

Souhrnná zpráva z provedených rozhovorů

Podprodukt č. I.2

Vytvořeno v rámci projektu Zajištění podmínek pro kvalitní správu datového fondu a zajištění řízeného přístupu k datům Reg.č.: CZ.31.1.01/MV/23_62/0000062

Obsah

1	Úvod	3
2	Souhrn zjištěných poznatků	4
2.1	Definice oblastí správy dat veřejné správy	4
2.2	Stručná charakteristika jednotlivých oblastí:	5
2.3	Základní zjištění v jednotlivých oblastech na základě strukturovaných rozhovorů	5
2.3.1	Oblast řízení a organizace správy dat	5
2.3.2	Oblast stanovení datových potřeb	5
2.3.3	Oblast popisu dat a datových řešení	6
2.3.4	Oblast realizace datových řešení	6
2.3.5	Oblast zajištění a využívání dat	6
2.4	Doplnění pohledu na základě postřehů z rozhovorů mimo strukturu otázek	7
2.4.1	Obecné postřehy	7
2.4.2	Rozmanitost „informatické struktury“ úřadů a jejich agendových působení	7
2.4.3	Souhrnné společné poznatky	8
3	Současný stav – přehled zjištěných informací	8
3.1	Oblast řízení a organizace správy dat	9
3.1.1	Organizace	9
3.1.2	Standardizace	11
3.1.3	Rozvoj kompetencí, spolupráce a podpora	13
3.1.4	Základní zjištění v oblasti řízení a organizace správy dat:	13
3.2	Stanovení datových potřeb	14
3.2.1	Prioritizace oblastí	14
3.2.2	Řízení datových potřeb	15
3.2.3	Využití architektury pro plánování změn	16
3.2.4	Základní zjištění v oblasti stanovení datových potřeb:	17
3.3	Popis dat a datových řešení	17
3.3.1	Popis dat	17
3.3.2	Modelování dat a datových řešení	18
3.3.3	Evidence v RPP	19
3.3.4	Kategorizace dat	22
3.3.5	Základní zjištění v oblasti popisu dat a datových řešení:	23
3.4	Realizace datových řešení	24
3.4.1	Aspekty řešení změn IS	24
3.4.2	Datová integrace	25
3.4.3	Základní zjištění v oblasti realizace datových řešení:	27

3.5	Zajištění a využívání dat	27
3.5.1	Využití dat	27
3.5.2	Identifikované nástroje využívané pro přípravu a analýzu dat v úřadech:.....	29
3.5.3	Kvalita dat	29
3.5.4	Bezpečnost / řízení rizik.....	29
3.5.5	Provoz datových úložišť a řešení	30
3.5.6	Základní zjištění v oblasti realizace datových řešení:	31
4	Metodika rozhovorů.....	31
4.1	Stanovení struktury vedených rozhovorů.....	31
4.2	Šablona zápisu výsledků rozhorů	33
4.3	Dotazované organizace.....	33
5	Přílohy	34
5.1	Příloha č. 1.....	35

1 Úvod

V gesci Digitální a informační agentury (DIA) a pod vedením odboru Hlavního architekta eGovernmentu (OHA) probíhají aktivity směřující k rozvoji kvalitní správy dat (data managementu) ve veřejné správě (VS).

Všechny aktivity jsou realizovány v rámci projektu „Zajištění podmínek pro kvalitní správu datového fondu a zajištění řízeného přístupu k datům“, jehož předmětem je:

- provedení datových auditů v organizacích státní správy ČR,
- vypracování „Strategie pro správu dat ve veřejné správě České republiky pro období (2024–2030)“,
- příprava paragrafového znění zákona o správě dat,
- metodická a konzultační podpora organizací státní správy při implementaci „Standardu správy dat“,
- příprava podmínek pro implementaci druhé kapitoly [nařízení o evropské správě dat \(tzv. "Data Governance Act"\)](#), tedy stanovení pravidel pro sdílení údajů, které podléhají některému způsobu ochrany, a není možné je zveřejňovat jako otevřená data, ale jejich zpřístupnění je možné za splnění určitých podmínek,
- rozvoj nástrojů pro vytváření slovníků pojmů a vytváření sémantických datových modelů.

Jedním z prvních cílů projektu je vytvoření strategického dokumentu „Strategie pro správu dat ve veřejné správě České republiky pro období (2024–2030)“, který bude v závěru roku 2023 předložen vládě ke schválení. Schválená strategie se spolu s výsledky provedených auditů stane základem pro následnou dlouhodobou metodickou podporu organizací VS, a také pro přípravu legislativních změn, které pevně zakotví povinnosti spojené s řádnou správou dat ve VS.

Významnou součástí přípravy strategického i legislativního rámce v letošním roce byly i realizované rozhovory k zmapování aktuálního stavu data managementu ve 34 organizacích státní správy ČR.

Cílem realizovaných rozhovorů bylo:

- systematicky získat informace o aktuálním stavu správy dat jednotlivých orgánů VS (úřadů) s cílem následujícího využití:
 - rozhovory a agregované získané informace se stanou základem pro finální upřesnění „standardu správy dat“ a pro přípravu „Strategie pro správu dat ve veřejné správě České republiky pro období (2024–2030),
 - rozhovory a získané informace se současně stanou „zpětnou vazbou“ pro vedení úřadu o stavu správy dat v jejich úřadě a prioritních oblastech, základních nedostatcích, vyžadujících zlepšení.
- přispět ke zvýšení povědomí klíčových aktérů správy dat v organizacích o základních principech, oblastech, rolích a procesech správy dat (datový audit bude pro řadu organizací prvním setkáním se systematickou správou dat),
- připravit podmínky pro přípravu navazujících aktivit zaměřených na konzultační a metodickou podporu kvalitní správy dat v organizacích, která bude poskytována v dalších fázích projektu,
- zmapovat stávající stav každé organizace prizmatem „standardu správy dat“, aby se stal výchozí základnou pro navržení optimálních (připravených na míru organizaci) navazujících aktivit k dosažení standardu,
- definovat a vymežit věcnou náplň dlouhodobé konzultační a metodické podpory pro zavádění kvalitní správy dat v organizacích ze strany OHA.

2 Souhrn zjištěných poznatků

Rozhovory byly směřovány ke zjištění aktuálního stavu připravenosti veřejné správy k naplnění požadavků digitalizace jejích služeb poskytovaných veřejnosti (a nejen jí).

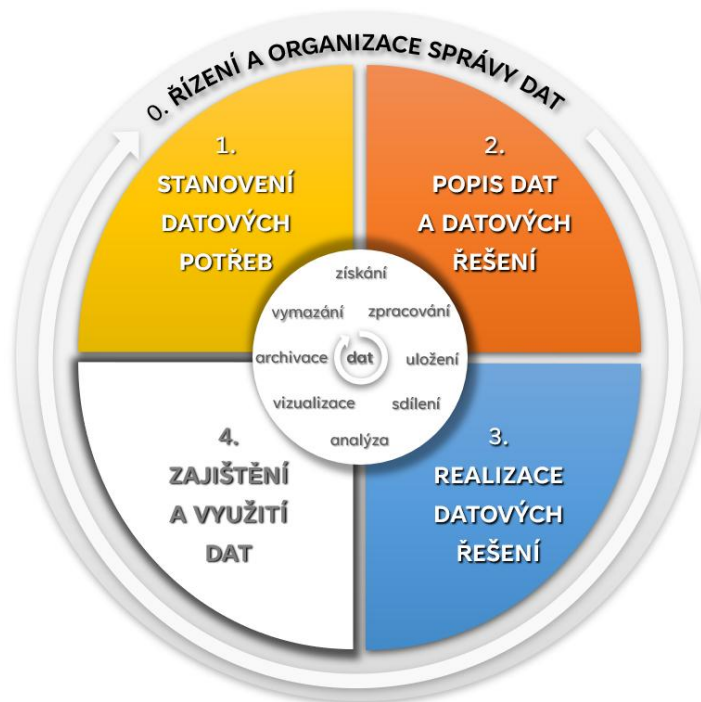
2.1 Definice oblastí správy dat veřejné správy

Realizaci rozhovorů s organizacemi státní správy předcházela příprava, která byla zaměřena na výběr a specifikaci relevantních oblastí důležitých pro dobrou správu dat ve veřejné správě. K tomu byly využity 2 zdroje informací:

- světově nejrozšířenější metodický rámec pro data management DAMA/DMBoK („Data Management Body of Knowledge“),
- znalosti a zkušenosti dobré praxe data managementu získané v průběhu šesti společných workshopů od členů expertní skupiny, sestavené pro tento účel z osob působících v oblasti komerčních společností a akademické sféry.

Výsledkem společné přípravy byla specifikace pěti doporučených oblastí správy dat pro veřejnou správu ČR. Jedna oblast je zaměřena na řízení a organizaci správy dat a v zásadě odpovídá centrální oblasti Data Governance dle DMBoK, zbylé čtyři volně kopírují průběh životního cyklu dat a datových řešení, z nichž oblast „Zajištění a využití dat“ pokrývá reálnou každodenní správu dat a jejich interní i externí využívání v organizacích VS.

Obrázek č. 1: Schéma pěti doporučených oblastí správy dat



Zdroj: vlastní DIA

2.2 Stručná charakteristika jednotlivých oblastí:

0. Řízení a organizace správy dat

- Organizační, procesní a kompetenční zajištění správného chování, rozhodování a kontroly v oblasti péče o data tak, aby byla chráněna jejich důvěryhodnost a hodnota jakožto cenného zdroje.

1. Stanovení datových potřeb

- Stanovené relevantní oblasti dat v organizaci, jejich prioritizace, identifikace příležitostí a potřeb v těchto oblastech, definice a prioritizace požadavků na datová řešení a vytvoření plánu aktivit směřujících k jejich naplnění.

2. Popis dat a datových řešení

- Evidence dat, kategorizace, popis, datové modely, modely datových toků a IT komponent zapojených do řešení datových potřeb (primárních systémů, datových úložišť, datových sad a rozhraní, nástrojů pro ETL, analýzu a prezentaci dat).

3. Realizace datových řešení

- Způsoby implementace nových a rozvoj stávajících datových řešení (získávání/vytváření, ukládání, zpracování, zpřístupňování, sdílení dat).

4. Zajištění a využití dat

- Provozování datových řešení (reálné získávání/vytváření, ukládání, zpracování a zpřístupňování konkrétních dat), zajištění jejich dostupnosti, bezpečnosti, kvality a interoperability. Tvorba datových produktů (reportů, analýz, odpovědí na dotazy apod.) pro podporu informovaného rozhodování ve VS („evidence based decision making“), zpřístupnění dat veřejnosti a podpora řízeného přístupu k datům VS odborné veřejnosti.

2.3 Základní zjištění v jednotlivých oblastech na základě strukturovaných rozhovorů

Základem zjišťování informací byly rozhovory vedené se 34 úřady státní správy v definovaných pěti oblastech zaměřených na dobrou správu dat. Způsob provedení rozhovorů, detailní získané poznatky a způsob jejich dalšího zpracování je uveden v kapitole „Současný stav – stručný přehled zjištěných informací“. Zastřešující zjištění reprezentující výsledky rozhovorů jsou uvedena v další části této kapitoly.

2.3.1 Oblast řízení a organizace správy dat

Základní charakteristiky stávajícího stavu oblasti:

- absence jednoznačně stanovené manažerské odpovědnosti za data celého úřadu a jejich správu,
- nejednoznačné společné vnímání pojmu „agenda“,
- absence datových rolí a jejich formalizace,
- nedostatečná interní legislativa zaměřená na data,
- neexistence publikovaných standardů VS v oblasti dat,
- podceněný rozvoj kompetencí (vzdělávání, spolupráce, podpora) v oblasti dat, ale také v celé oblasti ICT.

2.3.2 Oblast stanovení datových potřeb

Základní charakteristiky stávajícího stavu oblasti:

- nesoulad mezi vymezením věcných datových oblastí a stanovením věcných odpovědností za jejich data,
- nesystémová prioritizace věcných datových oblastí bez navazujících dopadů do správy dat,
- systematické nevyužívání EA při správě změnových požadavků a neexistence současného i budoucího stavu v oblasti dat v EA.

2.3.3 Oblast popisu dat a datových řešení

Oblast přehledu o datech organizací a jejich popisech spadá mezi nejvíce podhodnocené oblasti správy dat. Organizace neudržují přehled o svých datech, jejich významech, metadatových popisech, zdrojích dat, neudržují správnou evidenci údajů v RPP atd.

Základní charakteristiky stávajícího stavu oblasti:

- nedostatečná specifikace pojmu „agenda“ a souvisejícího věcného datového vymezení,
- neexistence a podcenění významu konceptuálních modelů pro popis dat a jejich významů, a také pro odvození evidovaných údajů,
- nepřipravenost organizací ke konceptuálnímu modelování (vhodní pracovníci, znalosti, dovednosti, nástroje), nedostatečné proškolení a chybějící podpora pro tvorbu konceptuálních modelů,
- nedostatečná znalost způsobů sdílení dat a účelu PPDF, VDF, OD a řízeného přístupu, nedostatečná znalost informační architektury VS,
- neznalost principů a účelu RPP, jeho souvislostí a významu pro sdílení dat VS,
- nedostatečné tempo uvedení stavu evidence údajů v RPP do stavu vyžadovaného Zákonem č.111/2009 Sb. pro zajištění sdílení dat VS (přírůstek evidovaných údajů za období září 2022 až listopad 2023 činí 1,1 %),
- pomalé reakce na straně „centra“ (potvrzení správnosti modelů, výběrů evidovaných údajů apod).
- neexistence standardů pro popis dat a nástrojů pro práci s metadaty,
- kategorizace dat téměř v 90 % případů vychází z GDPR a kybernetického zákona,
- správa číselníků je podhodnocena, schází systémové přístupy.

2.3.4 Oblast realizace datových řešení

Základní charakteristiky stávajícího stavu oblasti:

- podceněný vyžadovaný standard věcného plnění dodávek řešení od dodavatelů (důraz je kladen na prevenci datové kvality, dodávku (izolovaných) datových modelů bez vyžadovaných standardů a požadavků na zajištění konzistence s již existujícími modely),
- interní i externí integrace je založena především na přímém propojení informačních systémů,
- nejčastějšími důvody nesystémového přístupu k integraci jsou technologická nepřipravenost ISVS pro nové způsoby integrace (absence vhodných konektorů), nedostatečné povědomí a znalost nových způsobů sdílení dat, nedůslednost v plnění povinností stanovených legislativou a strategickými dokumenty, a také obtížné nalezení potřebných dat v prostředí VS (kdo je má, kde jsou a jakým způsobem je lze získat),
- sdílení dat prostřednictvím PPDF není zatím běžné, úplné sdílení uvedlo pouze 7 % úřadů, částečné v některých oblastech 33 %.

2.3.5 Oblast zajištění a využívání dat

Tato oblast je řešena relativně dostatečným způsobem, dále uvedené poznatky jsou spíše okrajové.

Základní charakteristiky stávajícího stavu oblasti:

- otevřená data nejsou publikována na základě publikačních plánů, ale převážně způsobem ad hoc,
- zkušenosti s poskytováním neveřejných údajů způsobem řízeného přístupu má zatím menšina organizací, přičemž se většinou jedná o nestandardizovaný způsob zpřístupnění,
- skartace dat a příprava dat k archivaci do Národního archivu je otevřeným problémem, zejména v případě archivace digitálního obsahu, kdy organizace nemají připravené postupy pro jeho předání k archivaci.

