, teda „každý s každým“. Tento spôsob zdieľania dát je náročný na riadenie životného cyklu dát, nielen technicky, ale aj finančne, vedie k nekonzistencii a nesprávnej interpretácii dát, a v konečnom dôsledku k nesprávnym rozhodnutiam.

Prvým krokom k zmene tohto stavu je CSRÚ, ktorý zabezpečuje dátovú integráciu medzi jednotlivými orgánmi verejnej správy, poskytuje prístup k referenčným registrom a základným číselníkom, má implementované prvky dátovej kvality.

Riziká súčasného stavu

Zoznam rizík, ktoré vyplývajú zo skutočnosti, ak by sa súčasný stav nezmenil:

- Nekvalita dát, ich nejednoznačnosť, nekonzistentnosť znižuje výkonnosť verejnej správy, znižuje sa miera transparentnosti a zvyšuje sa nedôveryhodnosť verejnej správy.
- Redundancia dát, hlavne evidenčných, je potenciálnym zdrojom nekvality údajov a zvyšuje náklady na prevádzkovanie duplicitných informačných systémov.
- Požiadavky verejnosti a orgánov EU na dostupnosť údajov a zrýchlenie akcieschopnosti VS sa budú zvyšovať, náročnosť realizácie zmien v požadovanom čase a kvalite bude enormná.

3.3.2 Najlepšie skúsenosti

Zoznam najlepších skúseností (best-practice) zo zahraničia, ktoré sú relevantné pre strategickú prioritu, respektíve sa z nich vychádza:

- 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,
- Kľúčové je propagovať dôležitosť kvalitnej dokumentácie dát pomocou štandardov na metadáta na základe najlepších praktík,
- Metadáta sa musia štandardizovať naprieč celou verejnou správou,
- Vytváranie IT infraštruktúry pre webové služby a registre podľa aktuálnych medzinárodných štandardov, aby bolo možné vymieňať, opakovane používať a harmonizovať dáta aj metadáta aj v rámci EÚ,
- Zakomponovanie popisu metadát do každej etapy 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.

3.3.3 Strategický prístup k riešeniu

Realizácia „Riadenia dát“ predstavuje kľúčový krok k zefektívneniu verejnej správy. Chceme, aby sa verejná správa správala k dátam ako k vzácnemu zdroju, disponovala dátami, zdieľala ich a zabezpečila ich koordinované poskytovanie. Navrhovaná stratégia postupu je nasledovná:

- Dáta ostanú v kompetencii a vlastníctve povinnej osoby, ktorá však s nimi bude disponovať podľa jednotných pravidiel.
- Vytvorí sa centrálna kompetencia pre biznis funkcie riadenia dát, ako sú Dátová kvalita, Centrálna správa referenčných údajov, Metadata manažment, Manažment riadenia zmien, Manažment riešenia problémov. Táto kompetencia bude organizačne podporená zriadením Dátovej kancelárie verejnej správy.
- Dátová vrstva na úrovni jednotlivých povinných osôb bude spravovaná podľa centrálne nastavenej metodiky, pričom informačné systémy verejnej správy budú integrované na dátovej úrovni systémom MUK.
- Pre realizáciu riadenia dát budú nasadené SW nástroje, ktoré podporia dátovú integráciu a implementáciu jednotlivých princípov.
- **Zavedú sa najlepšie praktiky pre technické riešenia ochrany osobných údajov** ako pre riadenie prístupov k dátam a pre anonymizáciu a pseudonymizáciu vrátane identifikovania presných scenárov, kedy musia byť použité.
- Zavedie sa právny rámec pre udeleniu súhlasu použitia osobných a citlivých údajov, ako aj identifikovanie technologických riešení, ktoré tento koncept podporujú (takzvané systémy pre manažment osobných údajov).
- Vytvorí sa pravidlá riadenia dát v navzájom poprepájaných európskych privátnych aj verejných cloudoch, certifikácie podnikov pracujúcich s dátami akreditovanou certifikačnou autoritou a jednotné zásady správania sa pre podniky v dátovej ekonomike.

Pre využitie maximálneho potenciálu ukrytého v dátach je nutné aplikovať stratégiu od „Zero Data“ to „Big Data“:

- Zero Data – Čo sú kľúčové hodnoty reformy? Ako chceme merať výkonnosť a kvalitu? Čo sú naše strategické priority? Aký je kľúč k úspechu?

