

Metodika poskytování dat ve veřejném datovém fondu

Vytvořeno v rámci projektu

Rozvoj datových politik v oblasti zlepšování kvality
a interoperability dat veřejné správy

CZ.03.4.74/0.0/0.0/15_025/0013983

Klíčová aktivita: 03 Návrhy a realizace opatření pro zlepšování kvality otevřených dat
Indikátor: 8 05 00 Počet napsaných a zveřejněných analytických a strategických dokumentů (vč. evaluačních)



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost



MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY

Verze výstupu: 01



Obsah

1. Úvod k metodice a její charakteristika	5
2. Zachycení VDF v legislativě a jeho základní principy	6
2.1 Legislativní pozadí veřejného datového fondu	6
2.2 Principy veřejného datového fondu	7
3. Logické komponenty VDF a jejich význam	8
3.1 Zastřešující pohled na logickou strukturu VDF	9
3.2 Stručná charakteristika společných komponent VDF	9
3.3 Stručná charakteristika komponent VDF jednotlivých agend (OVM)	10
4. Komponenty VDF a jejich vzájemné vazby	12
4.1 Provázání komponent VDF k zajištění evidence údajů v RPP a NKOD	12
4.2 Provázání komponent VDF při čtení údajů z VDF	14
4.3 Provázání komponent VDF při publikaci údajů do VDF	15
5. Popis hlavních procesů VDF	18
5.1. Publikace veřejných údajů	18
5.1.1 Pravidla pro evidenci veřejných údajů v RPP	18
5.1.2 Pravidla pro údaje zpřístupněné veřejným datovým fondem	18
5.1.3 Popis procesu publikace veřejných údajů	19
5.2 Publikace veřejných číselníků	20
5.2.1 Pravidla pro evidenci veřejných číselníků	20
5.2.2 Pravidla pro publikaci veřejných číselníků do VDF	21
5.2.3 Popis procesu publikace číselníků	22
5.3 Čtení veřejných údajů a veřejných číselníků z VDF	23
6. Specifikace jednotlivých komponent VDF	24
6.1 Společné komponenty VDF	24
6.1.1 RPP registr práv a povinností	24
6.1.2 Systém pro zpřístupňování údajů z RPP	25
6.1.3 Úložiště číselníků a evidence RPP a datových sad RPP	26
6.1.4 Lokální katalog RPP	27
6.1.5 Katalog uživatelů	27
6.1.6 Notifikační hub	27



6.1.7 NKOD – Národní katalog otevřených dat	28
6.2 Komponenty jednotlivých agend (AIS)	28
6.2.1 Systém pro zpřístupnění údajů.....	29
6.2.2 Systém pro čtení veřejných údajů z VDF	30
6.2.3 Úložiště datových sad	30
6.2.4 LKOD – Lokální katalog otevřených dat.....	31
6.3 Technické zajištění přístupu k VDF pro výkon VS	32
6.3.1 Referenční rozhraní veřejné správy.....	32
6.3.2 ISGOD – součást referenčního rozhraní	32
7. Příprava datových sad k publikaci	34
7.1 Legislativní vymezení oblastí a typů informací k publikaci do VDF	34
7.2 Identifikace konkrétních agend a konkrétních množin údajů	35
7.3 Rozdělení a seskupení dat k publikaci do ucelených datových sad.	36
7.4 Příprava a použití prvků repozitáře otevřených formálních norem pro publikaci údajů	39
7.5 Příprava obsahu datové sady ve vybraném základním otevřeném formátu	43
8. Externí odkazy a standardy	44
8.1 Metodiky, pravidla, existující OFN.....	44
8.2 Komunikační standardy komponent	44
9. Zdroje a terminologie	44



1. Úvod k metodice a její charakteristika

Metodika poskytování dat prostřednictvím veřejného datového fondu (VDF) je poměrně specifická z pohledu pokrývaného věcného rozsahu. Tato skutečnost pramení již ze samotné definice VDF v Informační koncepci ČR v dílčím cíli 5.10. V něm se uvádí, že VDF je metoda sdílení veřejných informací v prostředí VS, a jelikož se jedná o metodu, je nutné pro její konkretizaci a praktické zabezpečení vytvořit poměrně rozsáhlý systém, tvořený:

- provázanými již existujícími, ale i neexistujícími prvky architektury VS ČR,
- všemi veřejnými údaji VS publikovanými ve formátu otevřených dat,
- pravidly a postupy vytvářejícími a zabezpečujícími metodu sdílení údajů v duchu její definice,
- standardy zajišťujícími plnou interoperabilitu sdílených údajů napříč VS.

Z důvodu komplexnosti takového systému vyžaduje plnohodnotné a bezproblémové zavedení metody VDF do výkonu veřejné správy úplné pochopení:

- role VDF v architektuře VS,
- fungování VDF jako systému na základě vazeb mezi jeho jednotlivými komponentami,
- znalost principů VDF,
- znalost funkcionalit jednotlivých komponent VDF.