2.4 Doplnění pohledu na základě postřehů z rozhovorů mimo strukturu otázek

2.4.1 Obecné postřehy

- Místy nedostatečně zvýrazněný význam a hodnota dat pro organizaci a VS, zazněl i názor, „že data jsou až to poslední čemu se věnují, na prvním místě je informační systém.“
- Aplikační podpora a rozvoj agendových systémů jsou převážně (> 75 %) zajišťovány externími dodavateli, útvary ICT sehrávají roli spíše „prostředníků“ mezi odbornými útvary a dodavateli. Dodávaná řešení jsou z velké míry poznamenána individuálními přístupy jak ze strany dodavatele, tak i ze strany zadavatele.
- Všeobecně téměř všechny organizace zápasí s nedostatečným personálním obsazením, s jeho kvalitou a podfinancováním oblasti ICT. Z tohoto pohledu je většina nových požadovaných povinností z „centra“ nereálných. Organizace nejsou připraveny technologicky, nemají dostatek pracovníků, nedisponují potřebnými kompetencemi, nemají potřebné finance, novým požadavkům často patřičně nerozumí, podpora z „centra“ není dostatečná (schází vysvětlující školení, příklady, návody, podpora). Vyplynávající pragmatická praxe je taková, že pokud je něco nereálné v jejich podmínkách, tak to nedělají, a pokud není zbytí (tlak, nebo sankce), tak se to dělá pouze formálně.
- Ze strany organizací je poměrně často vnímán nesoulad legislativy a prosazované digitalizace VS (buď legislativa předbíhá technologickou připravenost organizací, nebo ji naopak legislativa brzdí v případech, kdy organizace jsou již technologicky připraveny a také dohodnuty si vzájemně data poskytovat, ale stávající legislativa to neumožňuje). Na vrub legislativy jdou také zmínky ohledně jejího zaměření zejména na výkon správy, avšak bez odpovídajícího důrazu na data.
- V případě některých úspěšnějších organizací (z pohledu správy a využití dat) stojí za úspěchy vždy dobrý (omezený) tým vedený lídrem (tzv. „srdcaři“), a to často i bez potřeby výraznější podpory nejvyššího vedení.
- Pozitivní vliv na přístup ke správě dat se ukázal u organizací, které jsou procesně řízeny, a to jak z pohledu vnímání úlohy dat, tak i z pohledu jejich kategorizace, nastavených způsobů jejich správy, stanovení odpovědností, ale i z pohledu častějšího ošetření oblasti dat odpovídajícími interními směrnici.
- Překvapivým zjištěním byl rozsah povědomí o otevřených datech, které je poměrně silné ve všech organizacích (často se pod pojmem „data“ zpodobaným nejprve vybavila oblast „otevřených dat“). Všeobecně aktivity kolem otevřených dat byly hodnoceny (bez přímého dotazování, spíše asociacemi dotazovaných pracovníků) velmi kladně, a to jak z pohledu osvěty, vzdělávání, workshopů, dostupných informací na POD, tak i z pohledu dlouhodobě poskytované podpory týmem otevřených dat. Vypovídajícím dokladem efektivity a správnosti projektu OD je i skutečnost, že roli „koordinátora otevřených dat“ má dobrovolně ustanoveno 78 % organizací, bez toho, že by taková povinnost byla nějakým způsobem stanovena z „centra“.

2.4.2 Rozmanitost „informatické struktury“ úřadů a jejich agendových působení

Rozdělení úřadů na základě provedených rozhovorů:

- Velké úřady s velkým počtem agend
 - finančně i personálně silnější,
 - mají potenciál více specializovaných rolí,
 - podstatnou roli v některých případech sehrává rozsáhlost rezortu a „blízkost“ vykonávaných agend (například srovnání MPSV a MD, nebo MZd a MPO).
- Menší úřady s malým počtem agend
 - finančně i personálně omezenější,
 - mají omezené množství specializovaných rolí,
 - vyplývající nutnost koncentrace rolí na jednoho pracovníka.

- Úřady s distribuovanými systémy agendy – náročné na správu a často i na aplikační podporu.
- Úřady s lokálním systémem agendy, ale s distribuovanými pracovišti (s přenesenou působností i bez ní) – snadná a bezproblémová správa i aplikační podpora.
- Úřady provozující agendy s využívanými výstupy napříč veřejnou správou – obvykle dobře řešena evidence v RPP, kvalitní data a časté sdílení (i když stále převažuje mimo PPDF).
- Úřady provozující v podstatě izolované agendy – obvykle jednoduché a bezproblémové řešení.
- Specifikem je MZV, které provozuje více než tisíc aplikací (v jeho terminologii „agend“) v Lotus Notes, základem informace je „dokument“ (LN je dokumentová databáze) a navíc se potýká s nedostatkem odborných znalostí na většině pracovišť, způsobenou pravidelnou rotací pracovníků na zahraničních úřadech.

2.4.3 Souhrnné společné poznatky

- nedostatečně vnímaná role, význam a hodnota dat pro organizaci a VS,
- technologická nepřipravenost (a nevyváženost) organizací pro vzájemné sdílení údajů prostřednictvím PPDF (externí integrace),
- nedostatečná technologická připravenost infrastruktury veřejné správy pro plnohodnotné sdílení údajů prostřednictvím PPDF (včetně výběrových dávkových přenosů),
- všeobecný nedostatek vhodných a kompetentních pracovníků v organizacích,
- všeobecný nedostatek finančních zdrojů pro úpravy infrastruktury a agendových systémů,
- neznalost a nedostatečné chápání nových požadavků z „centra“ v kontextu celé informační architektury veřejné správy,
- nedostatečná odborná podpora pro realizaci nových požadavků z „centra“, schází vysvětlující školení, příklady, návody, připravená řešení, ...,
- absence shodně chápaných pojmů při vzájemné komunikaci jak s „centrem“, tak často mezi organizacemi navzájem,
- způsob a míra zapojení externích dodavatelů ovlivňuje:
 - charakter činností odborných a informatických útvarů, včetně požadovaných kompetencí, ve všech oblastech správy dat:
 - věcná,
 - technická,
 - provozní,
 - aplikační,
 - konceptuální modelování a existenci standardizovaných konzistentních modelů,
 - datové modelování a forma datových modelů,
 - katalogizaci dat, jejich popisy a metadata,
 - technologickou připravenost ke sdílení údajů,
- schází standard vytvoření poptávky řešení pro zajištění předpokladů pro dobrou správu dat (modely a jejich standardizace, zajištění konzistence s již existujícími modely, zajištění napojení na PPDF, případně VDF, publikace OD atd.).

3 Současný stav – přehled zjištěných informací

Poznání současného stavu správy dat v organizacích VS je založeno na sérii strukturovaných rozhovorů provedených se zástupci 34 vybraných centrálních úřadů a organizací státní správy. Kvalita a podrobnost získaných informací z rozhovorů je rozmanitá, a je do jisté míry poplatná výběru konkrétních účastníků jednotlivých organizací, který byl plně v jejich kompetenci. Z každého provedeného rozhovoru byl vyhotoven zápis a pro zajištění zastřešujícího (celkového a sjednocujícího)

pohledu na stav správy dat v organizacích VS, byly zápisy z rozhovorů dále kvalitativně ohodnoceny pomocí jednoduché tříúrovňové hodnotící stupnice.

Každý prvek (téma oblasti standardu správy dat) validovaného strukturovaného zápisu byl poměřován jednoduchou hodnotící škálou pro získání jeho kvalitativního hodnocení. Ke stanovení kvalitativní hodnoty bude využito následné hodnotící stupnice:

Pokud obsah sdělení odpovídal (v kontextu „standardu správy dat“):

- kvalitnímu (plnému) naplnění posuzovaného prvku – hodnocení „ANO“,
- částečnému plnění (neurčité nebo nepřesné sdělení, ale v každém případě nesrovnatelné s hodnocením „ANO“) – hodnocení „částečně“,
- žádnému plnění – hodnocení „NE“.

Tímto způsobem kvalitativně ohodnocené odpovědi byly zachyceny v sešitu excel a nad nimi vytvořeny všechny dále vložené vizualizace agregovaných odpovědí.

Následně uvedené vytěžené informace z provedených rozhovorů je nutné brát s mírnou rezervou z následujících důvodů:

- rozhovory nebyly směřovány do detailní hloubky jednotlivých témat,
- rozsah témat rozhovorů byl poměrně rozsáhlý,
- nevyhnutelné subjektivní ovlivnění přístupu tazajících i dotazovaných (včetně tendence uvádět ze strany dotazovaných co bude, než co aktuálně je),
- výběrem účastníků rozhovorů ze strany organizací,
- a také z velké míry ne vždy jasně a významově stejně používanými pojmy z obou stran.

Samotné následné posuzování odpovědí pro jejich kvalitativní hodnocení proběhlo spíše pozitivním způsobem, takže prezentované výsledky budou pravděpodobně optimističtější než aktuální realita.

Výčet zkoumaných oblastí správy dat během realizovaných rozhovorů:

- řízení a organizace správy dat,
- stanovení datových potřeb,
- popis dat a datových řešení,
- realizace datových řešení,
- zajištění a využívání dat.

3.1 Oblast řízení a organizace správy dat

Oblasti zájmu rozhovorů:

- Organizace (role, jejich odpovědnosti a pravomoci v oblasti dat).
- Standardizace (zakotvení rolí, odpovědností, procesů a pravidel pro práci s daty v řídicích dokumentech úřadu).
- Rozvoj kompetencí, spolupráce a podpora (vzdělávání v oblasti dat, spolupráce lidí zabývajících se daty uvnitř/vně úřadu, podpora ze strany vedení úřadu).

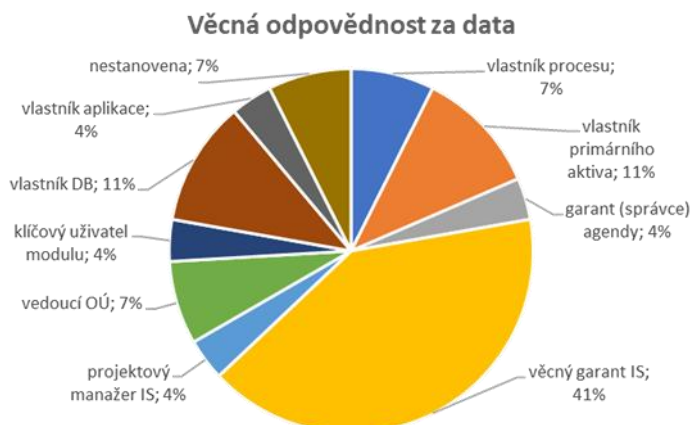
3.1.1 Organizace

Pozitivním zjištěním bylo, že v řadě organizací je vnímána podpora v oblasti dat ze strany nejvyššího vedení, konkrétně ve 30 % organizací je hodnocena jako maximální, v 35 % jako dostatečná a pouze ve 35 % jako nedostatečná. Často ale bylo uváděno, že podpora není stabilní a je nejvíce viditelná v případě problémů, zejména v případech, kdy jsou problémy řešeny veřejně.

Z pohledu jednoznačně stanovené manažerské odpovědnosti za data celého úřadu a jejich správy existuje pouze jeden úřad, který se k ní zatím jenom přibližuje. Pro většinu úřadů se spíše jedná o obtížně uchopitelný problém.

Existuje poměrně různorodé vnímání pojmu „agenda“ (např. dle legislativy, jako informační systém, oblast dat, hlavní proces, aplikace, apod), což má vliv na řadu dalších oblastí, jako je věcná odpovědnost za data, modelování agend a evidence jejich údajů (dle zákona č.111/2009 Sb.), vymezení oblastí dat, jejich prioritizaci, a také na samotnou komunikaci o těchto tématech. Nejednotnost chápání pojmu „agenda“ je viditelná v odpovědích na stanovení věcné odpovědnosti za data.

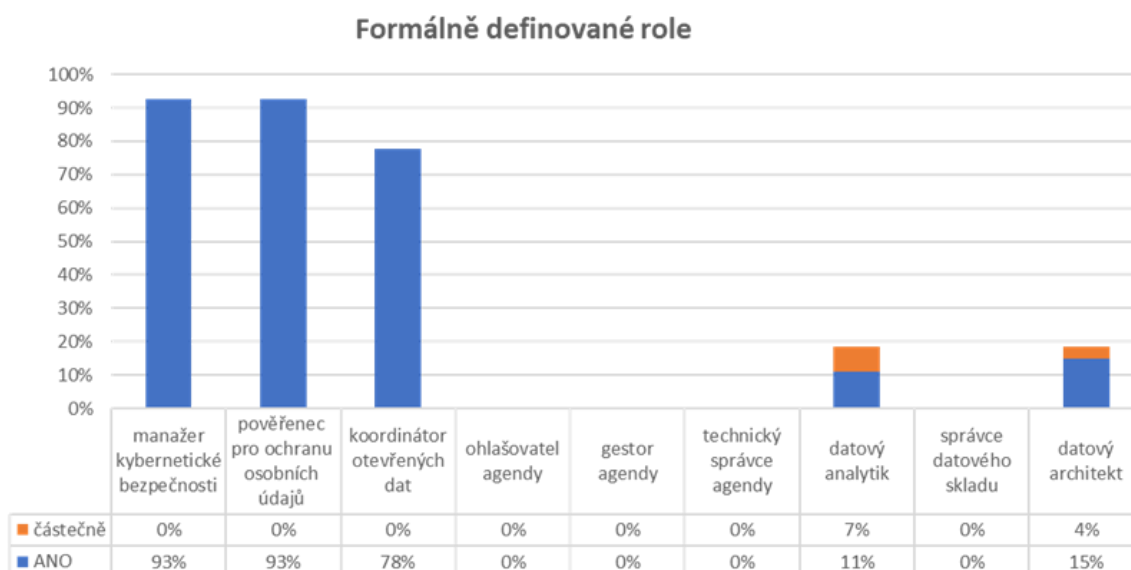
Graf č. 1



Zdroj: vlastní DIA

Tato nejednotnost se také ukázala ve zjištění existence datových rolí v organizacích, neboť žádná organizace nemá formálně zavedené role, které by byly nějakým způsobem spojeny s pojmem „agenda“.

Graf č. 2

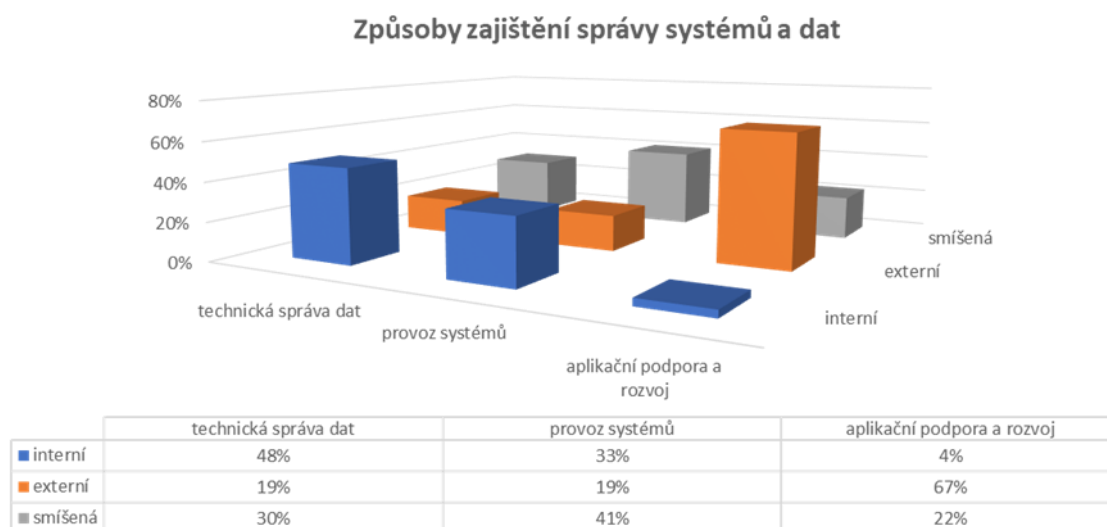


Zdroj: vlastní DIA

Z uvedených datových rolí v obrázku je zřejmé, že ačkoliv si organizace začínají uvědomovat význam dat, tak na organizační úrovni to zatím není patrné. Jediné role nějak spojené s daty vyplývají z povinnosti zavedení GDPR a kybernetické bezpečnosti. Potěšující je, že téměř 78 % organizací má zavedenou roli „koordinátora otevřených dat“, bez toho, že by jim to nařizovala nějaká legislativa nebo směrnice.

V oblasti zajištění správy systémů a dat se u organizací jedná o velkou různorodost z pohledu zapojení a kompetencí externích dodavatelů a interních útvarů organizací. Dotýká se v podstatě všech oblastí správy dat (věcné, technické, provozní, aplikační), a také s následným dopadem do konceptuálního a datového modelování, katalogizace dat, jejich popisů, metadat a technologické připravenosti ke sdílení údajů atd.