- **Malé dáta:** Ako sa cítia občania a podnikatelia pri využívaní služieb? Čo sú primárne zdroje zásadných pochybení? Nástroje: Metóda prípadov použitia, Analýza kľúčových príčin a podobne.
- **Stredné dáta:** Čo sú kľúčové hnacie prvky spokojnosti občanov a podnikateľov? Aká bola výkonnosť za posledné tri roky? Nástroje: Vzorkovanie, Kontrolný graf, Regresia, ANOVA a podobne.
- **Big Data:** Kto ochorie? Kto bude súčasťou fraudu? Ako využiť behaviorálnu vedu pri nastavovaní služieb? Nástroje: Clustering, Prediktívne metódy a podobne.

Riadenie dát je potrebné realizovať v dvoch paralelných vetvách:

- Upraviť súčasný stav, tak aby sa odstránila architektúra „peer to peer“. Informačné systémy verejnej správy sa modernizujú tak, aby umožnili riadenie dát a boli integrované na dátovej vrstve s modulom úradnej komunikácie.
- Pri vytváraní nových informačných systémov, aby sa zamedzilo architektúre „peer to peer“ a všetkým konsekvenciám, ktoré takéto riešenia prinášajú.

Pre existujúce agendové informačné systémy:

Prvoradou úlohou je potreba zmonitorovať aktuálny stav údajov v AIS jednotlivých orgánov verejnej správy. K tomu je potrebné stanoviť role a zodpovednosti vlastníkov a stewardov údajov. Bude zmapované, aké údaje sú evidované v jednotlivých systémoch, ktoré údaje sú primárnym vlastníctvom príslušného orgánu, ktoré údaje sú získavané od iných orgánov verejnej správy. Budú popísané dátové toky.

Ďalšou úlohou je definovať rozsah a priority dátovej integrácie existujúcich IS. Postupné odstraňovanie „peer to peer“ prepojení, nahradením jednotného prístupu cez MUK. Táto aktivita bude spočiatku finančne aj časovo náročná, nakoľko si vyžiada úpravu na strane poskytovateľa aj konzumenta dát. Postupným využívaním štandardizovaného prístupu k zdieľaným údajom bude náročnosť klesať.

Pre novovznikajúce informačné systémy:

- Vznikajúce IS majú svojho vlastníka a data stewardov.
- Sú konzumentmi dát iných IS a poskytovateľmi dát pre iné IS. Vzájomná komunikácia medzi IS bude prebiehať cez MUK. Pre dáta, ktoré nový informačný systém poskytuje, budú implementované procesy dátovej kvality a riadenia životného cyklu.
- Zavedením a dodržiavaním všetkých častí Data Governance budú dodržané dátové architektonické princípy.

Princípy

Zoznam architektonických princíпов, na ktorých je riešenie postavené:

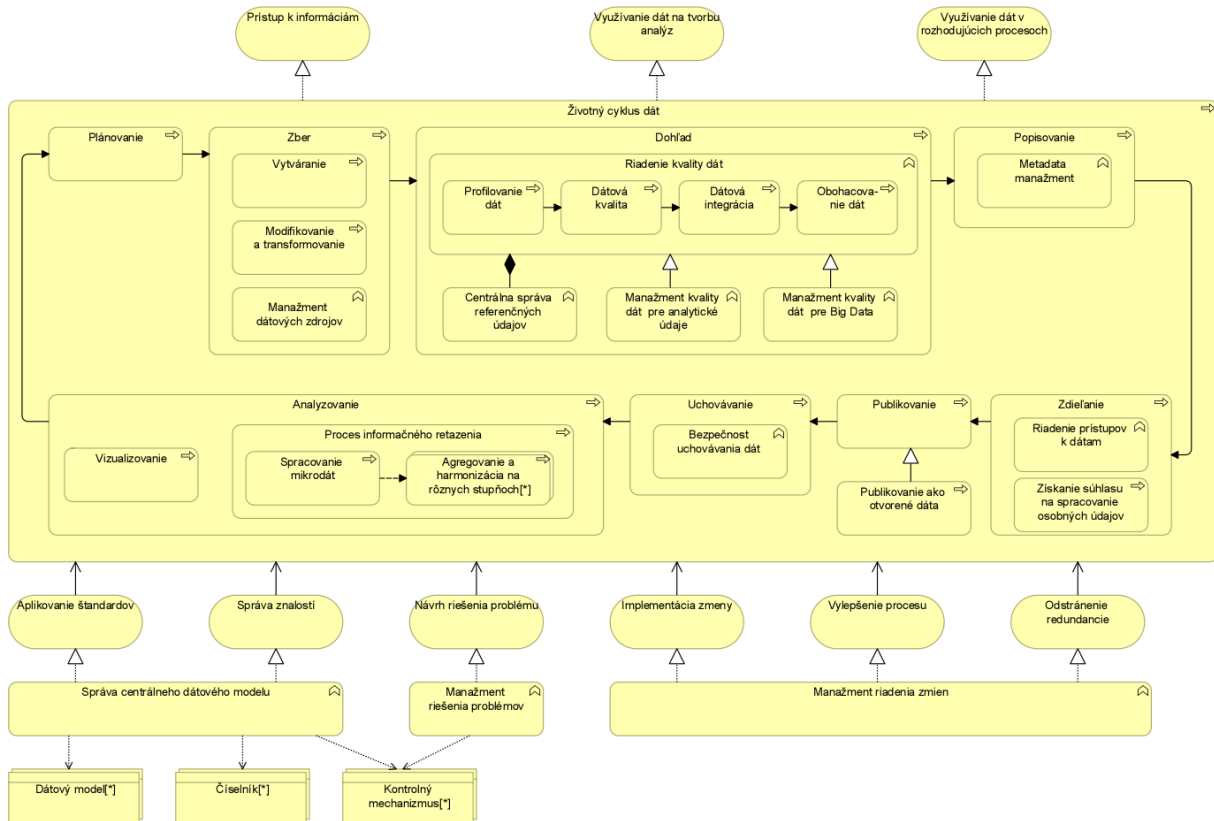
- **Údaje sú aktívom**, ktoré má hodnotu a je podľa toho riadené a spravované. Údaje sú hodnotným zdrojom VS, majú reálnu merateľnú hodnotu. V zjednodušenom ponímaní, účelom údajov je pomáhať pri rozhodovaní, podporovať rozhodovanie. Presné a včasné údaje sú rozhodujúce pre presné a včasné rozhodnutia.
- **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 a použití. Správca údajov zodpovedá za ich dôvernosť, integritu a dostupnosť.
- **Spoločné údajové ontológie a definícia konceptov** - 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.