Chápání role jednotlivých komponent všemi zainteresovanými subjekty je velmi důležité, neboť každý orgán veřejné moci (OVM) se v průběhu postupné implementace VDF stane jeho důležitým spolutvůrcem.

Metodika z uvedených důvodů věnuje pozornost také znalostní stránce a má následující členění:

Zaměření na vysvětlení VDF:

- Kapitola 2. - Zachycení VDF ve stávající legislativě, přiblížení základních principů zahrnutých do VDF.
- Kapitola 3. - Představení komponent VDF a jejich rozdělení na komponenty společné (zajišťované MV ČR) a komponenty jednotlivých agend (zajišťované OVM).
- Kapitola 4. - Popis vzájemných vazeb mezi komponentami VDF.
- Kapitola 5. - Komplexní popis tří hlavních procesů VDF:
 - publikace veřejných údajů,
 - publikace veřejných číselníků,
 - čtení veřejných údajů a veřejných číselníků z VDF.

Zaměření na realizaci VDF:

- Kapitola 6. - Specifikace jednotlivých komponent VDF:
 - společné komponenty VDF (zajišťované MV ČR),
 - komponenty jednotlivých agend, AIS (zajišťované OVM),
 - technické zajištění přístupu k VDF pro výkon VS.
- Kapitola 7. - Příprava datových sad k publikaci.
- Kapitola 8. - Externí odkazy na potřebné standardy.

Metodika odpovídá a zachycuje stávající stav rozpracování VDF. S postupem implementačních prací bude metodika doplňována a upřesňována v souladu s rozšiřováním funkcionalit VDF, typů publikovaných veřejných údajů a rozšiřováním publikačních formátů otevřených dat.

Metodika je určena:

- všem OVM, zejména rolím a pracovním místům:
 - gestor agendy,
 - ohlašovatel agendy,
 - správce ISVS,



- poskytovatel číselníku,

2. Zachycení VDF v legislativě a jeho základní principy

Veřejný datový fond (VDF) je definován v [Informační koncepci ČR](#) (IKČR) jako [dílčí cíl 5.10](#), je součástí eGovernmentu VS ČR a jeho poslání a role je definována následujícím způsobem:

“Veřejný datový fond tvořený publikovanými veřejnými údaji veřejné správy je základní metodou pro sdílení veřejných informací mezi veřejnoprávními subjekty navzájem i pro sdílení veřejných údajů mezi veřejnoprávní a soukromoprávní sférou v ČR. Veřejný datový fond se od pouhé publikace automatizovaně čitelných otevřených dat posune též k publikaci právně závazných, platných a pravidelně aktualizovaných datových sad s jasně definovanou zodpovědností OVS za takové sady.”

2.1 Legislativní pozadí veřejného datového fondu

VDF není v legislativě jako pojem explicitně zmíněn. Legislativa pouze zavádí pojmy důležité v rámci VDF a pravidla evidence a sdílení údajů ve VDF.

VDF je ale implicitně ukotven v zákoně č. 111/2009 Sb., o základních registrech (ZoZR), zákoně č. 106/1999 Sb., o svobodném přístupu k informacím (InfoZ) a zákoně č. 365/2000 Sb. o informačních systémech veřejné správy (ZoISVS).

Základním pojmem je údaj vedený nebo vytvářený v rámci agendy.

Podle [§ 51 odst. 6, písm. k\) ZoZR](#) je pro každý údaj vedený nebo vytvářený v rámci agendy v RPP vedena jeho přístupnost veřejnosti a v případě, že údaj není přístupný veřejnosti, je povinností uvést číslo a název právního předpisu a označení jeho ustanovení, na jehož základě není údaj přístupný veřejnosti (resp. seznam takových referencí na ustanovení právních předpisů).

Pokud tedy neexistují legislativní překážky ke zveřejnění údaje, jedná se o údaj přístupný veřejnosti (dále jen veřejný údaj). Veřejný údaj je speciálním případem údaje.

Způsob poskytnutí údaje zveřejněním upravuje [§ 4b InfoZ](#), dle kterého se veřejný údaj poskytuje jako informace ve všech formátech a jazycích, ve kterých byla informace vytvořena; při zveřejnění takové informace v elektronické podobě musí být jeden z těchto formátů otevřený a, je-li to možné, též strojově čitelný. Je-li to možné a vhodné, zveřejní povinný subjekt spolu s informací též metadata, která se k ní vztahují. Formát i metadata by měly co nejvíce splňovat otevřené formální normy.

V případě údajů spravovaných v AIS se z principu jedná o informace vedené ve strukturované podobě ve vnitřních databázích informačních systémů, které jsou spravovány databázovými systémy. Jejich zveřejnění v otevřeném a strojově čitelném formátu je tedy technicky možné vždy. Zároveň je možné zveřejnit spolu s informací i metadata a při volbě formátu zveřejněné informace a reprezentaci metadat plně postupovat podle otevřených formálních norem.