Graf č. 3

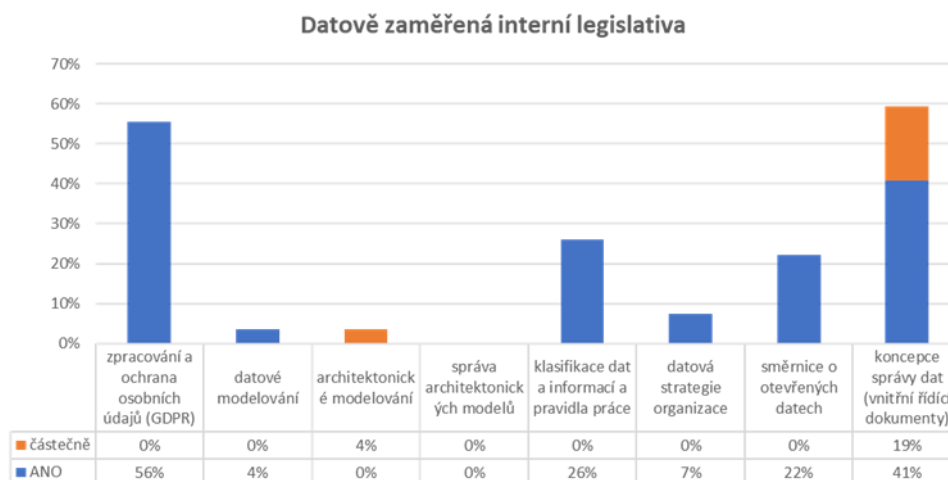


Zdroj: vlastní DIA

3.1.2 Standardizace

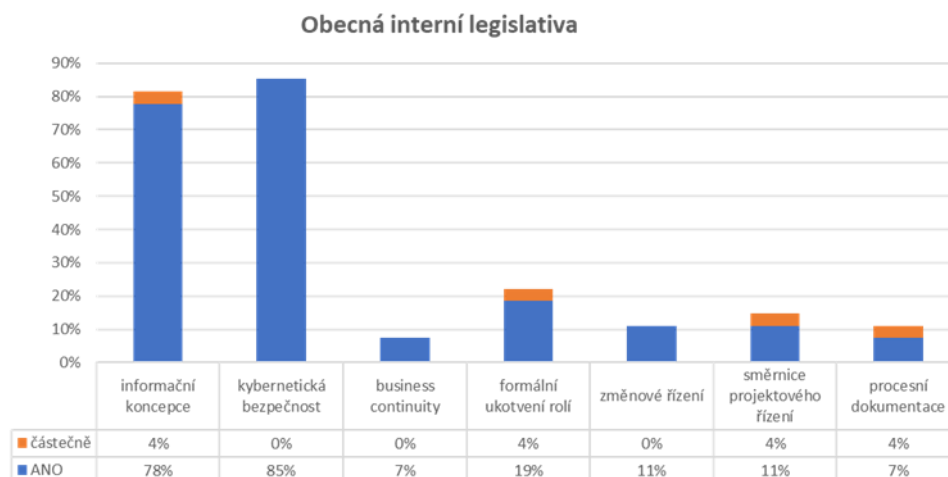
K ukotvení rolí, odpovědností, procesů a pravidel pro práci s daty zatím v organizacích neexistuje dostatečná interní legislativa zaměřená datově, a stejně jako v případě datových rolí existují převážně obecněji zaměřené dokumenty k GDPR a kybernetické bezpečnosti. Oblasti ICT se věnují informační koncepce, ve kterých specifické zaměření na data vesměs schází, u některých organizací lze ale rozpoznat náznaky zaměření i na data, jedná se však o jednotlivé případy. Určitou výjimku představují občas zmiňované dokumenty zastřešené pojmem „koncepce správy dat“, ale účastníci rozhovorů o jejich obsahu mluvili neurčitě, často se jednalo o směs blíže nespecifikovaných vnitřních řídicích dokumentů.

Graf č. 4



Zdroj: vlastní DIA

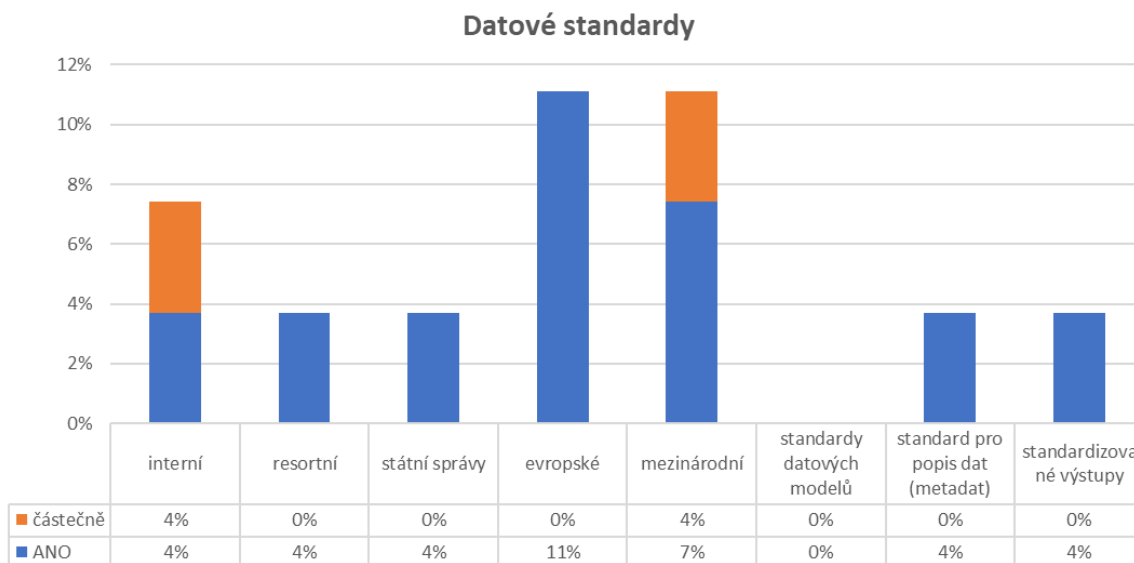
Graf č. 5



Zdroj: vlastní DIA

Z rozhovorů rovněž nevyplynulo (až na pouhých několik výjimek – 7 %), že by organizace nějakým způsobem metodicky řešily práci s daty a využívaly standardy, mimo požadavků napojení na evropské systémy (11 %), nebo vykazování směrem k mezinárodním institucím (7 %). V těchto případech jsou řešení většinou založena na transformaci (exportu/importu) z interních systémů do externě stanovených formátů. Nadřazené organizace se rovněž nezabývají standardizací ani v rámci resortů (4 %), které metodicky vedou pouze okrajově, ani v rámci státní správy (4 %). Svou část viny na tom pravděpodobně má i „centrum digitalizace“, neboť standardy nejsou výrazněji z „centra“ prosazovány, vydávány, a také související metodické pokyny jsou v tomto směru formulovány často vágním způsobem.

Graf č. 6

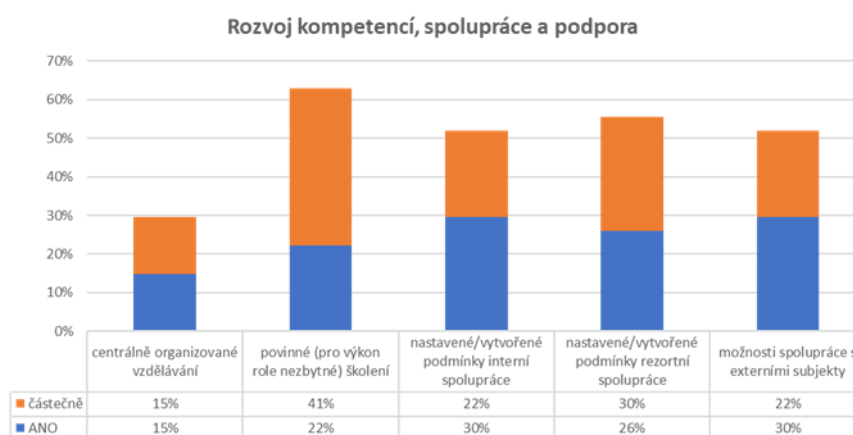


Zdroj: vlastní DIA

3.1.3 Rozvoj kompetencí, spolupráce a podpora

Tato část je poměrně nepřesná, neboť účastníci rozhovorů příliš nerozlišovali mezi rozvojem kompetencí zaměřeným na data a všeobecným pohledem na oblast vzdělávání a spolupráce. I přesto nejsou výsledky povzbudivé, neboť v rámci rozhovorů téměř 70 % účastníků prohlásilo, že centrálně organizovaná školení se mívá účinkem, a dokonce zaznělo, že centrálně organizovaná školení nevychází z potřeb odborných útvarů a s oblastí ICT se doslova „nepotkávají“. Možnosti interní i externí spolupráce byly hodnoceny z více než 50 % jako ne zcela skvělé, ale dostatečné.

Graf č. 7



Zdroj: vlastní DIA

3.1.4 Základní zjištění v oblasti řízení a organizace správy dat:

- Absence jednoznačně stanovené manažerské odpovědnosti za data celého úřadu a jejich správy.
- Nejednoznačné vnímání pojmu „agenda“.
- Schází datově zaměřené role (pouze v 5 úřadech „datový analytik“ a ve 4 „datový architekt“) a jejich formalizace (implementace do stávajících pracovních pozic). Stávající

role jsou převážně pouze stanovené zákonem nebo vyhláškami (manažer kybernetické bezpečnosti, pověřenec pro ochranu osobních údajů). Výjimkou je koordinátor otevřených dat, který není stanoven legislativou, a přesto jej má téměř 80 % úřadů ustanoveno.

- Interní legislativa zaměřená na data se částečně objevuje ve směrnících zaměřených na zpracování a ochranu osobních údajů (58 %), kybernetickou bezpečnost (85 %) a v informačních koncepcích (81 %). Jedna organizace má směrnici k datovému modelování, 6 úřadů ke klasifikaci dat, informací a práci s nimi, 2 úřady mají samostatnou datovou strategii úřadu, 6 úřadů má směrnici k otevřeným datům.
- Standardy v oblasti dat, a tedy důraz a pozornost zaměřená na interoperabilitu dat a snadnost jejich sdílení je zcela přehlížena.
- Rozvoj kompetencí (vzdělávání, spolupráce, podpora) je výrazně podceněn a plnohodnotné centrálně organizované vzdělávání se vyskytuje pouze v 15 % organizacích. Jelikož odpovědi zahrnovaly spíše vzdělávání v oblasti ICT, dá se předpokládat, že v oblasti dat bude situace výrazně horší. V rozhovorech zazněly zmínky, že centrálně organizované vzdělávání a potřeby v oblasti ICT se nepotkávají.

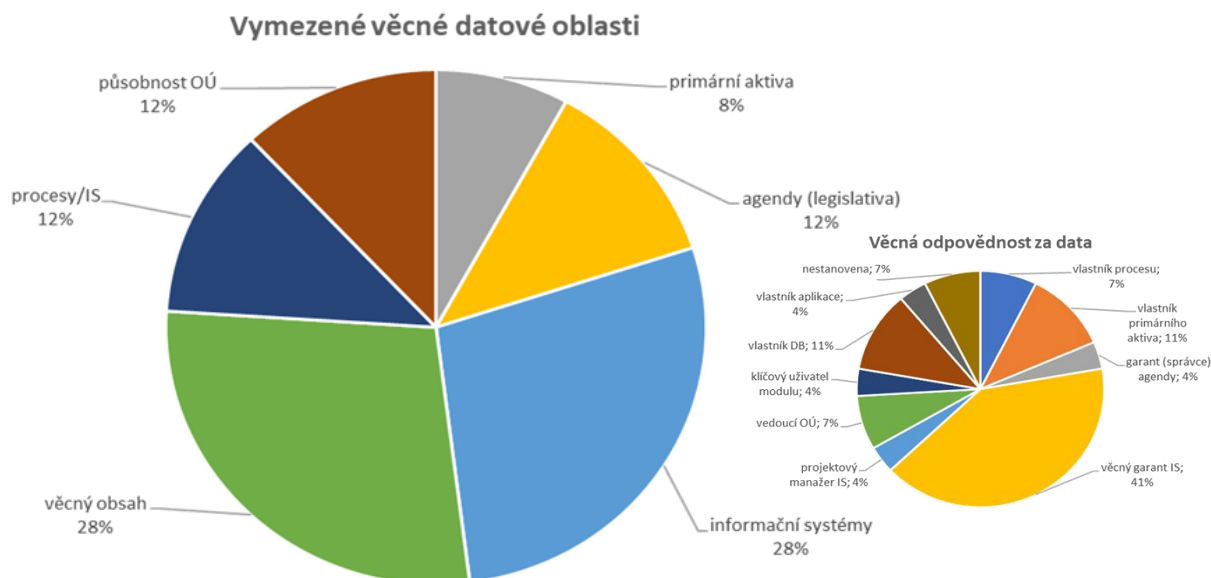
3.2 Stanovení datových potřeb

Oblasti zájmu rozhovorů:

- Prioritizace oblastí (využívání a důležitost dat v jednotlivých věcných oblastech).
- Řízení potřeb (identifikace a řízení požadavků na data / datové výstupy).
- Využití architektury pro plánování změn (zohlednění datových potřeb v architektuře / návrhu cílového stavu úřadu).

3.2.1 Prioritizace oblastí

Graf č. 8.



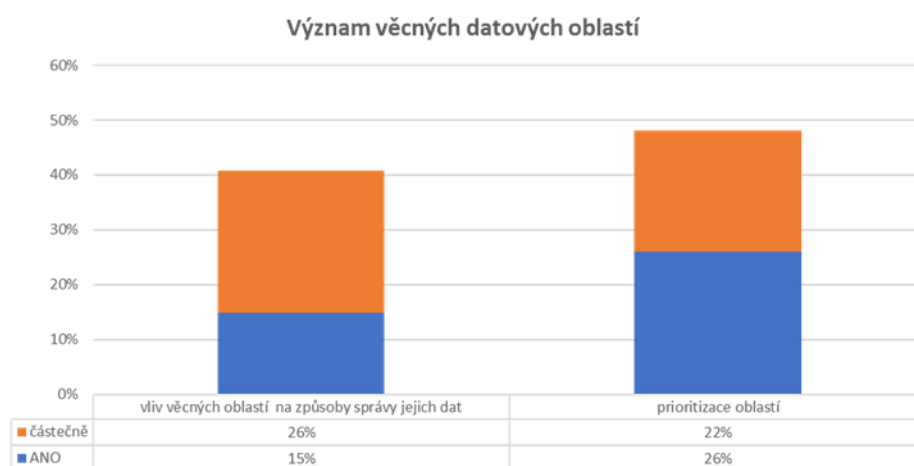
Zdroj: vlastní DIA

Vymezení věcných datových oblastí je obdobně různorodé jako vnímání odpovědnosti za data. Na první pohled je ale zřejmé, že při vymezení datových oblastí je mnohem větší důraz kladen na věcný obsah

dat, než na odpovědnost vyplývající z ryze „ICT pohledů“. Vymezení oblastí vyplývající z věcného obsahu dat (věcný obsah, agenda, působnost OÚ) představují celkový podíl 52 %, ve srovnání s věcnou odpovědností za data (garant agendy, vedoucí OÚ, klíčový uživatel a vlastník procesu) s celkovým podílem pouhých 22 %. V případě stanovení odpovědnosti za data je naopak míra „ICT pohledu“ (věcný garant IS, vlastník DB, vlastník aplikace a projektový manažer) celkem 60 %, ve srovnání s věcnými oblastmi, kde jejich míra představuje pouze (informační systémy) 28 %. Výsledky jsou zcela opačné, faktem je ale skutečnost, že věcná odpovědnost za data by měla být primárně ve shodě s vymezením věcných datových oblastí.

Přístup k jednotlivým věcným datovým oblastem závisí na jejich charakteru, pokud se jedná o oblasti „životně důležité“, nebo oblasti se sdílenými údaji napříč celou veřejnou správou, pak organizace uplatňují jejich prioritizaci s dopadem do způsobu jejich správy. V opačných případech všechny organizace uvádějí, že spravují všechny oblasti jednotně a s patřičnou péčí bez ohledu na způsob prioritizace. V takových případech není zřejmé, jaký je význam případné prioritizace.

Graf č. 9

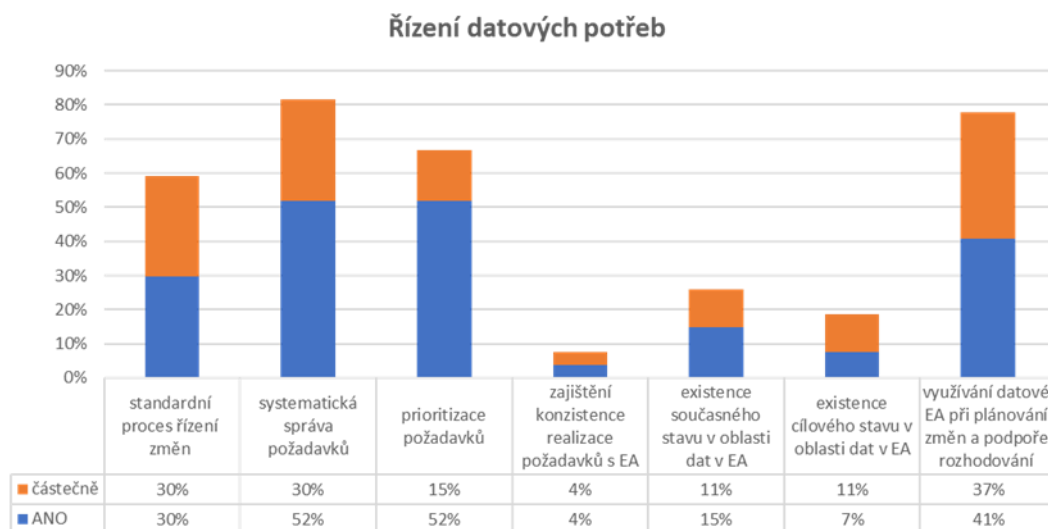


Zdroj: vlastní DIA

3.2.2 Řízení datových potřeb

Z pohledu identifikace a řízení požadavků na data / datové výstupy většina organizací spravuje požadavky a jejich prioritizaci systematicky, více než polovina z nich k tomu plně (nebo částečně) používá standardní proces pro řízení změn. Souhrnně lze tuto oblast považovat za uspokojivě řešenou. Pouze informace o zajištění konzistence realizace požadavků s EA je v rozporu s dále uvedenou informací o využívání datové EA při plánování změn, více na toto téma v části „Využití architektury pro plánování změn“.