- Aplikovanie princípu *súkromie ako súčasť dizajnu riešenia* („Privacy by design“).

3.3.4 Biznis vrstva

Obrázok 4: Pohľad na biznis vrstvu v oblasti Riadenia dát



Požiadavky

Kategorizácia typov dátových zdrojov

- V budúcom období musia byť definované presné metodológie a procesy pre riadenie dát
 - podľa zdroja:
 - administratívne zdroje, produkujúce administratívne dáta v procese výkonu verejnej moci;
 - dátové laked a dátové sklady obsahujúce analytické dáta pre sekundárne využívanie v oblasti analýz (ide o dáta zozbierané z ostatných dátových zdrojov);
 - štatistické prieskumy, určené pre štatistické zisťovania;
 - senzory;
 - internet of things;
 - industry 4.0;
 - podľa typu:
 - interné;
 - externé;
 - podľa formy:
 - štruktúrované;

- neštruktúrované;
- prípadne ďalšie kategorizácie:
 - administratívne údaje;
 - analytické údaje;
 - surové údaje a Big Data;
 - geografické údaje;
 - osobné údaje;
 - citlivé údaje a podobne;

Kategorizácia metadát

V kontextu datových skladů se rozlišují tři hlavní kategorie metadat – business metadata, technická metadata a operační/provozní metadata. Každá z těchto kategorií je využívána jinými typy uživatelů a je zpracována pomocí jiných nástrojů.

Business metadáta popisujú BI prostredie, použité objekty a výstupy (reporty, dashboardy, predikcie) pomocou jednotných a previazaných business pojmov. Tieto informácie o dátach sú potrebné pre ich výmenu v budúcnosti, prehľadávanie (ako aj vôbec možnosť dozvedieť sa, že také dáta existujú) a sprístupňovanie tak, aby dátam používatelia porozumeli. Primárni používatelia dát totiž môžu disponovať nezdokumentovanou vedomosťou, ktorá ich prevedie cez správnu analýzu, avšak sekundárni používatelia sa musia spoľahnúť len na formalizované metadáta, aby mohli využiť plný potenciál dát a neurobiť chyby pri svojich analytických záveroch. Do kategórie business metadát tiež patrí ďalšie vysvetlenie princípov, metodík a pravidiel aplikovaných v prostredí BI.

Technickými metadátami sa chápe napríklad popis štruktúry dát (modely, tabuľky, stĺpce, vzťahy), životného cyklu dát (data lineage - od vzniku až po prezentáciu užívateľovi) a ďalšej dokumentácie riešenia (metodológiu, vzorkovanie, ukazovatele kvality, ...). Vďaka týmto informáciám je možné ľahšie vykonávať analýzu dopadov zmien, riadenie kvality dát a podobne.