Z důvodu dosažení interoperability mezi informačními systémy veřejné správy a také z důvodu dosažení přeshraniční interoperability je nutné při zveřejňování údajů postupovat podle otevřených formálních norem vydávaných MV ČR a při zveřejňování metadat je nutné postupovat dle otevřené formální normy rozhraní katalogů otevřených dat.

Dále je dle [§ 51 odst. 6, písm. k\) ZoZR](#) pro údaj vedený nebo vytvářený v rámci agendy, jehož možné hodnoty jsou vymezeny číselníkem, tento číselník veden v RPP.

Z [§ 54 odst. 1, písm. a\) ZoZR](#) vyplývá, že primárně musí být volen číselník, který již je v RPP veden. Pouze pokud žádný vhodný číselník v RPP veden není, uvádí ohlašovatel agendy svůj vlastní číselník, tj. nový číselník ohlašovatele agendy.

Dle [§ 69 odst. 4 ZoZR](#) MV ČR zveřejnilo otevřenou formální normu pro vytváření číselníků a podle [§ 51 odst. 11 ZoZR](#) jsou číselníky veřejně přístupné způsobem umožňujícím dálkový přístup.



Ze spojení s [§ 5a odst. 2 InfoZ](#) pak vyplývá, že číselník musí být zveřejněn jako otevřená data podle otevřené formální normy vyplývající z [§ 69 odst. 4 ZoZR](#).

Z pohledu ZoZR odlišujeme dva možné typy veřejných číselníků.

- Dle [§ 50 odst. 2 ZoZR](#) to jsou číselníky ČSÚ.
- Dle [§ 54 odst. 1, písm. a\) ZoZR](#) to jsou číselníky ohlašovatelů agend.

Jinými slovy, možné hodnoty údaje lze vymezit buď číselníkem ČSÚ nebo číselníkem ohlašovatele agendy.

2.2 Principy veřejného datového fondu

Obečným výchozím principem VDF je princip [P13 eGovernmentu “Otevřená data jako standard”](#) (Open Data by Default):

“Veřejné údaje evidované orgány veřejné správy ve spravovaných ISVS musí být zveřejňovány jako otevřená data. Pro neveřejné údaje musí být jako otevřená data zveřejňována jejich anonymizovaná podoba, souhrn nebo statistika. V případě, že orgány veřejné správy sdílejí veřejné údaje (včetně anonymizované podoby neveřejných údajů, souhrnů nebo statistik) musí je sdílet jako otevřená data.”

VDF zpřístupňuje veřejné registrované údaje jednotlivým OVM a SPUÚ pro čtení bez omezení přístupu jako otevřená data dle [§ 3 odst. 11 InfoZ](#).

Upřesnění dále uvedených pravidel i samotné metodiky práce s VDF vychází ze čtyř základních principů:

- P1 (distribuovanost) - VDF zastřešuje a popisuje skutečné datové zdroje poskytující údaje prostřednictvím referenčního rozhraní veřejné správy a stanovuje pravidla poskytování a čerpání údajů, ale nevynucuje jejich centralizaci na jedno místo.
- P2 (garance) - pro OVM a SPUÚ čerpající údaje z VDF je garantována technická dostupnost a formální správnost (věcná správnost, úplnost, platnost a pravidelná aktualizace) těchto údajů.
- P3 (otevřená data) - údaje přístupné z VDF jsou zpřístupněny jako otevřená data dle [§ 3 odst. 11 InfoZ](#) bez výjimky.
- P4 (interoperabilita) - údaje jsou z VDF zpřístupněny v podobě, která zajišťuje schopnost různých programových vybavení vzájemně si poskytovat služby a efektivně spolupracovat.

3. Logické komponenty VDF a jejich význam

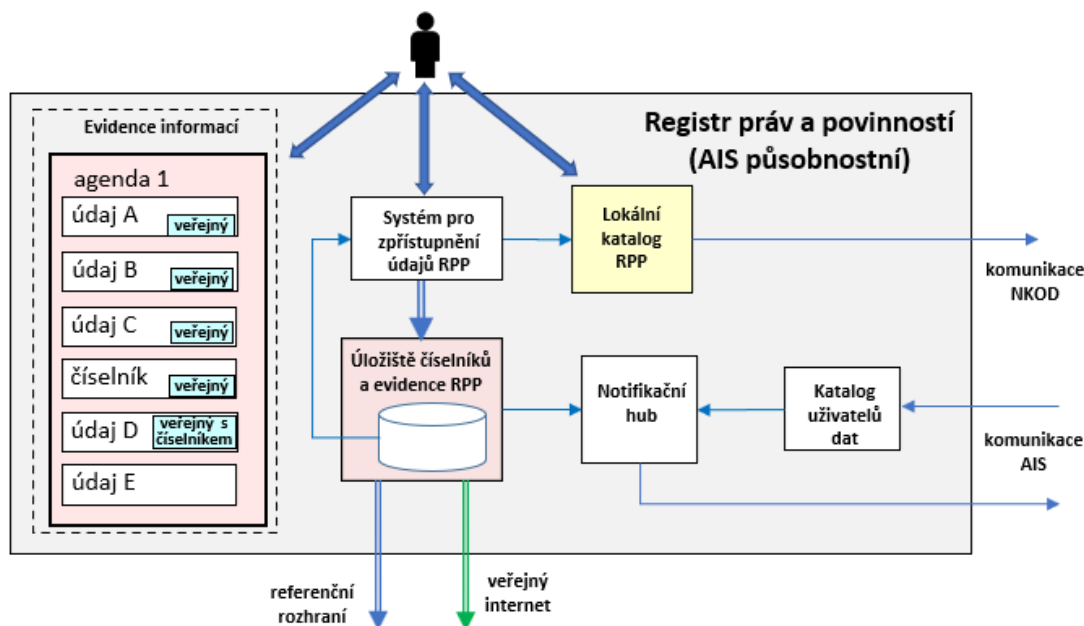
Veřejný datový fond (VDF) je tvořen:

- všemi veřejnými údaji veřejné správy,
- nezbytnou softwarovou podporou zajišťující jejich sdílení a garantovanou dostupnost,
- stanovenými pravidly a definovanými standardy pro zajištění interoperability sdílených údajů,
- poskytovatelem dat definovanými a implementovanými procesy pro zajištění jejich právní závaznosti, platnosti, aktuálnosti a jasně definovanými zodpovědnostmi za sdílené údaje.

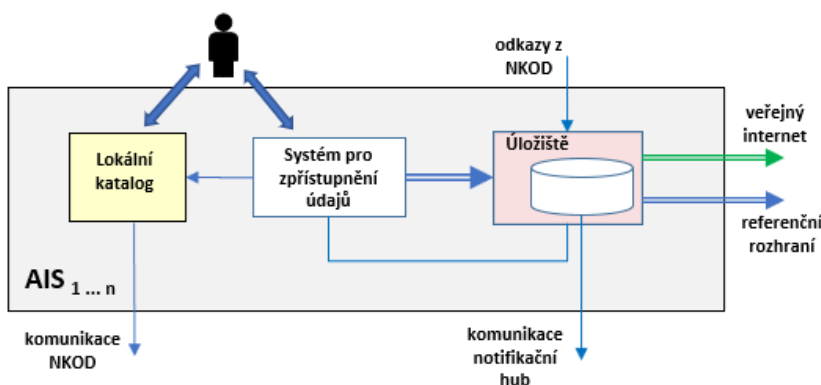
Je tedy zřejmé, že VDF není nějaký další ISVS, který by bylo možné popsat jako jeden celek s přesnou funkční specifikací, ale jedná se o komplexní systém, zahrnující všechny stávající informační systémy veřejné správy a způsoby jejich komunikace.

Z tohoto důvodu metodika nepopisuje jednotlivé oblasti VDF jako přesně vymezené a definované fyzické komponenty, ale pouze jako logické komponenty VDF se všemi jejich potřebnými specifikacemi s tím, že způsoby jejich fyzické realizace (vytvoření nových ICT systémů nebo modulů, formy zakomponování do stávajících ISVS, ...) jsou plně ponechány na rozhodnutích jednotlivých správců ISVS a ohlašovatelů agend.

Realizace společných komponent VDF v AIS Působnostním:



Doporučený způsob realizace budovaných komponent jednotlivými OVM v ostatních AIS:





Způsob realizace komponent VDF formou jejich zahrnutí do příslušného AIS jako jejich nedílné součásti nabízí řadu výhod:

- omezení řešení nutnosti komunikace komponent mimo AIS,
- jednodušší získání a správu metadat publikovaných datových sad,
- jednodušší a systematický přístup k výběru údajů k publikaci (po jednotlivých agendách),
- jednodušší, efektivní a systematickou správu informací o veřejných údajích v RPP.