Graf č. 10



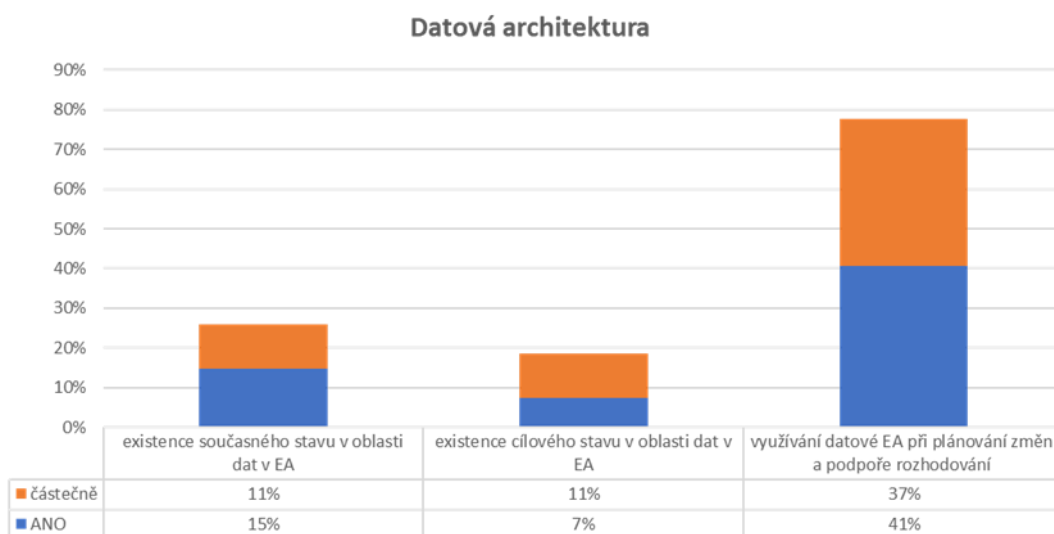
Zdroj: vlastní DIA

3.2.3 Využití architektury pro plánování změn

Řízení datových potřeb z pohledu zajištění jejich konzistence s datovou vrstvou EA je problematické, neboť většina organizací datovou vrstvu v EA nemá modelovanou, a to ani současnou, ani její budoucí podobu. Překvapující může proto být zjištění u využívání datové vrstvy EA při plánování změn a podpoře rozhodování, ke kterému se hlásila většina organizací. Její využívání uvádělo 41 % organizací jako plnohodnotné a 37 % organizací jako částečně, avšak vzhledem ke skutečnosti, že pouze 26 % organizací má vytvořenou datovou architekturu současného stavu v EA a jenom 18 % i cílovou datovou architekturu, nelze uváděnou informaci považovat za reálnou. Pravděpodobně půjde o zkreslení, neboť organizace dnes v informačních koncepcích uvádějí především hrubé a zastřešující EA, což může vyvolávat dojem, že s nimi i aktivně pracují a na základě nich se také i rozhodují.

Souhrnně lze tedy konstatovat, že oblast EA a zejména její datové vrstvy jsou stále ještě na začátku plnohodnotného využití.

Graf č. 11



Zdroj: vlastní DIA

3.2.4 Základní zjištění v oblasti stanovení datových potřeb:

- Převažuje nesoulad mezi vymezením věcných datových oblastí v úřadech a stanovením věcných odpovědností ze jejich data.
- Způsob prioritizace věcných datových oblastí není řešen systémově a je bez konkrétního dalšího využití.
- Enterprise architektura (EA) a zejména její datové vrstvy stále hrají nepodstatnou roli jak ve správě a realizaci změnových požadavků, tak i při plánování a rozhodování v oblasti dat. Důvodem je zatím převažující neexistence současného i budoucího stavu v oblasti dat v EA.

3.3 Popis dat a datových řešení

Oblasti zájmu rozhovorů:

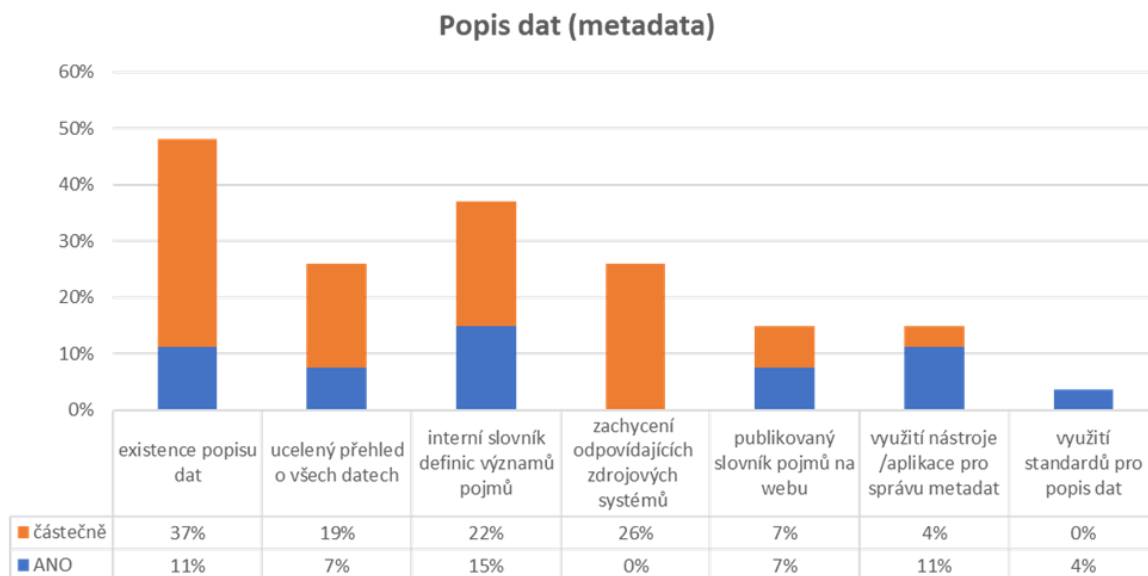
- Popis dat (přehled o spravovaných datech, podoba a rozsah popisu dat).
- Evidence v RPP (stav evidence agendových údajů v Registru práv a povinností).
- Modelování dat a datových řešení (modely dat, datových toků a komponent využívaných při získávání, ukládání a zpracování dat).
- Kategorizace dat (používané kategorie dat, číselníky a jejich správa, data přebíraná zvenčí, data sdílená napříč agendami).

3.3.1 Popis dat

Přehled o datech organizací a jejich popisech spadají mezi nejvíce podhodnocené oblasti správy dat. Organizace neudržují přehledy o svých datech, jejich významech a metadatových popisech. Výjimkou jsou organizace ČSÚ a NA (Národní archiv), které jsou primárně zaměřené na práci s metadaty a jakýkoliv údaj bez popisných metadat je pro ně v podstatě nepoužitelný. Výkon jejich agend je orientován na práci s metadaty a k tomu jsou navrženy i jejich informační systémy. Organizace většinou nespravují vlastní interní slovníky významů pojmů, a vesměs takové případné slovníky nejsou ani dostupné na webu organizace. V případě provozování portálů organizací jsou částečně publikovány popisy a vysvětlivky zveřejňovaných informací, avšak zdaleka je nelze považovat za řádně spravované slovníky, i z toho důvodu, že na portálech nebývají data, ale až informace vzniklé jejich zpracováním. Naprostá většina organizací rovněž nevede informace o zdrojích dat (jejich dohledatelnosti). Až na uvedené výjimky (ČSÚ a NA), organizace nemají k dispozici nástroje pro správu dat a ani neznají (nepoužívají) vhodné standardy.

Význam metadat organizace vesměs nedohlíží, chápou jejich význam především ve spojení se spisovými službami a nově částečně i v kontextu otevřených dat. Častou obhajobou stávajícího neutěšeného stavu bylo odkazování na dokumentace konkrétních ISVS s tím, že v nich jsou většinou jednotlivé datové položky uvedeny a z jejich názvů je pro (zainteresované) zřejmý i jejich význam. Neplatí to samozřejmě automaticky pro všechny ISVS, vždy záleží na obsahu dodávky řešení od externích dodavatelů, a také poměrně často v tom sehrává podstatnou roli i stáří systémů. Historické systémy představují problém a nedostatek vhodných informací.

Graf č. 12



Zdroj: vlastní DIA

3.3.2 Modelování dat a datových řešení

Pro modelování dat a datových řešení platí totéž co pro popisy dat. V 90 % platí, že aplikační podpora a rozvoj systémů je buď téměř (67 %) nebo částečně (22 %) v rukou externích dodavatelů, a modelování dat automaticky také. Existence a dostupnost modelů je pak otázkou obsahu dodávky řešení od externích dodavatelů, nastavených smluvních podmínek a také otázkou stáří systémů.

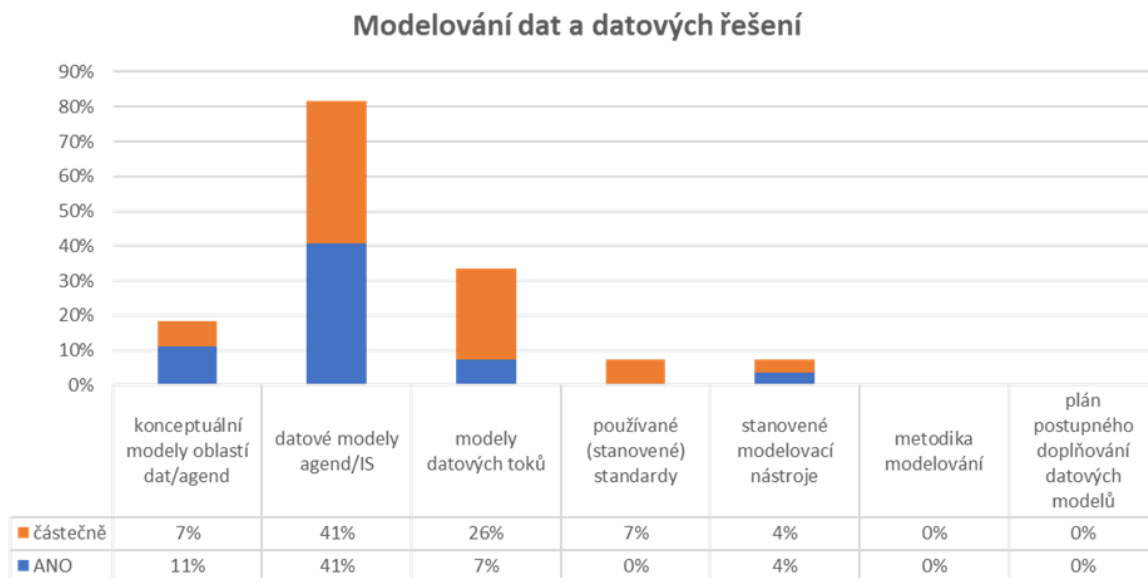
To, co by mělo být nezávislé na dodavateli, a naopak plně „v rukou“ věcných správců datových oblastí (agend), znalých odpovídající legislativy a často i jejich tvůrců, představují technologicky nezávislé konceptuální modely. Situace v této oblasti není uspokojivá, pouze 11 % organizací má konceptuální modely vytvořeny, 7 % uvádí, že je má jen částečně. Metodika evidence údajů v RPP se o konceptuální modely opírá, neboť představují ideální východisko pro správné odvození evidovaných údajů v RPP. Navíc slouží pro vytvoření uceleného kontextu spravované oblasti a je vždy východiskem pro návrh vhodného datového modelu ISVS (relační, hierarchický, grafový atd.).

Datové modely jsou hodnoceny jako převážně dostupné, avšak je nutné zdůraznit, že z pohledů účastníků rozhovorů byly jako datové modely považovány nejrozumnější formy exportů tabulek ze systémů, stejně jako modely popsané v dokumentech.

Modely datových toků konkrétních dat v podstatě organizace nemají, v lepším případě mají zachyceny toky celých datových objektů mezi aplikacemi na „vysoké úrovni“.

Modelování jako takové je zcela podceněno, nemají stanoveny modelovací standardy, modelovací nástroje, metodiky modelování, ani plán případného doplnění chybějících skutečných datových modelů. Svou vinu na tom bohužel nese i „centrum“, které se této oblasti nevěnovalo dostatečným způsobem a neposkytlo organizacím patřičnou metodickou a poradenskou podporu.

Graf č. 13



Zdroj: vlastní DIA

3.3.3 Evidence v RPP

K hodnocení stavu evidence údajů v RPP a zachycení jejího pokroku (trendu) byly využity informace z roku 2022 (listopad), kdy v rámci projektu „KODI“, byla provedena [„Analýza existujících informačních systémů veřejné správy vč. jejich příslušné legislativy a identifikace datových sad pro veřejný datový fond“](#), a podobným způsobem (využití stejné [excelové přílohy](#)) byly získány také informace za rok 2023.

Rozhovory byly využity ke zjištění souvisejících problémů spojených s praktickým prováděním evidence v RPP.

Aktuální stav evidence údajů agend dle jednotlivých ohlašovatelů je zachycen v následující tabulce. Žlutě podbarvené organizace nebyly součástí rozhovorů.

Organizace	Počet ohlášených agend	Počet agend s údaji	Počet evidovaných subjektů/objektů	Počet evidovaných údajů
ČESKÁ NÁRODNÍ BANKA	2	0	0	0
Český báňský úřad	6	1	1	1
Český statistický úřad	4	1	1	20
Český telekomunikační úřad	4	0	0	0
Český úřad zeměměřický a katastrální	4	4	79	478
Digitální a informační agentura	16	4	16	94
Ministerstvo dopravy	9	7	26	158
Ministerstvo financí	29	2	13	80
Ministerstvo kultury	36	0	0	0
Ministerstvo obrany	16	0	0	0

Ministerstvo práce a sociálních věcí	10	1	4	36
Ministerstvo pro místní rozvoj	12	1	1	5
Ministerstvo průmyslu a obchodu	37	1	1	16
Ministerstvo spravedlnosti	26	6	9	97
Ministerstvo školství, mládeže a tělovýchovy	15	2	9	35
Ministerstvo vnitra	56	16	17	405
Ministerstvo zahraničních věcí	5	0	0	0
Ministerstvo zdravotnictví	22	6	26	160
Ministerstvo zemědělství	31	3	7	62
Ministerstvo životního prostředí	25	7	8	14
Národní bezpečnostní úřad	1	0	0	0
Národní sportovní agentura	1	0	0	0
Národní úřad pro kybernetickou a informační bezpečnost	2	0	0	0
Nejvyšší kontrolní úřad	1	0	0	0
Rada pro rozhlasové a televizní vysílání	2	0	0	0
Správa státních hmotných rezerv	3	0	0	0
Státní úřad pro jadernou bezpečnost	4	0	0	0
Úřad pro ochranu hospodářské soutěže	5	0	0	0
Úřad pro ochranu osobních údajů	2	0	0	0
Úřad průmyslového vlastnictví	7	0	0	0
Úřad vlády České republiky	2	0	0	0
Celkový součet	395	62	218	1661

Souhrnné změny sledovaných údajů mezi roky 2022 a 2023

Sledovaný údaj	2022	2023	Meziroční změna %
Počet ohlašovatelů agend	30	31	3,33 %
Počet ohlášených agend	388	395	1,80 %
Počet agend s evidovanými údaji	63	62	-1,59 %
Počet ohlašovatelů s evidovanými údaji	14	15	7,14 %
Počet evidovaných subjektů a objektů práva	215	218	1,40 %
Počet evidovaných údajů	1643	1661	1,10 %

Zvýšení počtu ohlašovatelů souvisí se vznikem Digitální a informační agentury (DIA) a převzetím některých agend od MV ČR.