Procesné/operačné metadáta sú prevádzkového charakteru, spadajú sem logy, informácie o zpracovaní automatizovaných úloh, výsledky monitoringu (využitie CPU, veľkosť DB) a pod. Tieto informácie sú využívané pre prevádzku a každodennú údržbu BI prostredia.

Definovanie kritérií kvality dát a jej meranie

Kvalita dát má priamy dopad na spoľahlivosť navrhovaných politík a regulácií ako aj na platnosť vedeckého výskumu. Väčšinou sa meria využitím kritérií ako:

- Dostupnosť;
- Presnosť;
- Úplnosť;
- Relevantnosť;
- Včasnosť;
- Integrita;
- **Relevantná prepojitelnosť (možnosť kombinovať dáta z rôznych dátových zdrojov) a porovnateľnosť rôznych datasetov;**
- Konzistentnosť a mnohé ďalšie, ktoré musia byť presne zadefinované.

Splniť tieto kritériá neznamená len dáta sprístupniť, ale aj zrozumiteľne zdokumentovať dátové štruktúry, proces tvorby dát, štatistické metodológie (ak boli použité), dátové zdroje, kontext a mnoho ďalších aspektov.

Kvalita dát má aj mäkké aspekty ako je prehľadnosť a zrozumiteľnosť ich prezentácie, preškolenie zamestnancov, metodické pokyny a podobne.

Spravovanie dát verejnosťou

- Verejnosť môže pristupovať k údajom, ktoré o nej verejná správa eviduje:
 - K dispozícii je prehľad všetkých dátových objektov, ktoré sa týkajú klienta na jednom mieste,
 - 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 (otvorené dáta – neanonymné, otvorené dáta – anonymizované, neanonymizované, pseudonymizované alebo anonymizované iba pre potreby verejnej správy, iba pre potreby vybranej inštitúcie a tak ďalej).
- Klient (FO a PO) poskytuje nové údaje verejnej správe len raz, nie každej inštitúcii samostatne.

Výmena údajov medzi orgánmi verejnej moci

- Údaje o klientovi, ktoré vlastní a spravuje iný orgán verejnej správy, si inštitúcie poskytujú navzájom bez potreby interakcie s klientom (FO/ PO):
 - Pre každý proces bude navrhnutý biznis kontext: zoznam dátových objektov (z iných inštitúcií verejnej správy), ku ktorým je možné napriamo pristupovať – na základe tohto kontextu budú nastavené oprávnenia pre prístup k údajom,
 - Biznis kontext bude centrálné spravovaný,
 - Klient (FO/PO) bude môcť mať výmenu informácií pod kontrolou (bude informovaný a vo vybraných situáciách bude dávať súhlas na zdieľanie údajov).
- Automatizácia komunikácie medzi jednotlivými AIS, odstránenie manuálnych činností pracovníkov verejnej správy (prepísovanie údajov z jednej aplikácie do druhej), ktoré môžu byť zautomatizované.

Dátový model verejnej správy

- Je dostupný dátový model, je známy stav dát v rámci ich životného cyklu.
- Údaje majú jednoznačné a zrozumiteľné definície, majú vlastníka a správcu. Sú dostupné, zdieľané, navzájom referencovateľné.
- S ohľadom na Big Data a ukladanie neštruktúrovaných dát v takzvaných Data Lakes je však dôležité vedieť pracovať aj s ad hoc dátovými modelmi, čo je možné hlavne vďaka kvalitnému a dôslednému Metadáta manažmentu.

Nový prístup k ochrane osobných údajov a cezhraničnej výmeny

Všeobecné nariadenie COM(2012)11 o ochrane údajov posilní dôveru v digitálne služby, nakoľko by malo chrániť jednotlivcov v súvislosti so spracovaním osobných údajov všetkými spoločnosťami, ktoré ponúkajú svoje služby na európskom trhu. 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 verejnemu blahu. Občania postupne získajú právo kontrolovať tok svojich dát, ktoré vlastní, a udeľovať súhlas na ich spracovanie – prvou oblasťou pre túto zmenu môže byť zdravotníctvo alebo štatistické zisťovanie. Do prostredia eGovernmentu sa nasadí pilotné riešenie Systému na správu osobných údajov (Personal Information Management System), ktoré naštartuje akceptáciu trhom a spotrebiteľmi a podporí tak vznik ďalších komerčných riešení a aplikácií.

Služby

Zoznam služieb

- Kontrola kvality dát,
- Vyhodnotenie kvality dát,
- Publikovanie a zdieľanie údajov.