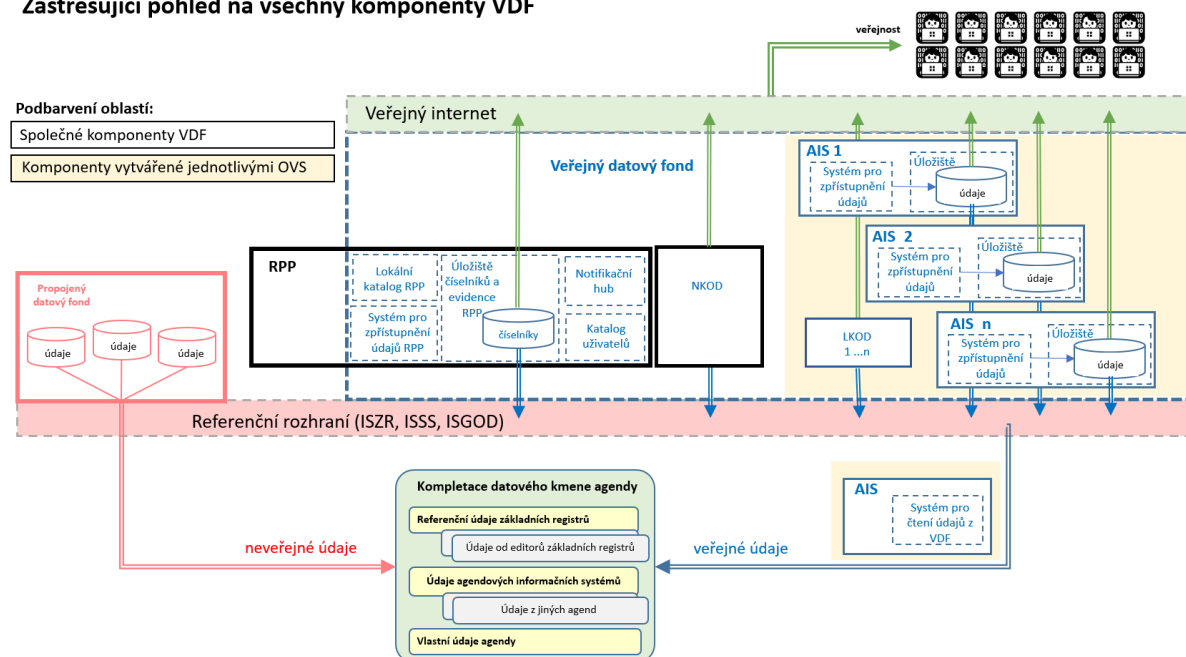
3.1 Zastřešující pohled na logickou strukturu VDF

Dále uvedený obrázek přináší základní zastřešující pohled na logickou strukturu VDF a v následujících kapitolách jsou uvedeny charakteristiky jejich jednotlivých logických komponent.

V zásadě jsou logické komponenty VDF rozděleny do dvou skupin:

- Společné komponenty – jsou zaměřeny především na správu a organizaci zpřístupňovaných údajů a s nimi souvisejících procesů, jsou plně zajišťovány MV ČR.
- Komponenty jednotlivých agend – jsou zaměřeny především na zpřístupnění a užívání údajů jednotlivých agend, jsou zajišťovány jednotlivými OVM, správci agend, správci ISVS.

Zastřešující pohled na všechny komponenty VDF



3.2 Stručná charakteristika společných komponent VDF

RPP – registr práv a povinností

[Registr práv a povinností](#) (AIS působnostní) spravuje Ministerstvo vnitra, obsahuje a spravuje informace pro řízení přístupu k údajům základních registrů a k údajům v agendových informačních systémech.

Pro potřeby VDF je RPP rozšířen o [evidenci doplňujících informací k agendám a jejich údajům](#).

Úložiště číselníků a evidence RPP – nová komponenta RPP



- Obsahuje a zpřístupňuje datové sady publikovaných veřejných číselníků ve formátu CSV, JSON-LD a pomocí SPARQL endpointu.
- Obsahuje a zpřístupňuje obsah (údaje) RPP jako datové sady ve formátech JSON-LD a prostřednictvím SPARQL endpointu.

Lokální katalog RPP – nová komponenta RPP

- Lokální katalog otevřených dat RPP obsahuje katalogizační záznamy pro:
 - datové sady zastřešující verze číselníků,
 - datové sady s verzemi číselníků,
 - datové sady s údaji, které jsou evidovány a spravovány v RPP (obsah RPP).
- Zpřístupňuje katalogizační záznamy k jejich katalogizaci v NKOD.

Katalog uživatelů dat – nová komponenta RPP

- Eviduje využívané veřejné údaje v jednotlivých agendách (zajišťuje ohlašovatel agendy).

Notifikační hub – rozšíření RPP

- Nástroj zajišťující s pomocí katalogu uživatelů dat notifikační službu, která informuje subjekty registrované v katalogu uživatelů dat VDF o změnách publikovaných veřejných údajů.

Systém pro zpřístupnění údajů RPP

- Připravuje obsah RPP a veřejné číselníky do podoby datových sad a dávkově je předává do úložiště datových sad k publikaci.

NKOD (Národní katalog otevřených dat)

- Spravuje a zpřístupňuje katalogizační záznamy všech datových sad dostupných ve VDF.
- Pravidelně přebírá katalogizační záznamy ze všech lokálních katalogů VDF.

ISGOD – komponenta referenčního rozhraní

- Rozhraní VDF pro přístup k publikovaným údajům, které jsou uloženy distribuovaným způsobem v úložištích jednotlivých agendových systémů, fyzické přístupy jsou zajišťovány samotnými úložišti AIS a NKOD.
- ISGOD je pouze logickým pojmenováním a zastřešením celého distribuovaného rozhraní VDF.

3.3 Stručná charakteristika komponent VDF jednotlivých agend (OVM)

Systém pro zpřístupnění údajů

- Aktivně v pravidelných intervalech definovaných správcem ISVS exportuje obsah veřejných údajů do podoby datových sad a dávkově je předává do úložiště datových sad.

Systém pro čtení veřejných údajů z VDF

- Čte veřejné údaje z VDF jako otevřená data prostřednictvím ISGOD v podobě jednotlivých distribucí datových sad ve formátech definovaných datovými specifikacemi.