Graf č. 14



Zdroj: vlastní DIA

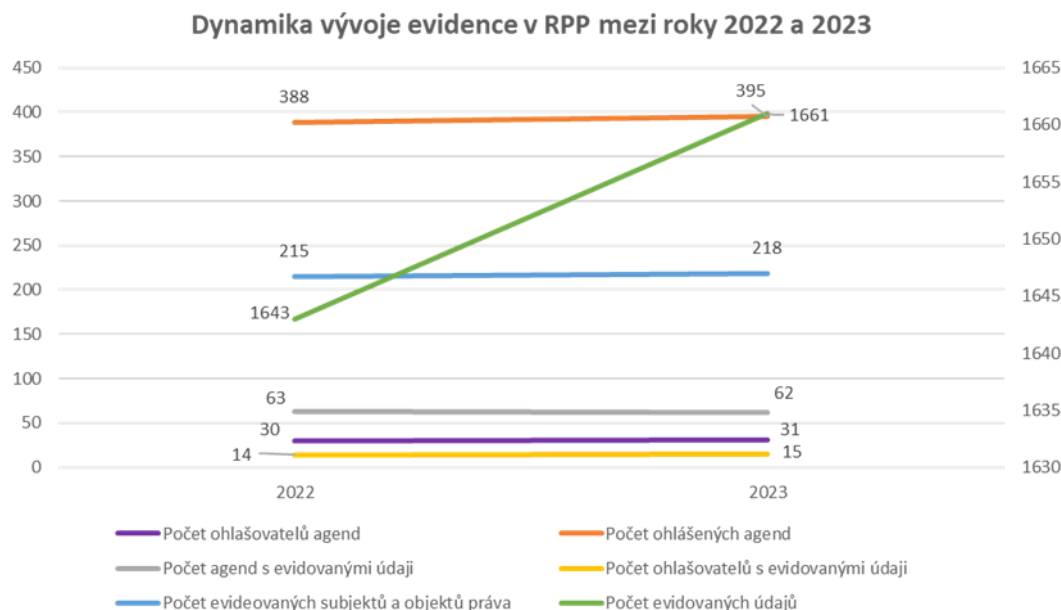
Graf č. 15



Zdroj: vlastní DIA

Z uvedených grafů a tabulky je zřejmé, že se během celého roku 2023 v evidenci údajů v RPP nic podstatného neudálo a stav je v podstatě setrvalý a neměnný.

Graf č. 16



Zdroj: vlastní DIA

Identifikované základní problémové oblasti evidence údajů v RPP:

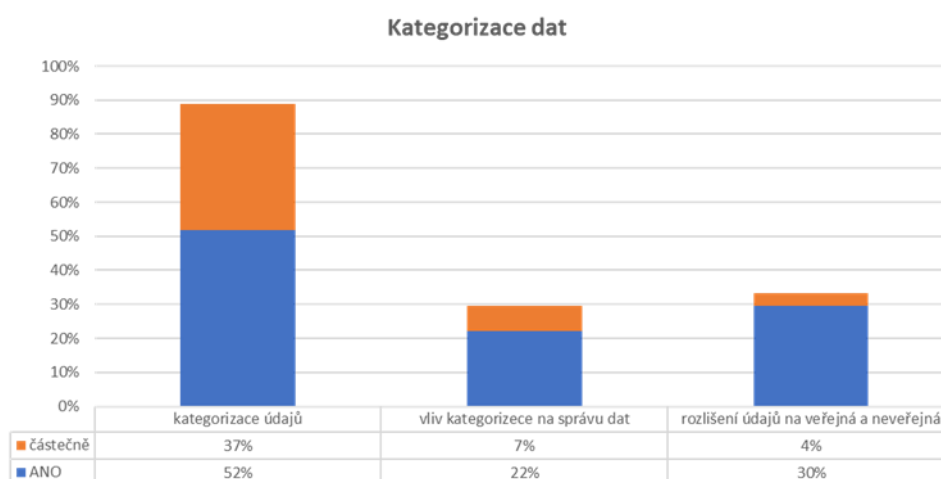
- Nedostatečná specifikace pojmu „agenda“ a souvisejícího věcného datového vymezení (organizace žijí v domněnku, že je nutné evidovat všechna data).
- Neznalost principů a účelu RPP, jeho souvislostí se sdílením dat.
- Nedostatečné tempo uvedení stavu evidence údajů v RPP do stavu vyžadovaného Zákonem č.111/2009 Sb.

- Neexistence a podcenění významu konceptuálních modelů pro odvození evidovaných údajů.
- Nepřipravenost organizací ke konceptuálnímu modelování (vhodní pracovníci, znalosti, dovednosti, nástroje).
- Nedostatečné proškolení a chybějící podpora pro tvorbu konceptuálních modelů.
- Absence jednoduchého modelovacího nástroje.
- Komplikované rozdělení povinností spojených s přípravou informací k evidenci údajů ve vztahu ohlašovatele agendy a faktického vykonavatele (podřízené resortní organizace).
- Malá povědomost o způsobech sdílení dat a účelu PPDF, VDF, OD a řízeného přístupu.
- Pomalé reakce na straně „centra“ (potvrzení správnosti modelů, výběrů evidovaných údajů apod).

3.3.4 Kategorizace dat

Organizace kategorizují data téměř v 90 % případů, avšak kritéria pro jejich kategorizaci vycházejí především z GDPR a kybernetického zákona, případně často organizace rozlišují data na provozní a data sloužící k výkonu agend. Prakticky má kategorizace minimální dopad na samotnou správu dat, spíše se uplatňuje až v případě bezpečnosti dat a případných rizik. Zajímavým zjištěním bylo, že rozlišování dat na veřejná a neveřejná není tak obvyklé, jak by se dalo očekávat zejména v souvislosti s otevřenými daty.

Graf č. 17

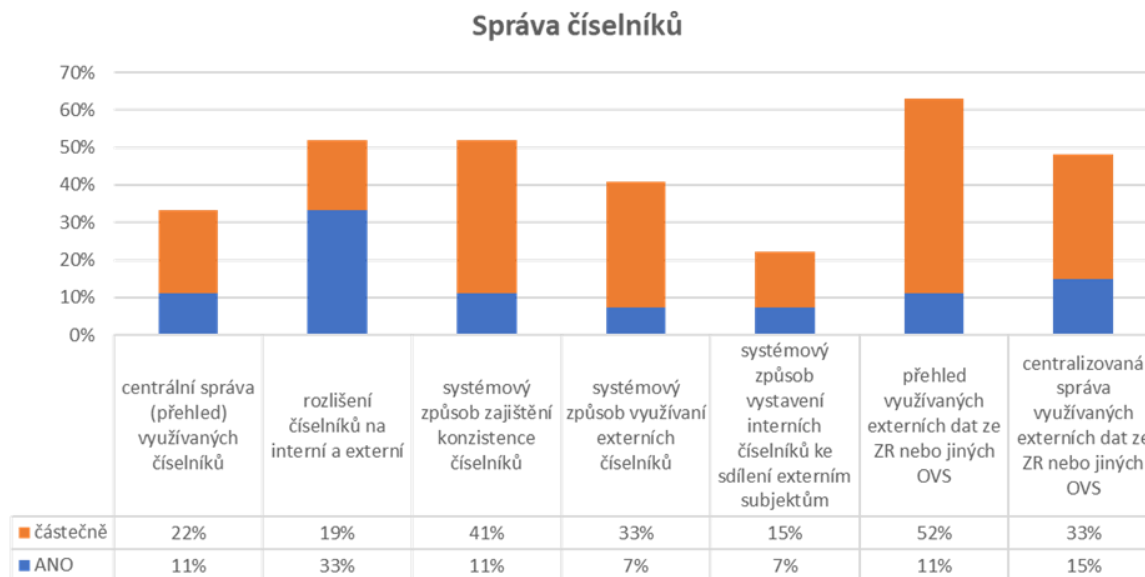


Zdroj: vlastní DIA

Samostatnou kategorií jsou samozřejmě číselníky, i když organizace je jako specifickou kategorii nevnímají. Správa číselníků je všeobecně podceněna a není systémově řešena, jak by se očekávalo. Centrální správu číselníků má úplně zavedenou pouze 11 % organizací a částečně v některých oblastech dalších 22 %. Úplnou konzistenci všech číselníků zajišťuje pouze 11 % organizací, částečně v některých oblastech až 41 %, ale je nutné zdůraznit, že obecně téměř všechny organizace připouští nekonzistence číselníků, jejich vzájemné nesdílení. Je běžné, že jednotlivé IS používají duplicitně některé číselníky s tím, že jejich synchronizace neprobíhají on-line, a obdobně to platí i pro externě využívané číselníky. Pozitivem je fakt, že některé organizace si již zvykly publikovat své číselníky pro externí využití ve formátu otevřených dat.

Úplný přehled využívání dat ze ZR nebo jiných AIS si organizace důsledně nevedou, stejně jako neřeší centrální správu (zajištění) externích dat pro využití interními systémy. Je běžné, že každý interní systém přistupuje k externím datům individuálně.

Graf č. 18



Zdroj: vlastní DIA

3.3.5 Základní zjištění v oblasti popisu dat a datových řešení:

- Přehled o datech organizací a jejich popisech spadají mezi nejvíce podhodnocené oblasti správy dat. Organizace neudržují přehledy o svých datech, jejich významech a metadatových popisech.
- Naprostá většina organizací rovněž nevede informace o zdrojích dat (jejich dohledatelnosti).
- Význam metadat organizace vesměs nedohlíží, částečně chápou jejich význam především ve spojení se spisovými službami a nově i v kontextu otevřených dat. Často považují informace o datových položkách v dokumentacích IS za dostatečné pro pochopení jejich významu.
- Historické systémy představují z pohledu popisu dat a dostupnosti vhodných informací o datech největší problém.
- Ve VS v podstatě neexistují standardy pro popis dat stejně jako nástroje pro práci s metadaty (výjimkou jsou pouze ČSÚ a NA).
- Pouze 11 % organizací má vytvořeny konceptuální modely, 7 % uvádí, že je má jen částečně. Metodika evidence údajů v RPP se přitom o konceptuální modely opírá jako o východisko pro správné odvození evidovaných údajů v RPP.
- Existence a dostupnost datových modelů je otázkou obsahu dodávky řešení od externích dodavatelů, nastavených smluvních podmínek a také otázkou stáří systémů. Datové modely jsou hodnoceny jako převážně dostupné, avšak jako datové modely jsou úřady vnímány i nejrůznější formy exportů tabulek ze systémů a modely popsané v dokumentech.
- Modely datových toků konkrétních dat organizace v podstatě nemají, v lepším případě mají zachyceny toky celých datových objektů mezi aplikacemi na „vysoké úrovni“.
- Modelování dat je zcela podceněno, nejsou stanoveny modelovací standardy, modelovací nástroje, metodiky modelování.

- V evidenci údajů v RPP se během celého roku 2023 ve srovnání s rokem 2022 nic podstatného neudálo a stav je v podstatě setrvalý a neměnný. Změnou je pouze rozšíření počtu ohlašovatelů o agenturu DIA, rozšíření počtu evidovaných subjektů a objektů práva o 3 a nárůst počtu evidovaných údajů o 18. Je ale nutné zdůraznit, že úřady během rozhovorů uváděly nárůst aktivit v této oblasti a úzkým místem se tak pravděpodobně stává „centrum digitalizace“.
- Během rozhovorů byla s evidencí údajů v RPP zachycena řada souvisejících problémů:
 - Nedostatečná specifikace pojmu „agenda“ a souvisejícího věcného datového vymezení (organizace žijí v domněnku, že je nutné evidovat všechna data).
 - Neznalost principů a účelu RPP, jeho souvislostí se sdílením dat.
 - Neexistence a podcenění významu konceptuálních modelů pro odvození evidovaných údajů.
 - Nepřipravenost organizací ke konceptuálnímu modelování (vhodní pracovníci, znalosti, dovednosti, nástroje).
 - Nedostatečné proškolení a chybějící podpora pro tvorbu konceptuálních modelů.
 - Absence jednoduchého modelovacího nástroje.
 - Komplikované rozdělení povinností spojených s přípravou informací k evidenci údajů ve vztahu ohlašovatele agendy a faktického vykonavatele (podřízené resortní organizace).
 - Malá povědomost o způsobech sdílení dat a účelu PPDF, VDF, OD a řízeného přístupu.
 - Pomalé reakce na straně „centra“ (potvrzení správnosti modelů, výběrů evidovaných údajů apod).
- Kategorizace dat téměř v 90 % případů vychází z GDPR a kybernetického zákona a má minimální dopad na samotnou správu dat, spíše se uplatňuje až v případě bezpečnosti dat a případných rizik.
- Správa číselníků je podhodnocena, není systémově řešena, jejich úplnou konzistenci zajišťuje pouze 11 % organizací, většina organizací připouští naopak nekonzistence číselníků, jejich vzájemné nesdílení, duplicity apod.
- Přehled využívání dat ze ZR nebo jiných AIS si organizace důsledně nevedou, stejně jako neřeší centrální správu (zajištění) externích dat pro využití interními systémy, je běžné, že každý interní systém přistupuje k externím datům individuálně.

3.4 Realizace datových řešení

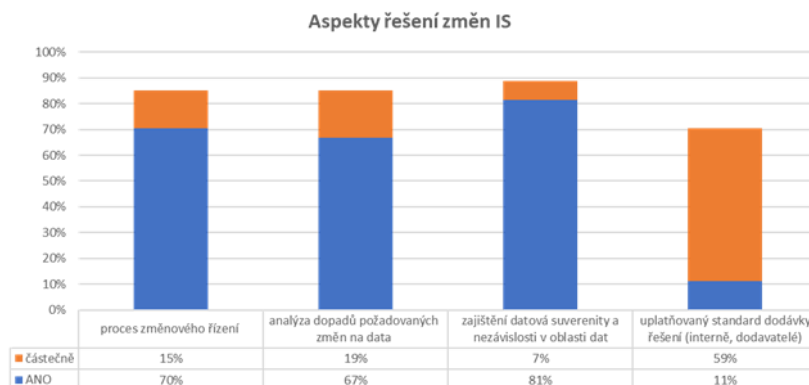
Oblasti zájmu rozhovorů:

- Aspekty řešení změn IS (způsob řešení projektů/změn zaměřených primárně na dodání dat nebo datových výstupů).
- Datová integrace (způsob výměny dat interně mezi agendami a externě s jinými úřady).

3.4.1 Aspekty řešení změn IS

Způsoby řešení projektů/změn zaměřených na dodání dat nebo datových výstupů jsou ze strany organizací řešeny poměrně standardizovaným způsobem, většina organizací má k tomu implementovaný standardní proces změnového řízení včetně související analýzy dopadů požadovaných změn na data. Rovněž se zdá, že je dnes již běžným standardem zajištění datové suverenity a nezávislosti. Bohužel je to ale pouze jeden prvek standardu dodávky řešení. V ostatních částech standardní dodávky, již organizace nejsou tak důsledné, ve větším rozsahu standard uplatňuje pouze 11 % organizací, částečně v některých bodech až 59 %.

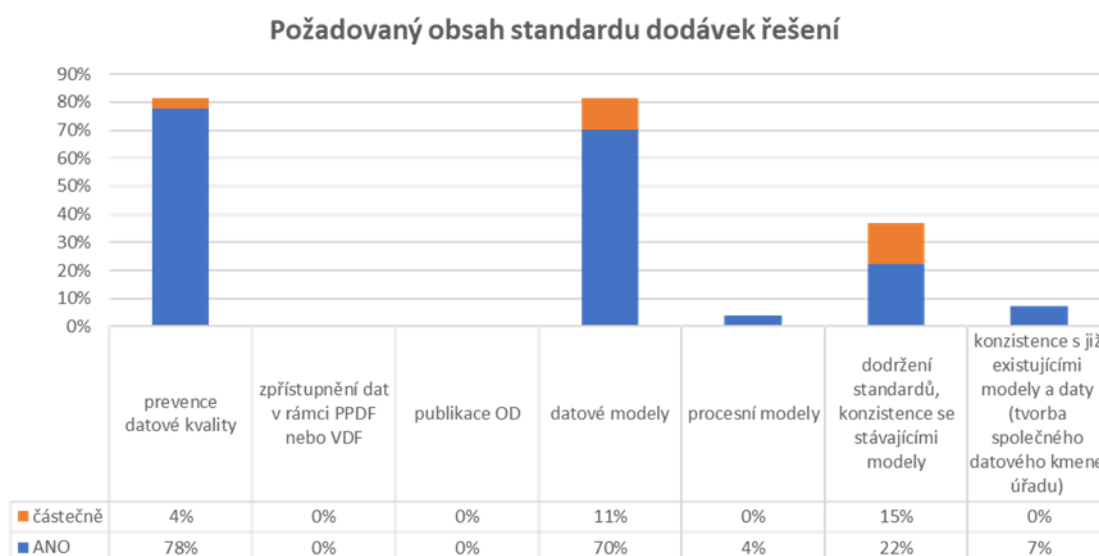
Graf č. 19



Zdroj: vlastní DIA

V detailnějším pohledu na některé důležité prvky standardu dodávky řešení je situace docela tristní. Organizace kladou důraz na prevenci datové kvality (zejména na vstupu), dodávce (izolovaných) datových modelů, avšak bez striktně předepsaných standardů, a pouze okrajově při zajištění konzistence s již existujícími modely, případně ve snaze o postupné vytvoření datového kmene organizace. Zajištění zpřístupnění dat prostřednictvím PPDF není věnována žádná pozornost a od dodavatelů to není vyžadováno.