Procesy a funkcie

- **Správa centrálneho dátového modelu**, v rámci ktorej vznikne tiež kompetencia na medzirezortné porozumenie dát ako aj na presadzovanie štandardov na pomenovávanie dát a na obsah jednotlivých polí (číselníky, prípustné hodnoty, kontrolné mechanizmy).
- **Životný cyklus dát – publikovanie ako otvorené dáta:**
 - Plánovanie
 - Zber:
 - Identifikovanie dátových zdrojov,
 - Poverenie vlastníkov, ktorí zbierajú a aktualizujú dáta,
 - Zavedie validačných pravidiel pre dáta a biznis pravidiel pre vytvorenie, zmenu alebo zber údajov,
 - Dohľad
 - Popisovanie
 - Zdieľanie
 - Definovanie pravidiel bezpečnosti prístupu a ich dodržiavanie,
 - Vytvorenie štandardného dopytu a reportov,
 - Poskytovanie funkcionality používateľom,
 - Manažovanie využívania systému,
 - Monitorovanie výstupnej dátovej kvality,
 - Poskytovanie vhodných metadát,
 - Publikovanie
 - Zaevidovanie do katalógov ako Katalóg otvorených dát a do registrov
 - Nastavenie procesov pre vytváranie aplikačných rozhraní (API)
 - Uchovávanie
 - Vytvorenie a dodržiavanie pravidiel uchovávania,
 - Vymazávanie dát v súlade s biznis pravidlami.
 - Analyzovanie
- **Manažment kvality dát** - 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:
 - Profílovanie dát – ide o proces, v ktorom sa získa porozumenie toho, ako existujúce dáta korešpondujú s nastavenými kritériami kvality. Kontroluje sa predovšetkým úplnosť dát – jednak, či sú k dispozícii údaje o každom objekte a jednak, či je vyplnené každé pole. Ďalším rozmerom je presnosť dát v poliach - či vyplnené informácie majú povolené rozsahy hodnôt, správny formát, či sú štandardizované, v súlade s číselníkmi a či úspešne prešli automatickou kontrolou. Podľa nastavených kritérií kvality sa môže ďalej hodnotiť integrita, včasnosť, konzistentnosť a mnohé ďalšie parametre.
 - Dátová kvalita – na základe poznatkov z profílovania 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 riešenia 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:

- Vylúčiť dáta, hlavne pri ich kopírovaní do dátových skladov, ak sú problémy príliš vážne a neopraviteľné;
 - Prijíť 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“.
- Dátová integrácia a postupné zjednocovanie dátovej vrstvy verejnej správy – dáta o tých istých objektoch veľmi často existujú vo viacerých databázach. Prvou výzvou je rozpoznať, že sa jedná o ten istý objekt, druhá výzva je informácie o objekte z rôznych databáz zjednotiť tak, aby vznikol správne skonsolidovaný obraz. Cieľom celého riešenia je teda poskytovať užívateľovi konsolidovanú informáciu cez všetky Agendové (prípadne aj iné) systémy, v ktorých sú uložené jeho referenčné alebo nereferenčné dáta. Pretože sa predpokladá, že jednotlivé agendové systémy sa budú prekrývať v uložených informáciách o hlavných entitách (napr. klient), je potrebné tieto entity konsolidovať a vytvoriť tzv. "Golden record", podľa ktorého sa budú ostatné komponenty riadiť.
- Obohacovanie dát je poslednou fázou, ako dať dátam vyššiu kvalitu. Môže ísť o obohacovanie dát o danom objekte 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 objektu, na základe ktorých sa dá s istotou učiť, že ide o informácie o tom istom objekte. Takémuto kombinovaniu dát extrémne napomáha aj riadenie metadát, ako je popísané nižšie.

Kontrola dátovej kvality údajov prebieha v niekoľkých úrovniach:

- segmentovej: pre každý agendový systém sú implementované procesy pre dátovú kvalitu. Túto službu unifikovane poskytuje CSRÚ. Realizuje meranie a vyhodnocovanie KPI (ukazovatele výkonnosti), a dátový steward aplikuje proces opravy údajov, ktorých správa je v jeho kompetencii
- centrálnej: kontrola údajov naprieč súvisiacimi publikovanými datasetmi, (predpokladom je, že agendové informačné systémy publikujú a zdieľajú kvalitné, preverené údaje). Dátoví stewardi dotknutých datasetov spoločne riešia problém nekvality údajov.
- analytickej : implementácia procesov dátovej kvality v CDS a reportoch.