Úložiště datových sad

- Úložiště datových sad slouží k ukládání publikovaného obsahu registrovaných veřejných údajů vedených v daném publikujícím AIS (ISVS) v podobě distribucí datových sad.
- Provádí technickou kontrolu datových sad předaných "Systémem pro zpřístupnění údajů" a transformuje vytvořené datové sady do všech dalších formátů (distribucí) definovaných příslušnou datovou specifikací.



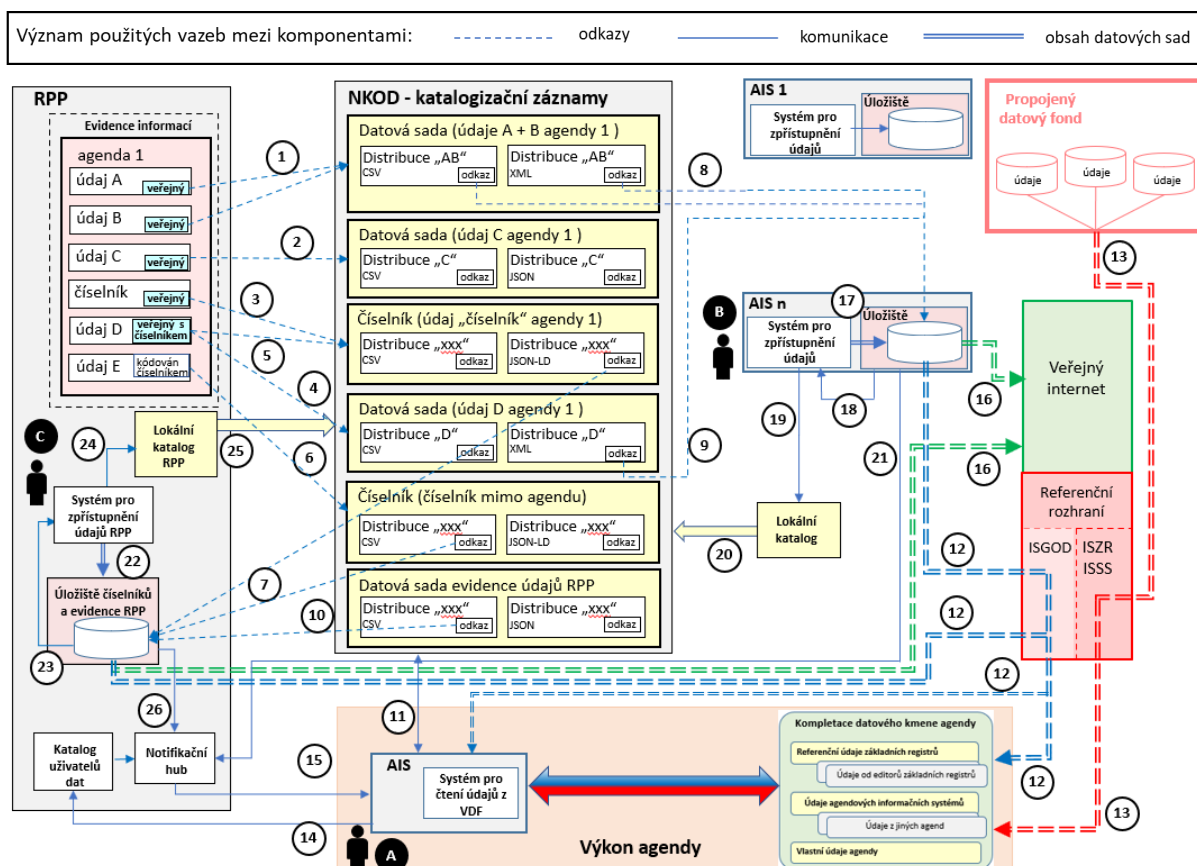
LKOD (Lokální katalog otevřených dat)

- Lokální katalog otevřených dat RPP obsahuje a spravuje katalogizační záznamy datových sad publikovaných veřejných údajů.
- Zpřístupňuje katalogizační záznamy k jejich katalogizaci v NKOD.
- Nabízí se dva možné způsoby jejich realizace:
 - LKOD pro každý konkrétní AIS (komponenta AIS) určený pro své publikované údaje,
 - LKOD konkrétní organizace VS, který katalogizuje všechny údaje publikované organizací.

4. Komponenty VDF a jejich vzájemné vazby

Celkový kontext a vzájemné propojení všech komponent VDF je zachycen na následujícím obrázku. Jedná se o statický pohled na strukturu komponent VDF a jejich vzájemné vazby. Dynamickému pohledu – popisu procesů VDF, je věnována následující kapitola 5.

Vstupní body A, B, C reprezentují situace a interakce pracovníků VS spojené s přípravou publikace údajů do VDF a se čtením údajů z VDF.