Graf č. 20



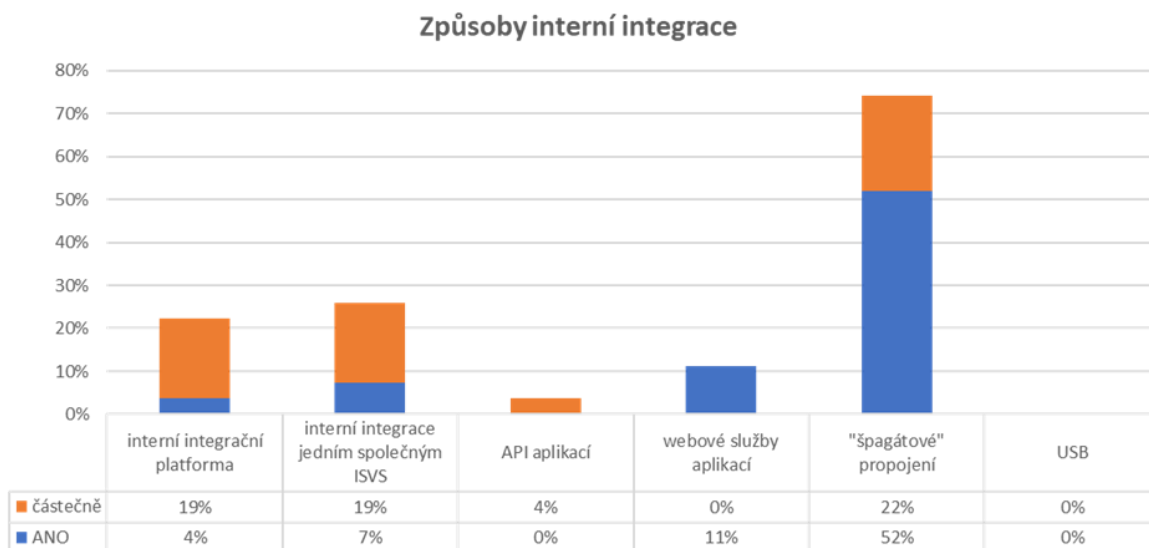
Zdroj: vlastní DIA

3.4.2 Datová integrace

Oblast datové integrace je rozdělena na integraci interních systémů a integraci s externími systémy.

Nejběžnějším způsobem zajištění integrace interních systémů je bohužel stále (až 74 %) jejich přímé propojení („špagátová“ integrace). Interní integrační platforma je spíše ojedinělá, stejně jako integrace jedním společným systémem (většinou v menších organizacích s několika agendami, nebo IS).

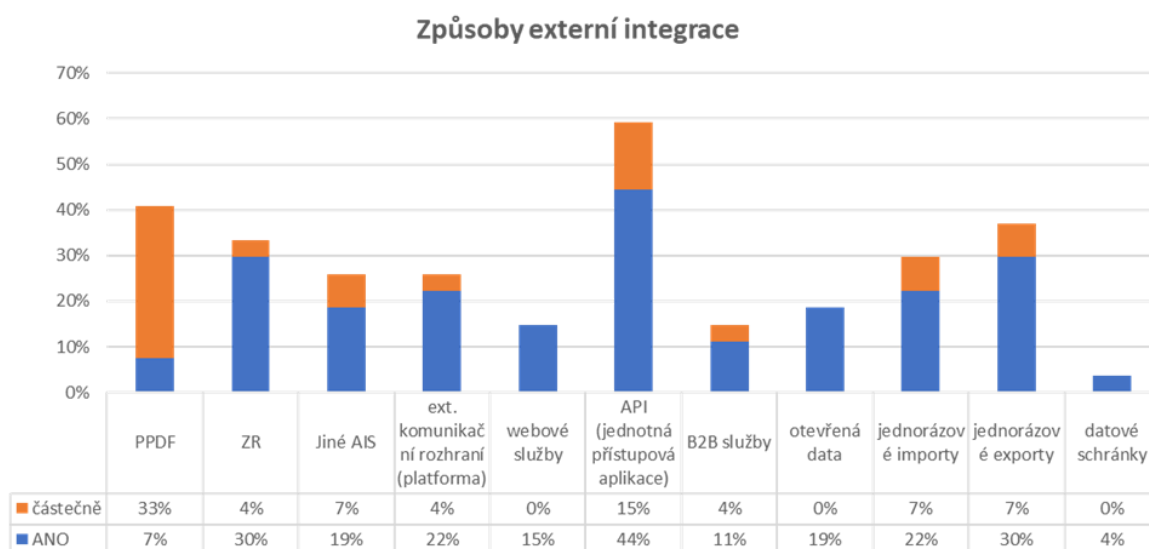
Graf č. 21



Zdroj: vlastní DIA

Integrace s externími systémy je mnohem bohatší, výjimkou nejsou ani integrace prostřednictvím datových schránek, jednorázovými importy a exporty, nebo otevřenými daty, což nelze považovat za digitalizaci veřejné správy v plném slova smyslu. Sdílení dat prostřednictvím PPDF není tak běžné, jak by se dalo očekávat, pojem VDF (Veřejný datový fond) je pro naprostou většinu úřadů neznámý, třebaže je specifikován v Informační koncepci ČR jako její specifický cíl 5.10. V každém případě platí, že integrace s externími systémy je skutečně pestrá z pohledu způsobů externího propojení systémů, a je z převážné části poplatná technologické nepřipravenosti (často i nemožnosti) k tomu vhodných rozhraní jednotlivých ISVS organizací. Z velké míry se také jedná o důsledek stávajících (nesystémových) způsobů interní integrace, nedostatečného povědomí o možnostech sdílení dat v architektuře veřejné správy ČR, a také nedostatečného (často formálního) plnění požadavků stanovených legislativou, koncepčními (strategickými) dokumenty a „centrem digitalizace“.

Graf č. 22



Zdroj: vlastní DIA

3.4.3 Základní zjištění v oblasti realizace datových řešení:

- Řešení projektů/změn zaměřených na dodání dat nebo datových výstupů jsou ze strany organizací řešeny téměř standardizovaným způsobem. Významným nedostatkem je vyžadovaný rozsah standardizovaného věcného obsahu, který uplatňuje pouze 11 % organizací. Obvykle je kladen důraz na prevenci datové kvality a dodávku (izolovaných) datových modelů, avšak bez striktně předepsaných standardů, a většího zajištění konzistence s již existujícími datovými modely, a bez větší snahy postupného vytváření datového kmene organizace.
- Nejběžnější způsob zajištění integrace interních systémů je stále (až 74 %) jejich přímé propojení („špagátová“ integrace). Společná interní integrační platforma je spíše ojedinělá, stejně jako integrace zajišťovaná jedním společným systémem.
- Integrace s externími systémy je bohatá na nejružnější (nesystémové) způsoby, výjimkou nejsou ani integrace prostřednictvím datových schránek, jednorázovými importy a exporty, nebo otevřenými daty. Nejčastějšími důvody nesystémového přístupu jsou technologická nepřipravenost ISVS pro nové způsoby integrace, nedostatečné povědomí a znalost nově vyžadovaných způsobů a nedůslednost plnění povinností stanovených legislativou či strategickými dokumenty.
- Sdílení dat prostřednictvím PPDF není běžné, úplné sdílení uvedlo pouze 7 % úřadů, částečné v některých oblastech 33 %.

3.5 Zajištění a využívání dat

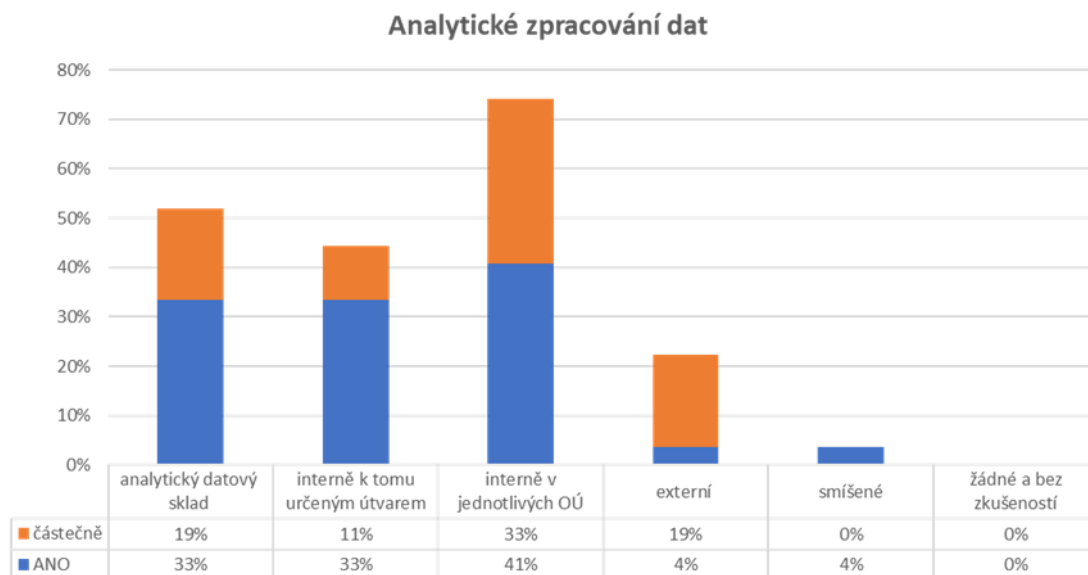
Oblasti zájmu rozhovorů:

- Využití dat (analýza dat / tvorba analytických výstupů, k tomu využívané nástroje).
- Kvalita dat (problémy s nedostatečnou kvalitou dat, způsob jejich řešení).
- Bezpečnost / řízení rizik (rizika/hrozby v oblasti dat, opatření k jejich snížení, klasifikace dat z hlediska citlivosti, archivace a výmaz dat).
- Provoz datových úložišť a řešení (správa datových úložišť a souvisejících komponent, využívání cloudových úložišť a nástrojů, řízení dostupnosti/výkonu).

3.5.1 Využití dat

Analytické zpracování dat a tvorba specifických výstupů zatím není až na výjimky (např. ČSÚ, ÚZIS apod) zcela usazenou disciplínou v organizacích veřejné správy. Ve 33 % organizací existují v různých podobách specializované útvary zaměřené na analytické zpracování dat, což také koresponduje s počtem organizací provozujících analytické datové sklady, stále ale zatím převažují analytické práce jako součást působení jednotlivých odborných útvarů.

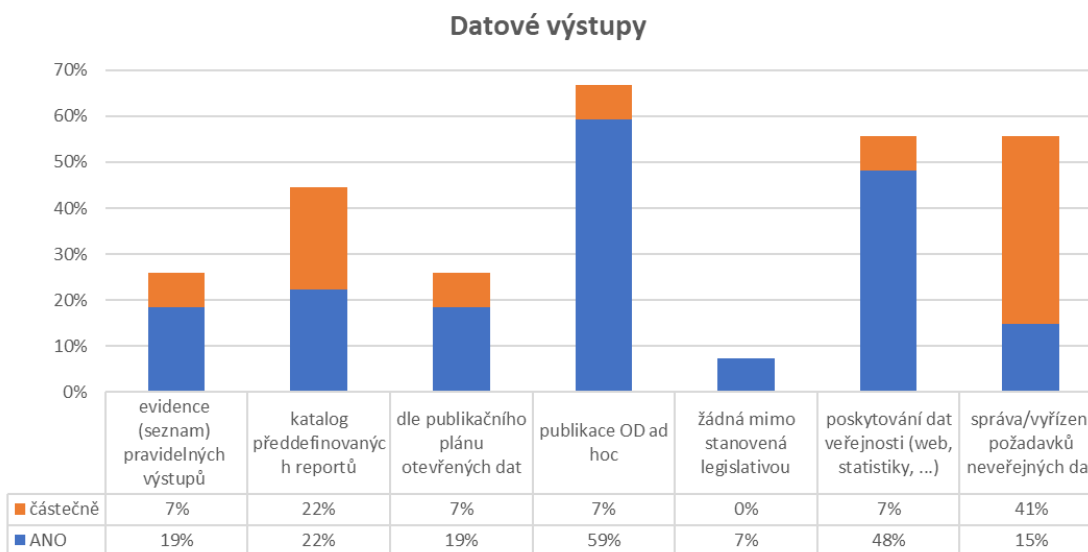
Graf č. 23



Zdroj: vlastní DIA

Datové výstupy jsou zatím poskytovány většinou na základě konkrétních požadavků, ale některé organizace již poskytují i standardizované výstupy v různých formátech (data, otevřená data, reporty, analýzy, neveřejná data), a také vedou a spravují katalog předdefinovaných reportů (datových výstupů). Součástí datových výstupů jsou dnes již také otevřená data, poskytování dat a informací veřejnosti na portálech organizací, a v několika případech i poskytování neveřejných údajů odborné veřejnosti. Plně organizovaný přístup deklarovalo asi 15 % organizací, přičemž jedna organizace přímo uváděla, že má k tomu vyhrazeny speciální prostory (Safe Room), a že speciální přípravu takových neveřejných dat má zpoplatněnu na základě vydaného ceníku. Asi 40 % organizací dále uvádělo, že nějakou zkušenost s poskytováním neveřejných údajů má také, ale nebyla schopna to doložit upřesňujícími informacemi (způsoby poskytnutí, forma, formáty, ...) s tím, že tyto aktivity se odehrávají na úrovni konkrétních věcných oblastí pod správou příslušných vlastníků dat. Nejedná se o centrálně spravovanou aktivitu.

Graf č. 24



Zdroj: vlastní DIA

3.5.2 Identifikované nástroje využívané pro přípravu a analýzu dat v úřadech:

V průběhu rozhorů byla zmíněna řada nástrojů, které organizace využívají ke své analytické práci. Výběr konkrétních nástrojů je s velkou pravděpodobností poplatný zkušenostem konkrétních analytických pracovníků s nástroji, dále skutečnosti, že součástí některých IS již jsou určité analytické nástroje, a především jejich cena a jejich případná volná dostupnost.

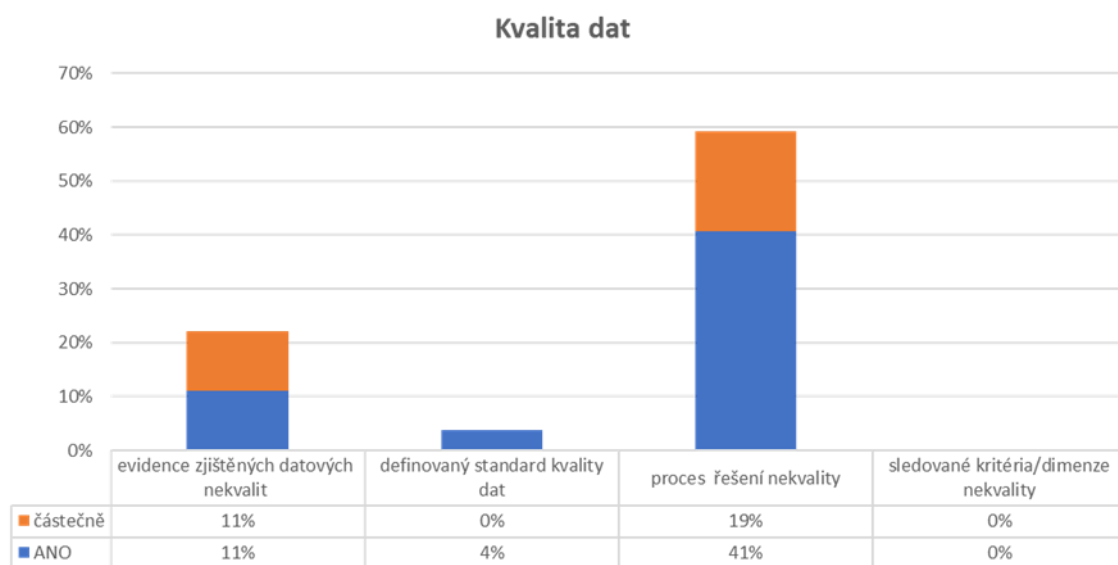
Seznam identifikovaných nástrojů:

Apache Superset, ARGIS, Attacama, Business Warehouse, Cognos, Excel, Google Data Studio, KNIME, MatLab, Oracle BI, PHP, PostGIS, PostgreSQL, Power BI, Power Query, programovací jazyk R, Python, SafeSW FME, SAP BusinessView, Spektrum, SPSS, SQL, TOAD, Tovek, vlastní řešení, SAS.