■ **Manažment kvality dát** pre analytické údaje;

■ **Manažment kvality dát** pre Big Data (zo senzorov, internet of things, industry 4.0);

■ **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)³, ISO-TS 17369 Statistical Data and Metadata Exchange (SDMX)⁴, Dublin Core Metadata Initiative (DCMI)⁵, ISO/IEC 11179⁶. Jeden štandard pre medatáda však nepokryje celý životný cyklus dát. Dôležité je vybudovať **distribovaný 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 poloaufomaticky, často je tiež potrebné zasiahnuť manuálne a dohľadať chýbajúci kúsok znalosti o dátach.

■ **Master data management a centrálna 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. Referenčné údaje

³ <http://www.ddialliance.org>

⁴ <http://sdmx.org/>

⁵ <http://dublincore.org/>

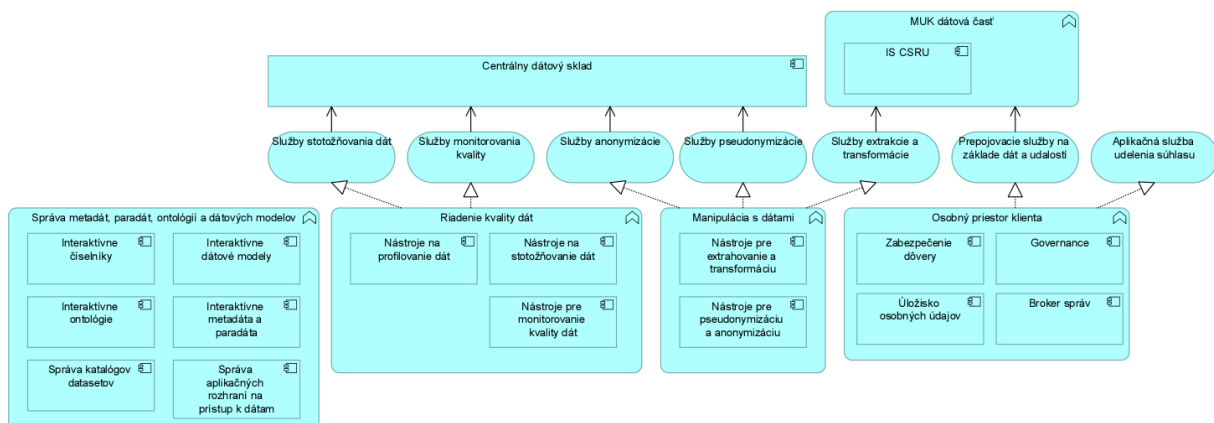
⁶ <http://metadata-standards.org/>

(i Master data) sú údaje netraskčného charakteru, popisujúce entity, vzťahy medzi nimi a ich hierarchie. Tieto údaje sú súčasťou mnohých IS, sú zbierané a spracovávané v rôznych procesoch.

- **Manažment riešenia problémov** – nastavenie procesov, pravidiel a roli umožní urýchliť vyriešenie incidentov.
- **Manažment riadenia zmien** – realizáciou procesov riadenia zmien bude zabezpečené minimalizovanie implementácie biznis procesov zbierajúcich duplicitné dáta a spravovanie redundancie dát.
- **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.
- **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.
- **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).

3.3.5 Aplikačná vrstva

Obrázok 5: Pohľad na aplikačnú vrstvu pre Riadenie dát



Požiadavky

Zoznam požiadaviek

- Každý informačný systém verejnej správy bude pripojený na platformu dátovej integrácie tak, aby údaje, ktoré sú vo vlastníctve iného orgánu verejnej moci alebo sú spravované v inom informačnom systéme, mohli byť prepoužité. Ak potrebný údaj existuje v niektorom informačnom systéme verejnej správy, informačné systémy si nevytvárajú vlastné, editovateľné evidencie.
- Ak dátový prvok môže nadobúdať iba povolené hodnoty, informačné systémy využívajú číselníky, vlastné alebo zdieľané, aby sa eliminovala nekvalita údajov, spôsobená typovaním do systému. Na aplikačnej úrovni sú kontrolované aj formáty jednotlivých dátových prvkov.
- Každý informačný systém má popísané svoje metadáta v interaktívnom katalógu.
- Je možné spravovať dlhodobé datasety v jednom formáte dát.