4.1 Provázání komponent VDF k zajištění evidence údajů v RPP a NKOD

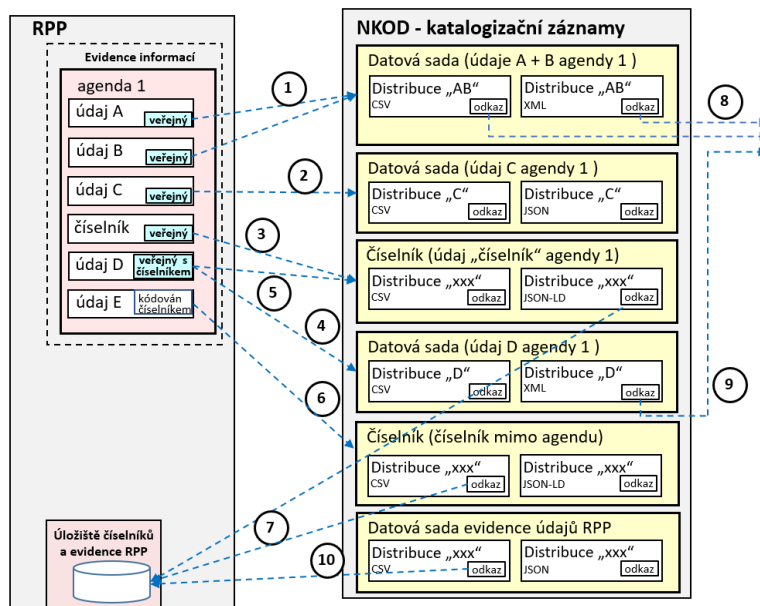
Propojení údajů evidovaných v RPP a v NKOD tvoří “organizační” základ fungování VDF.

RPP obsahuje informace o agendách VS a jejich údajích, NKOD obsahuje informace o datových sadách, ve kterých jsou veřejné údaje publikovány. Výsledný komplex informací vytváří úplnou a celkovou “Mapu údajů veřejné správy”.

Pro správnou činnost VDF, dosažení všech očekávaných výhod a přínosů z něho vyplývajících, a splnění [díličního cíle 5.10](#) v [Informační koncepci ČR](#), je nutná řádná a úplná evidence všech předepsaných metadat v RPP (metadata o existenci údajů a jejich veřejnosti) a NKOD (metadata popisující datové sady, prostřednictvím kterých jsou veřejné údaje publikovány).

Problematiku správy a evidence informací v RPP řeší samostatná metodika “[Metodika evidence údajů v RPP](#)”.

Následující vysvětlující text pod obrázkem slouží k osvětlení vzájemných vazeb mezi údaji v RPP a NKOD a ke zvýraznění jejich významu v kontextu VDF. K vysvětlení je v dalším textu využit popis několika typických situací.



- (1) V RPP je u agendy 1 evidováno, že její údaje A i B jsou veřejné. To znamená, že jsou dostupné jako otevřená data prostřednictvím VDF
 - Oba údaje jsou dostupné prostřednictvím stejné datové sady.
 - V RPP je v evidenci těchto údajů uvedeno IRI této datové sady v NKOD (datová sada se jmenuje „Údaje A + B agendy 1“).
 - Datová sada „Údaje A + B agendy 1“ je publikovaná a dostupná ve dvou distribucích (CSV, XML), katalogizační záznam obsahuje pro každou distribuci odkaz na její fyzické umístění (její URL).
- (2) V RPP je u agendy 1 evidováno, že údaj C je veřejný. Jedná se o stejnou situaci jako v příkladu (1), pouze s tím rozdílem, že údaj C je publikován v jiné samostatné datové sadě „Údaj C agendy 1“.
- (3) Agenda 1 vytváří a udržuje číselník, který je (musí být) registrován v RPP, je dostupný ve VDF a je publikován jako otevřená data (datová sada „Číselník agendy 1“) [podle otevřené formální normy pro číselníky](#).
 - Příslušná datová sada s číselníkem je publikována ve dvou distribucích (CSV, JSON-LD) a obsahuje pro každou distribuci odkaz na její fyzické umístění (její URL).
 - Distribuce číselníků jsou vždy uloženy v Úložišti veřejných číselníků (součást RPP).
- (4) V RPP je u agendy 1 evidováno, že údaj D je veřejný a je publikován v samostatné datové sadě. Jedná se o stejnou situaci jako v příkladu (1).
- (5) U údaje D agendy 1 je také v RPP evidováno, že je kódován číselníkem (v příkladu je uvedena situace s číselníkem, který agenda přímo vytváří, ale vše uvedené platí, pro jakýkoliv využívaný číselník). Evidence údaje D v RPP proto také obsahuje IRI datové sady v NKOD, která obsahuje publikovaný číselník.