3.5.3 Kvalita dat

Z pohledu kvality dnes organizace neevidují závažnější problémy, většina nekvalit byla typicky spojena s historickými údaji, které jsou postupně čištěny a aktuálně je jejich kvalita již dobrá. Kvalita dat na vstupu je dnes řešena syntaktickými kontrolami již v samotných vstupních formulářích a většina organizací má pro řešení nekvality zavedeny procesní postupy. Definovaný standard kvality ve VS neexistuje, s čímž souvisí i nezjištěné informace o konkrétních kritériích sledované kvality. V rozhovorech pouze zaznělo, že z velké míry se sledují syntaktická pravidla zadávaných dat, kontrola vůči referenčním údajům a zejména věcná kvalita obsahu na úrovni věcně odpovědných pracovníků.

Graf č. 25

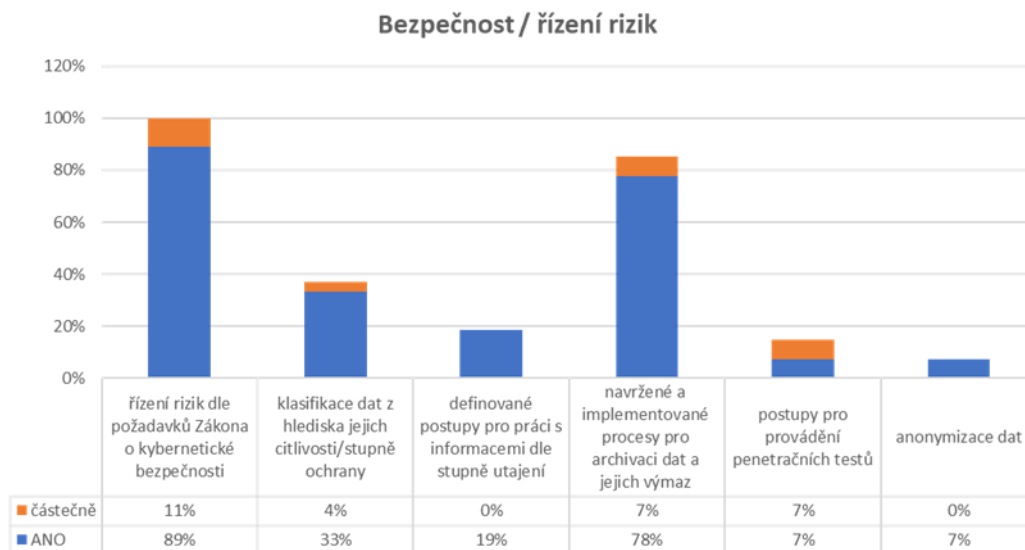


Zdroj: vlastní DIA

3.5.4 Bezpečnost / řízení rizik

Bezpečnost a řízení rizik se opírá o zákon o kybernetické bezpečnosti, naprostá většina organizací má k tomu implementovány související procesy. Penetrační testy občas provádí pouze několik organizací, k anonymizaci bylo obtížné získat bližší informace, až na pár výjimek (jako např. ČSÚ, ÚZIS – organizace zabývající se statistikami). Archivace dat na úrovni organizace je vesměs řešena uspokojivě, nejasnosti se týkají skartací, se kterými organizace zatím nemají velké zkušenosti. Ze strany Národního archivu zaznělo, že v případě archivace digitálního obsahu do NA, nejsou organizace často na ni předem připraveny a problémy řeší až následně. Mělo by spadat do povinností každé organizace mít již předem vše připravené, zejména pokud se jedná o archivaci digitálního obsahu.

Graf č. 26

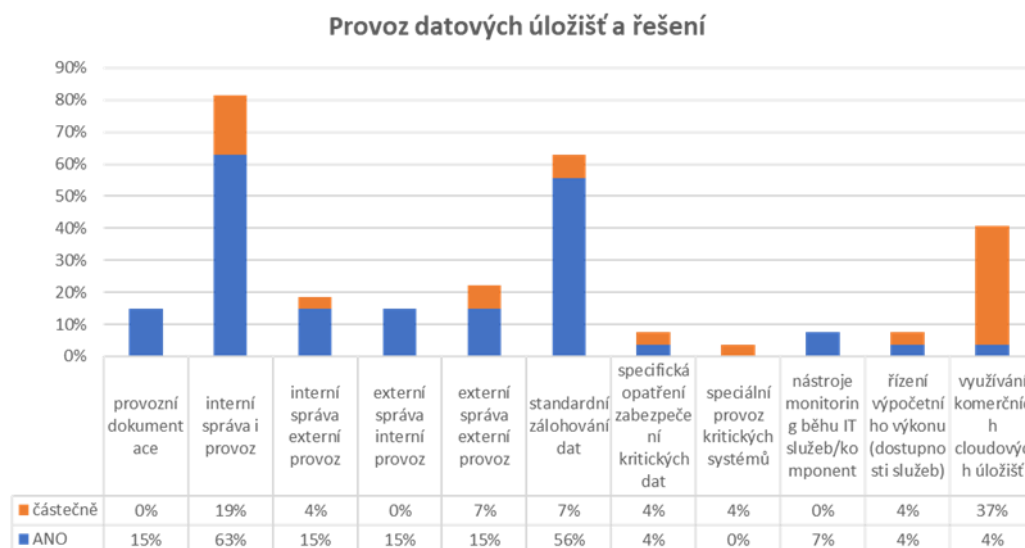


Zdroj: vlastní DIA

3.5.5 Provoz datových úložišť a řešení

Informace o provozu datových úložišť nevyžadují žádný komentář, jedná se pouze o zachycení aktuálního stavu. Zajímavou je pouze informace, že až v 80 % je správa i provoz datových úložišť zajišťován interně, na rozdíl od správy a provozu IS a agend, a že komerční cloudy se využívají minimálně a zejména v oblasti komunikace.

Graf č. 27



Zdroj: vlastní DIA

3.5.6 Základní zjištění v oblasti realizace datových řešení:

- Využívání dat a jejich analytickému zpracování se dnes již v nějaké formě věnuje většina úřadů, i když se zatím většinou jedná o jednotlivce v jednotlivých odborných útvech organizace.
- Formy a typy výstupů jsou různé, většinou vznikají na základě konkrétních požadavků, některé úřady a některé IS spravují i standardní předdefinované výstupy.
- Otevřená dat jsou publikována většinou způsobem ad hoc, řádný publikační plán má vypracováno pouze 19 % organizací.
- Veřejnosti poskytuje informace i na svých portálech asi polovina úřadů.
- S poskytováním neveřejných dat odborné veřejnosti má systematickou zkušenost jen pár úřadů, v ostatních úřadech k tomu občas také dochází, ale nedefinovaným systémem.
- Kvalita dat ani bezpečnost dat nepředstavuje pro úřady problém.
- Nedořešená je otázka skartace dat a přípravy dat k archivaci do Národního archívu. Problematická je zejména archivace digitálního obsahu a nepřipravenost organizací pro jeho předání k archivaci.
- Provoz datových úložišť a datových řešení se zdá být bezproblémový.

4 Metodika rozhovorů

4.1 Stanovení struktury vedených rozhovorů

Definované oblasti správy dat VS se staly základem vymezení rozsahu a obsahu vedených strukturovaných rozhovorů. Každá oblast byla tematicky rozdělena na několik částí.

Oblast	Téma	Příklady otázek
0. ŘÍZENÍ A ORGANIZACE SPRÁVY DAT	Organizace	Jaké role u vás v oblasti dat existují / hrají významnou roli? Kdo za data odpovídá a kdo o ně pečuje po věcné/odborné stránce a po technické stránce? Kdo u vás s daty aktivně pracuje (nad rámec standardního využívání dat v rámci výkonu agendy / poskytování služeb)?
	Standardizace	Máte role a odpovědnosti, procesy a pravidla související s péčí o data zakotveny v nějakých vnitřních řídicích dokumentech?
	Rozvoj kompetencí, spolupráce a podpora	Rozvíjíte nějak znalosti a dovednosti pracovníků v těchto rolích? Spolupracují tito lidé / tyto skupiny? Ví o sobě navzájem a sdílejí nějak své zkušenosti? (uvnitř úřadu nebo i s jinými úřady) Ví o nich vedení úřadu, vnímá a podporuje nějak jejich práci?
1. STANOVENÍ DATOVÝCH POTŘEB	Prioritizace oblastí	Rozlišujete v praxi nějaké věcné oblasti dat a jsou pro vás některé oblasti důležitější než jiné? Proč? Pokud ano, má to nějaký vliv na způsob, jak o data v jednotlivých oblastech pečujete? Máte přehled o tom kdo, jaká data, kde, v jaké podobě a proč aktuálně využívá a jak jsou pro něj důležitá?
	Řízení potřeb	Máte přehled o tom kdo, jaká data, kde, v jaké podobě a proč (nově) potřebuje? Od koho a jak k vám přichází a jak řídíte požadavky související s daty (např. na nová data, sdílení dat, vytvoření reportů)?
	Využití architektury pro plánování změn	Využíváte při plánování změn a pro podporu rozhodování o nich (enterprise) architekturu? Máte nějak zachycen současný stav a navržen cílový stav v oblasti dat? Pokud ano, zohledňujete při návrhu cílového stavu požadavky související s daty?
2. POPIS DAT A DATOVÝCH ŘEŠENÍ	Popis dat	Jakým způsobem máte popsána data, která spravujete? Co vše zahrnuje váš popis dat? Pokrývá veškerá vaše data? Udržujete jej aktuální? Jakým způsobem a komu je existující popis dat přístupný? Kdo s ním aktivně pracuje a kdy?

	Evidence v RPP	Mají vámi vykonávané agendy evidovány nějaké agendové údaje v Registru práv a povinností (RPP)? Pokud ano, odpovídá evidence vašich agendových údajů v RPP aktuálnímu stavu dat a legislativy?
	Modelování dat a datových řešení	Využíváte nějaké datové modely? Vytváříte je interně? Jaké využíváte modelovací/zobrazovací nástroje? Máte nějak zachyceno, jak (a jaká) data tečou uvnitř vašeho úřadu a jak si data vyměňujete s okolím? Jsou z tohoto zachycení patrné všechny zapojené komponenty? (informační systémy, datová úložiště, datové sady, rozhraní, nástroje pro ETL, analýzu a prezentaci dat) Kdo s těmito informacemi pracuje a kdy?
	Kategorizace dat	Rozlišujete nějaké kategorie dat? Pokud ano, podle čeho? Udržujete si přehled o tom, jaké využíváte číselníky, odkud pochází a kdo/jak je udržuje aktuální? Udržujete si přehled o tom, jaká data přebíráte pro výkon agend ze základních registrů nebo od jiných úřadů? Pokud využíváte data o klientech vašich služeb (nebo i jiná) ze základních registrů nebo od jiných úřadů napříč více agendami/systémy, přebírá a drží si je každá agenda/systém samostatně?
3. REALIZACE DATOVÝCH ŘEŠENÍ	Datové aspekty řešení změn informačních systémů	Hledáte při návrhu změn IS aktivně možnosti získání a přepoužití dat již existujících ve vašem úřadu nebo jinde ve veřejné správě? Zajišťujete v rámci veřejných zakázek a smluv s dodavateli svou datovou suverenitu a nezávislost (tzn. neomezený přístup k datům IS)? Co vše z pohledu dat je součástí dodávek od externích dodavatelů? Jsou standardní součástí popisy a modely dat a datových řešení? Zvažujete při návrhu IS zpřístupnění relevantních dat v rámci propojeného nebo veřejného datového fondu? Jak rozhodujete o tom, která data zpřístupnit a jakou formou?
	Integrace	Jaká data si vyměňujete (sdílíte)... - interně mezi agendami organizace - externě mezi vaším a jinými úřady Jakými způsoby výměna probíhá? Prostřednictvím propojeného datového fondu, nebo mimo něj?
4. ZAJIŠTĚNÍ A VYUŽÍVÁNÍ DAT	Využití dat	Jaké máte zkušenosti s analýzou dat a tvorbou datových výstupů? Na jaké problémy nejčastěji narážíte při analytickém zpracování dat? Jste schopni dohledat a získat data potřebná pro analýzu? Jaké nástroje a IT komponenty se u vás pro zpracování a analýzu dat využívají? Jsou vyhovující pro vaše potřeby? Zpřístupňujete nějaká data odborné či široké veřejnosti na vyžádání nebo veřejně ve formě otevřených dat? Pokud ano, získali jste z takového sdílení nějaké informace zpětně užitečné pro váš úřad?
	Kvalita dat	Setkali jste se s tím, že nějaká data nemůžete použít (nebo jen obtížně) pro zamýšlený účel, protože nejsou dostatečně kvalitní (např. obsahují zjevné chyby, duplicitu, jsou nesrozumitelná)? Co bývá příčinou? Máte nějakou evidenci zjištěných problémů s kvalitou vašich dat? Máte zavedený proces (postupy) pro případné reklamace dat a odhalování/řešení příčin problémů s daty? Máte a udržujete si přehled o tom, jak jste na tom s kvalitou dat v jednotlivých oblastech?
	Bezpečnost / řízení rizik	Vyhodnocujete a řešíte rizika/hrozby v oblasti dat? Jakým způsobem? Jaká jsou nejčastější a nejzávažnější rizika, která řešíte? Jaká máte zavedena opatření k snížení rizik/zabezpečení dat? Používáte klasifikaci dat z hlediska jejich citlivosti/stupně ochrany? Jak klasifikaci v praxi vynucujete a používáte? Máte zavedeny procesy a pravidla pro archivaci dat a pro jejich výmaz po uplynutí doby, po kterou je můžete držet? Jaké?
	Provoz datových úložišť a řešení	Jaká využíváte datová úložiště a kdo je spravuje? Máte přehled o tom / je nějak zdokumentováno, co všechno jejich správa zahrnuje? Využíváte nebo zvažujete možnost využívání komerčních cloudových úložišť, případně cloudových nástrojů pro zpracování dat?

		Řídíte u svých datových úložišť dostupnost (výkon)? Děláte to v závislosti na tom, jaká data v nich leží? Jaké klíčové IT komponenty kromě primárních informačních systémů využíváte pro řešení datových potřeb úřadu (datová úložiště, nástroje pro ETL, analýzu a prezentaci dat)? Které z nich spravujete interně kapacitami svého IT?
--	--	--

4.2 Šablona zápisu výsledků rozhorů

Zápis rozhovoru s každou organizací byl zaznamenán v jednotné struktuře pomocí šablony uvedené v [Příloze 1](#).

4.3 Dotazované organizace

Rozhovorů se účastnilo 34 úřadů VS, počet provedených rozhovorů s každým konkrétním úřadem byl různý, závisel na nominaci účastníků ze strany úřadu a její velikosti. Doba trvání rozhovoru byla stanovena na 120 minut. Záznamy všech provedených rozhovorů v jednom úřadu byly seskupeny do jednoho výsledného zápisu, který byl následně úřadem také validován.