Aplikačné služby

Zoznam služieb

- Prístup k publikovaným dátam (iný informačný systém, interný alebo externý používateľ),

- Notifikačné služby, kedy sa používatelia alebo systémy môžu automaticky dozvedieť o aktualizovaných alebo nových dátach,
- Poskytovanie tzv. vylepšených publikácií, ktoré poskytujú zistenia z dát, samotné dáta a metadáta v jednom balíčku, čím sa podporuje koncept opakovateľných a overiteľných dátových analýz. Uchováva sa tiež linka medzi primárnymi a sekundárnymi dátami.
- Poskytovanie interaktívnych číselníkov,
- Poskytovanie dátových katalógov,
- Vytváranie databázy otázok,
- Vytvorenie mapovaní na staré verzie číselníkov a metadát,

Aplikačné komponenty

Zoznam komponentov:

- MUK dátové časť:
 - Je nevyhnutným integračným prvkom. Umožňuje automatizovanú komunikáciu a zdieľanie údajov medzi jednotlivými informačnými systémami.
- Riadenie kvality dát:
 - Podporuje všetky procesy riadenia dátovej kvality: plánovanie, nasadenie, monitoring, tvorbu opatrení.
 - Poskytuje nástroje na profilovanie dát, stotožňovanie dát, monitoring a reporting kvality dát.
 - Umožňuje definovať a modifikovať predpisy a realizuje biznis pravidlá dátovej kvality na základe zvolených kľúčových dátovej ukazovateľov kvality.
- Správa metadát, paradát, ontológií a dátových modelov:
 - Úlohou tejto komponenty je udržiavať metadáta jednotlivých agendových informačných systémov, ich dátové modely, definície dát, dátových tokov.
 - Umožňuje vytvárať a spravovať ontológie a dátové modely.
 - Umožňuje dopĺňať, modifikovať, prehliadať metadáta informačných systémov, dátové modely a ontológie.
- IS CSRÚ:
 - Bude v nej prebiehať konsolidácia hlavných entít. Stane sa tak jedným zdrojom pravdy pre ostatné informačné systémy.
 - Vystavuje svoje on-line, off-line rozhrania. Ostatné komponenty sa budú na MDM komponentu integrovať cez dátovú platformu.
 - Jednotlivé agendové informačné systémy sú napojené na IS CSRÚ, kde sú konsolidované hlavné entity (napr. Klient, Produkt).
 - IS CSRÚ spracováva dáta transakčného charakteru.
 - Front end komponenty (napr. Portál UPVS) pristupujú k IS CSRÚ cez Integrovanú platformu (ESB), aby bolo možné v prípade nutnosti orchestrovať otázky do Agendových informačných systémov, poprípade iných komponentov
- Manažment riešenia problémov
 - Implementované nástroje umožnia evidovať a efektívne riešiť problémy súvisiace s kvalitou a informačnou hodnotou údajov.
- Manažment riadenia zmien
 - Systém, ktorý umožní evidovať, prioritizovať požiadavky na zmeny a riadiť zmeny v informačných systémoch.
- Správa aplikačných rozhraní na prístup k dátam
 - V rôznych fázach životného cyklu dát je umožnený rôzny prístup – MUK pre výmenu počas transakcií výkonu verejnej moci, Open Data pre publikovanie otvorených dát

- Správa katalógov datasetov:
 - Poskytuje nástroje pre vyhľadávanie a lokalizovanie dát a dátových zdrojov.

3.3.6 Technologická vrstva

Pre nové a aktualizované systémy bude vo vládnom cloude dostupná moderná dátová integrácia vo forme iPaaS ako súčasť MUK.

3.4 Problémy a riziká

Dôsledky

Zoznam dôsledkov, ktoré vyplývajú z realizácie strategickej priority pre jednotlivé subjekty:

- Všetky orgány štátnej moci budú pristupovať k dátam iných orgánov cez Modul úradnej komunikácie. Existujúce informačné systémy budú postupne pretransformované tak, aby boli schopné využívať integráciu cez Modul úradnej komunikácie.
- Niektoré informačné systémy stratia svoju opodstatnenosť.
- Naplnia sa požiadavky eGovernmentu „1x a dost“

Problémy

Zoznam problémov, ktoré bude potrebné vyriešiť počas realizácie strategickej priority:

- Ako vyriešiť efektívnu medzirezortnú spoluprácu, keďže program riadenia dát a ich kvality **je horizontálneho charakteru** a týka sa celej verejnej správy.
- Ako presvedčiť jednotlivé organizácie, že majú problém vo svojom riadení dát.
- Ako zabezpečiť udržateľné finančné a ľudské zdroje, keďže riadenie dát a ich kvality sa vo všeobecnosti považuje za veľmi náročnú úlohu vyžadujúcu veľa zdrojov, avšak návratnosť investícií je často veľmi ťažké objektívne kvantifikovať.
- Ako prioritizovať zdokumentovanie dát, dátových tokov v existujúcich IS.
- Ako nájsť a poveriť zodpovednosťami za konkrétne oblasti dátových stewardov.
- Ako realizovať opatrenia pre zvýšenie kvality dát a ako integrovať dáta ze zdrojů z rôznou východiskovou kvalitou dát.
- Ako integrovať informačné systémy, ktoré sú už morálne zastarané.

Riziká

Zoznam rizík, ktoré je dôležité mitigovať pri realizácii strategickej priority:

- Neochota jednotlivých orgánov verejnej správy zdieľať svoje údaje.
- Neochota jednotlivých orgánov zbavovať sa vlastných agend a informačných systémov, ktoré sú z pohľadu celku redundantné
- Nepodarí sa zaviesť procesy Data Governance do praxe, t.j. nepodarí sa využiť potenciál zdieľania kvalitných dát.
- Bude sa zvyšovať potreba na rozšírenia kapacity dátových centier, v dôsledku vytvárania redundantných agend a ukladania ich dát.

3.5 Legislatívne požiadavky

Legislatívne požiadavky

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

Legislatívny predpis	Navrhované opatrenie
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 eGovernmente, - aby na výmenu údajov medzi jednotlivými AIS musel byť použitý MUK. - aby každý zdieľaný dataset prešiel dátovou kontrolou
	Zriadenie centrálnej dátovej kancelárie
	Definovanie nových rolí a zodpovedností v oblasti Riadenia dát
Všeobecné nariadenie COM(2012)11 o ochrane údajov	Zohľadniť pri technických realizáciách

Požiadavky na Operačný program Efektívna verejná správa

- Vytvoriť katalóg dát (pojmov a definícií), dátových modelov existujúcich agendových informačných systémov
- Vytvoriť mapu dátových tokov medzi jednotlivými agendovými informačnými systémami.

3.6 Plánovanie a migrácia

Identifikácia projektov, ktoré je vhodné realizovať:

- Vybudovanie platformy dátovej integrácie – Rozvoj MUK 2,
- Integrácie všetkých informačných systémov verejnej správy do platformy dátovej integrácie
 - Bude súčasťou aktivít každého projektu, ktorý bude modernizovať alebo zavádzať príslušné ISVS.

3.6.1 Vybudovanie platformy dátovej integrácie – Rozvoj MUK 2

Popis a cieľ

Cieľom projektu Rozvoj MUK je v prvom rade zvýšenie efektivity vo verejnej správe automatizovanou výmenou dát, odstránením peer to peer integrácií a dobudovaním platformy dátovej integrácie, ktorá umožní výmenu dát, informácií medzi jednotlivými informačnými systémami jednotným spôsobom. Konzument dát sa integruje na modul úradnej komunikácie len raz, je odtienený od technológii a ich zmien na strane poskytovateľa dát napr. Pri výmene príslušného agendového informačného systému. Bude zabezpečená kontrola poskytovaných údajov, tak aby konzument dát dostal dôveryhodné údaje. Dátová platforma umožní nielen online komunikáciu, ale aj dávkové spracovanie údajov medzi jednotlivými informačnými systémami.

Aktivity

- Analýza a design
- Výber vhodných platforiem iPaaS do vládneho cloudu
- Implementácia pilotných projektov
- Implementácia celkového riešenia
- Testovanie
- Nasadenie

Navrhovaný realizátor

- Gestor: Ministerstvo financií SR,

Navrhovaný partneri

- Inštitúcie verejnej správy

Odhadovaná dĺžka realizácie projektu (projektov)

- 3 roky

Závislosti na iných projektoch a prioritách (OPII a OPIS)

- MUK
- Vládny cloud

Závislosti na iných projektoch a prioritách (OPEVS)

- Vytvorenie dátovej kancelárie VS

3.6.2 Integrácie všetkých informačných systémov verejnej správy do platformy dátovej integrácie

Popis a cieľ

Cieľom integračných aktivít je postupné odstránenie integrácií typu peer to peer a zabránenie ich vznikaníu. Integrácia všetkých informačných systémov verejnej správy do platformy dátovej integrácie bude riešená v jednotlivých projektoch, ktoré budú zavádzať alebo modernizovať existujúce informačné systémy. Ak nový informačný systém bude konzumentom dát informačného systému, ktorý ešte nie je zintegrovaný cez dátovú platformu, integrácia aj takéhoto poskytovateľa dát musí byť súčasťou projektu. Taktiež, ak bude potrebné rozšírenie funkcionality dátovej platformy, budú tieto aktivity zahrnuté v projekte.