Ministerstva a ústřední správní úřady (kategorie OVM: KO 129):

- MINISTERSTVO PRO MÍSTNÍ ROZVOJ
- MINISTERSTVO ŽIVOTNÍHO PROSTŘEDÍ
- MINISTERSTVO PRÁCE A SOCIÁLNÍCH VĚCÍ
- MINISTERSTVO VNITRA
- MINISTERSTVO ZAHRANIČNÍCH VĚCÍ
- MINISTERSTVO OBRANY
- MINISTERSTVO PRŮMYSLU A OBCHODU
- MINISTERSTVO ZDRAVOTNICTVÍ
- MINISTERSTVO SPRÁVEDLNOSTI
- MINISTERSTVO FINANČÍ
- MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
- MINISTERSTVO DOPRAVY
- MINISTERSTVO ZEMĚDĚLSTVÍ
- MINISTERSTVO KULTURY
- ČESKÝ STATISTICKÝ ÚŘAD
- ČESKÝ ÚŘAD ZEMĚMĚŘIČSKÝ A KATASTRÁLNÍ
- ČESKÝ TELEKOMUNIKAČNÍ ÚŘAD
- ENERGETICKÝ REGULAČNÍ ÚŘAD
- ÚŘAD PRŮMYSLOVÉHO VLASTNICTVÍ
- ÚŘAD VLÁDY ČESKÉ REPUBLIKY

Organizační složky státu (kategorie OVM: KO 191):

- STÁTNÍ ÚSTAV PRO KONTROLU LÉČIV
- ÚSTAV ZDRAVOTNICKÝCH INFORMACÍ A STATISTIKY
- GENERÁLNÍ ŘEDITELSTVÍ CEL
- GENERÁLNÍ FINANČNÍ ŘEDITELSTVÍ
- ÚŘAD PRO ZASTUPOVÁNÍ STÁTU VE VĚCECH MAJETKOVÝCH

Vybrané další úřady (kategorie OVM: KO 128, 199, 401):

- ČESKÁ SPRÁVA SOCIÁLNÍHO ZABEZPEČENÍ
- ČESKÁ INFORMAČNÍ AGENTURA ŽIVOTNÍHO PROSTŘEDÍ (CENIA)
- NÁRODNÍ ARCHIV
- STÁTNÍ ZEMĚDĚLSKÝ INTERVENČNÍ FOND
- STÁTNÍ FOND DOPRAVNÍ INFRASTRUKTURY
- TECHNOLOGICKÁ AGENTURA ČR
- GRANTOVÁ AGENTURA ČR
- SPRÁVA ZÁKLADNÍCH REGISTRŮ – DIGITÁLNÍ A INFORMAČNÍ AGENTURA
- ČESKÝ STATISTICKÝ ÚŘAD

5 Přílohy

5.1 Příloha č. 1

Zápis z rozhovoru o správě dat úřadu

Dne:			
Adresa:			
Zástupci úřadu:		Zástupci DIA:	

Oblast / téma a otázky	Zaznělo v rozhovoru	Interní doplnění
Otázky pro oblast 0. ŘÍZENÍ A ORGANIZACE SPRÁVY DAT téma Organizace - Máte definované a obsazené nějaké role, které se primárně zaměřují na práci s daty? Jaké? - Kdo u vás (manažersky) odpovídá za data (na úrovni jednotlivých oblastí / celkově)? - Kdo u vás reálně o data pečuje (po věcné/odborné a po technické stránce)? V čem (v jakých aktivitách) tato péče spočívá? - Kdo u vás s daty aktivně pracuje (nad rámec standardního využívání dat v rámci výkonu agend / poskytování služeb)? - Znáte a máte obsazené role typu: datový analytik, datový architekt, datový stevard, administrátor databáze, Chief Data Officer, Data Protection Officer? - Využíváte v rámci každodenní péče o data (mimo projekty a změnové aktivity) externí dodavatele? (např. při průběžné tvorbě/správě datových modelů) Pokud ano, jsou jejich poznatky a znalosti v oblasti dat předávány interním pracovníkům? Jakým způsobem?		
téma Standardizace - Máte tyto role a odpovědnosti zakotveny v nějakých vnitřních řídicích dokumentech? Jakých? (a můžeme je pro inspiraci vidět?) - Máte zavedenu nějakou strategii (nebo nějaká doporučení) pro práci s daty (datové role, procesy, vnitřní standardy, nástroje, metriky a otázky etiky práce s daty)?		
téma Rozvoj kompetencí, spolupráce a podpora Rozvíjíte znalosti a dovednosti pracovníků v oblasti datové problematiky (např. znalost a dovednost modelování / definice významu dat)? Spolupracují tyto lidé / tyto skupiny uvnitř úřadu? Ví o sobě navzájem a sdílejí nějak své zkušenosti? Zapojují se do nějakých skupin či aktivit mimo OVS tematicky orientovaných na data? (např. skupina datových analytiků z VS)		



Ví o nich vedení úřadu, vnímá a podporuje nějak jejich práci? Umožňuje jim rozvíjet specifické znalosti a dovednosti související s daty?		
Otázky pro oblast 1. STANOVENÍ DATOVÝCH POTŘEB téma Prioritizace oblastí - Rozlišujete v praxi nějaké věcné oblasti dat, která spravujete? Na základě čeho je rozlišujete? - Pokud ano, jsou pro vás některé oblasti důležitější než jiné? Proč? - Pokud ano, má to nějaký vliv na způsob, jak o ně pečujete? - Máte přehled o tom kdo, jaká data, kde, v jaké podobě a proč aktuálně využívá a jak jsou pro něj důležité? Udržujete ho nějak (systematicky)?		
téma Řízení potřeb - Máte přehled o tom kdo, jaká data, kde, v jaké podobě a proč (nově) potřebuje? Udržujete ho nějak (systematicky)? Využíváte spravovaná data agend k vlastním účelům mimo výkon agend (např. pro analýzy)? Vznikají při tom požadavky na doplňující data a jejich následnou správu? - Jak k vám přichází a jak řídíte požadavky související s daty, např.: - požadavky na dodání datových/analytických výstupů (např. na analytické výstupy/reporty pro vedení, nadřízenou organizaci), - požadavky na sdílení dat úřadu, - vnitřní: na výměnu dat mezi agendami uvnitř úřadu, vnější: na výměnu dat mezi úřady, úřadem a podniky, úřadem a občany, úřadem a EU/světlem. - Jak prioritizujete takové požadavky / rozhodujete o tom, které budete řešit a které ne (nebo později)?		
téma Využití architektury pro plánování změn Využíváte při plánování změn a pro podporu rozhodování o nich (enterprise) architekturu ? - Máte nějak zachycený současný stav a navržený cílový stav v oblasti dat? K jednomu budoucímu okamžiku, nebo ve více okamžicích? (tzn. nejen cílový, ale i tranzitní stav/-y). - Pokud ano, zohledňujete systematicky při návrhu cílového stavu prioritizované požadavky související s daty?		
Otázky pro oblast 2. POPIS DAT A DATOVÝCH ŘEŠENÍ téma Popis dat (metadata) - Kde a jakým způsobem máte popsaná data, která spravujete? Co vše zahrnuje váš popis dat? Pokrývá veškerá vaše data? Ve stejné nebo různé míře detailu? V jaké formě? Udržujete jej aktuální? - Máte identifikované pojmy / subjekty a objekty, o kterých vypovídají vaše data? Pokud ano, máte na jednotlivé pojmy navázané i informace o tom, z jakých zdrojových systémů pochází údaje/data o nich a případně v jakých datových rozhraních (nebo výstupech) jsou zpřístupňována uživatelům? [= otázka na existenci lokálního katalogu dat / slovníku pojmů / konceptuálního datového modelu].		



• Máte a udržujete nějakou definici významu pojmů používaných v úřadu a/nebo významu jednotlivých údajů? • Jaké využíváte pro popis dat nástroje? • Jakým způsobem a komu je existující popis dat přístupný? Kdo s ním aktivně pracuje a kdy? • Jak k vám přichází a jak řídíte požadavky související s popisem dat a datových řešení, např. • požadavky na to, která data, v jaké míře detailu a jakou formou a proč mají být popsána nebo namodelována, • dtto pro datové toky uvnitř i vně úřadu. • Držíte se při popisování dat nějakých standardů (a jakých)? Jakým způsobem zajišťujete kvalitu popisu dat?		
téma Evidence v RPP • Mají vámi vykonávané agendy evidované nějaké agendové údaje v RPP? Pokud ano, v jakém stavu je evidence vašich agendových údajů v RPP? • Pokud v RPP evidujete údaje, je to včetně jejich technické struktury (TSÚA)? • Kdo evidenci provádí (spolupráce vykonavatel-ohlašovatel agendy) a zajišťuje její soulad s aktuálním stavem dat a s aktuální stavem legislativy? Jakým způsobem je zajišťována evidence údajů v RPP v případech, kdy agendu, nebo její část, zajišťuje (vykonává) podřízená organizace nebo jakýkoliv jiný subjekt (státní podnik, ...)? • Je stávající podpora z "centra" v oblasti dostupných standardů a metodik dostatečná?		
téma Modelování dat a datových řešení • Máte nějaké datové modely? Vytváříte je interně? Držíte se při tvorbě datových modelů nějakých standardů (a jakých)? • Jaké využíváte modelovací nástroje? Kde jsou modely uloženy? Udržují se aktuální? Kdo s vytvořenými datovými modely následně pracuje a kdy? • Máte nějak zachyceno, jak (a jaká) data tečou uvnitř i vně vašeho úřadu? Jsou z tohoto zachycení patrné všechny komponenty, kterými data prochází (primární systémy, datová úložiště, datové sady a rozhraní, nástroje pro ETL, analýzu a prezentaci dat)? Kdo s těmito informacemi pracuje a kdy? • Máte nějaký plán, který by říkal, jaké modely mají vzniknout a kdy?		
téma Kategorizace dat (referenční a kmenová data, nestrukturovaná data) • Jakým způsobem rozlišujete různé kategorie dat? V čem je vaše správa jednotlivých kategorií dat specifická? Rozlišujete data na veřejná a neveřejná? Pokud ano, podle čeho? • Udržujete si přehled o tom, jaké využíváte číselníky, odkud pochází a kdo/jak je udržuje aktuální? V jaké podobě? • Snažíte se systematicky sjednocovat existující číselníky a používat „pro jednu věc“ jeden číselník napříč agendami? • Udržujete si přehled o tom, jaká data přebíráte pro výkon agend ze základních registrů nebo od jiných úřadů? V jaké podobě? Jakým způsobem zajišťujete, aby tato data zůstala aktuální?		



• Pokud využíváte data o klientech vašich služeb (nebo i jiná) ze základních registrů nebo od jiných úřadů napříč více agendami, přebírá a drží si je každá agenda samostatně? • Jakým způsobem a v jakém rozsahu spravujete nestrukturovaná data (dokumenty, e-maily, audiovizuální soubory atd.)? Máte zavedeny nějaké procesy zabývající se např. jejich potřebností, klasifikací, archivací či výmazem?		
Otázky pro oblast 3. REALIZACE DATOVÝCH ŘEŠENÍ téma Datové aspekty řešených změn IS • Věnujete v projektech / změnových aktivitách zaměřených na rozvoj IS pozornost dopadům řešených změn na data a jejich správu? Pokud ano, v jakých fázích řešení změny a co konkrétně děláte? • Hledáte při návrhu změn IS aktivně možnosti získání a přepoužití dat již existujících ve vašem úřadu nebo jinde ve VS? • Zajišťujete v rámci veřejných zakázek a smluv s dodavateli svou datovou suverenitu a nezávislost (tzn. neustálý přístup ke všem datům IS ve své správě, a to i po ukončení jejich životního cyklu, transparentní architekturu a dokumentaci všech dat uložených v IS)? • Co vše z pohledu dat je součástí dodávek od externích dodavatelů? Jsou standardní součástí popisy a modely dat a datových řešení? V jaké podobě? Jak je zajištěna jejich konzistence (soulad) s již existujícími popisy a modely? Vyžadujete po dodavatelích dodržení standardů? Kdo, kdy a jak akceptuje modely vytvořené dodavateli? • Zapojujete systematicky do řešení změn IS datové role? Jaké a jaká je jejich odpovědnost? (hlavně z věcné stránky) • Zohledňujete při návrhu IS (zejména už. rozhraní – typicky formulářů) preventivní aspekt datové kvality? (tzn. maximální možné vynucení validních údajů na vstupu, omezení na výběr z definovaných hodnot / číselníků, logické kontroly konzistence obsahu apod.) • Zvažujete při návrhu IS zpřístupnění relevantních dat v rámci propojeného nebo veřejného datového fondu? Zajišťujete jejich technické připojení do propojeného datového fondu nebo přípravu otevřených dat? Jak rozhodujete o tom, která data takto zpřístupnit a jakou formou?		
téma Datová integrace • Zvažujete při návrhu IS zpřístupnění relevantních dat v rámci propojeného nebo veřejného datového fondu? Zajišťujete jejich technické připojení do propojeného datového fondu nebo přípravu otevřených dat? Jak rozhodujete o tom, která data takto zpřístupnit a jakou formou? • Jakým způsobem zajišťujete výměnu (sdílení = poskytnutí i získání) dat (integrační platforma / jiné způsoby výměny) • Interně mezi agendami organizace: • pro výkon agend, • mimo výkon agend, • Externě mezi vaším a jinými úřady: • pro výkon agend,		



• mimo výkon agend? • Prostřednictvím PPDF? Mimo PPDF a pokud ano, jakým způsobem? • Zvažujete při plánování změn/projektů vytvoření společného datového kmene úřadu, který by obsahoval data využívaná více agendami úřadu (např. klientská data)?		
Otázky pro oblast 4. ZAJIŠTĚNÍ A VYUŽÍVÁNÍ DAT téma Využití dat • Máte nějaké zkušenosti s analýzou dat a tvorbou datových výstupů? Máte v úřadu někoho, kdo se tím systematicky zabývá? Pokud ano, o jaké výstupy jde a jaký mají účel? Jsou vzniklé datové výstupy někde evidované a dohledatelné? • Na jaké problémy nejčastěji narážíte při analytickém zpracování dat? Jste schopni dohledat data potřebná pro analýzu? Jak postupujete, pokud ne? • Jaké nástroje a IT komponenty se u vás pro zpracování a analýzu dat využívají? Jsou vyhovující pro vaše potřeby? • Zpřístupnění veřejnosti: Myslíte si, že některá z vámi spravovaných dat včetně neveřejných mohou být zajímavá pro odbornou veřejnost? (např. analýzy výzkumníků z akademické sféry) Zpřístupňujete už nějaká svá data pro tyto účely nebo veřejně ve formě otevřených dat? Pokud ano, získali jste z takového sdílení nějaké informace zpětně užitečné pro váš úřad? • Nabízíte nějakou možnost zpřístupnění vašich dat včetně neveřejných např. přímo ve vašich prostorách a pod dohledem ("safe room")?		
téma Kvalita dat [dvě stránky kvality v širším pojetí: kvalita popisu dat a dohledatelnost + kvalita obsahu dat] • Narazili jste někdy na problémy s daty, tzn. s tím, že nějaká data nemůžete použít (nebo jen obtížně) pro zamýšlený účel, protože nejsou dostatečně kvalitní (např. obsahují zjevné chyby, duplicity, jsou nesrozumitelná)? Co bývá příčinou? • Máte nějakou evidenci zjištěných problémů s kvalitou dat? Máte zavedený proces (postupy) pro případné reklamace dat a odhalování/řešení příčin problémů s daty? • Máte a udržujete si přehled o tom, jak jste na tom s kvalitou dat v jednotlivých oblastech? Pokud ano, využíváte při tom nějaké automatizované nástroje? • PRO POKROČILÉ: Pracujete systematicky s nějakými kritérii/dimenzemi kvality dat? Pokud ano, s jakými a jak?		
téma Bezpečnost / řízení rizik • Vyhodnocujete a řešíte rizika/hrozby v oblasti dat? Jakým způsobem? Jaká jsou nejčastější rizika, která řešíte? • Jaká máte zavedena opatření k snížení rizik/zabezpečení dat? • Používáte klasifikaci dat z hlediska jejich citlivosti/stupně ochrany? Jak to v praxi funguje a na co má tato klasifikace vliv? • Máte nějaké zkušenosti s anonymizací dat? Máte anonymizovaná data v neprodučních prostředích?		



[živ. cyklus] Máte zaveden procesy a pravidla pro archivaci dat a pro jejich výmaz po uplynutí doby, po kterou je můžete držet? Jaké?		
téma Provoz datových úložišť a řešení - Kdo a jak spravuje vaše datová úložiště? Máte přehled o tom, co všechno v rámci správy dělá? Existuje nějaká provozní a/nebo procesní dokumentace? - Máte nějak odstupňována opatření k zabezpečení dat podle míry jejich kritičnosti pro výkon agend a fungování úřadu? - Máte stanoveny a dodržujete důsledně procesy pro zálohování kritických dat? Kde/jak tato data zálohujete? - Využíváte nebo zvažujete možnost využívání komerčních cloudových úložišť? Pokud ano, jak se rozhodujete (na základě jakých hledisek) o tom, která data je možné v nich držet? - Řídíte u svých datových úložišť dostupnost (výkon)? Děláte to v závislosti na tom, jaká data v nich jsou uložena / v jaké fázi životního cyklu se nachází (tzn. máte např. archivovaná data na méně dostupných úložištích)? - Pokud využíváte data mimo výkon agend / provozujete nějaká čistě datová řešení (jejichž primárním nebo jediným cílem je dodání dat nebo datových výstupů), je pro vás jejich provozování v něčem specifické? - Jaké klíčové IT komponenty kromě primárních IS využíváte pro řešení datových potřeb úřadu (datová úložiště, datové sady a rozhraní, nástroje pro ETL, analýzu a prezentaci dat)? Které z nich spravujete interně kapacitami svého IT?		
+ BONUS (na úplný závěr celého dotazování): Máte nějaká doporučení pro nás (z centra) nebo pro ostatní úřady nějaké své " dobré praktiky " vyplývající z vašich zkušeností v oblasti správy a řízení dat?		