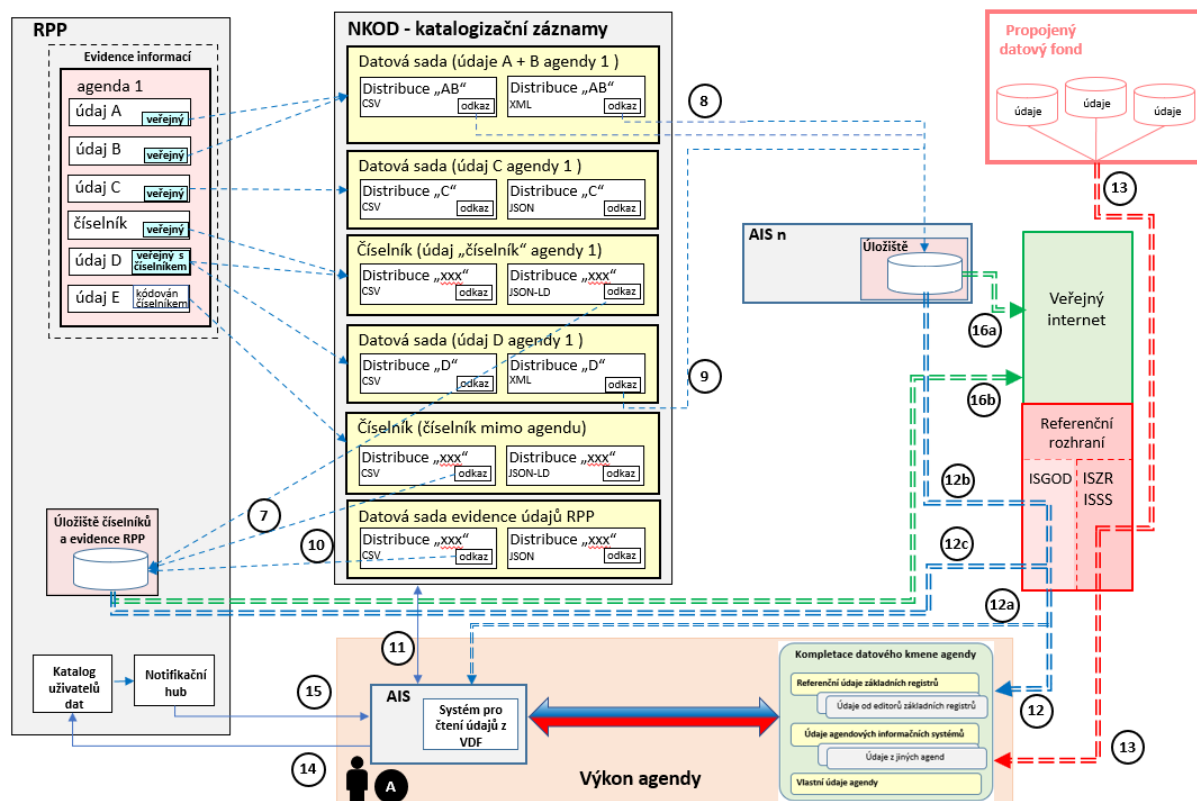
- (6) V RPP je u agentury 1 evidováno, že údaj E je neveřejný, a tudíž není možné jej zpřístupnit prostřednictvím VDF. U údaje E je ale evidováno, že je kódován číselníkem (v tomto případě se pro demonstraci jedná o číselník spravovaný mimo agendu 1). Evidence údaje v RPP obsahuje IRI datové sady v NKOD, který obsahuje publikovaný číselník, stejně jako v příkladu (5).
- (7) Katalogizační záznamy publikovaných číselníků obsahují odkazy (identifikátory IRI) na datové sady s číselníky uloženými v "Úložišti číselníků a evidence RPP".
- (8) Katalogizační záznam publikované datové sady "Údaje A + B agentury 1" obsahuje odkazy na její distribuce uložené v úložišti AIS, který datovou sadu publikoval ve formátu otevřených dat.
- (9) Katalogizační záznam publikované datové sady "Údaj D agentury 1" obsahuje odkaz na její distribuce uložené v úložišti AIS, který datovou sadu publikoval ve formátu otevřených dat.
- (10) Katalogizační záznam datové sady s publikovanou evidencí údajů RPP s odkazem na místo uložení její distribuce v "Úložišti číselníků a evidence RPP". Odkazovaná datová sada obsahuje metadata o údajích vedených v agendách.

4.2 Provázání komponent VDF při čtení údajů z VDF

V této části jsou uvedeny nejdůležitější vazby mezi komponentami VDF, které jsou nezbytné při ustavení a formování procesů pro čtení údajů z VDF. Uvedené vazby nerozlišují, zda se jedná o interakci správce ISVS (manuální) s komponentami VDF, nebo zda se jedná o automatickou vazbu mezi komponentami navzájem.

- Pro potřeby čtení údajů z VDF každý AIS obsahuje ústřední komponentu "Systém pro čtení údajů z VDF".

Schematické znázornění vazeb mezi komponentami při čtení datové sady





Význam jednotlivých vazeb mezi komponentami:

- (A) *Symbolické znázornění interakcí správce AIS při přípravě a nastavení systému pro čtení údajů z VDF.*
- (11) Komunikace "Systému pro čtení údajů z VDF" s NKOD, například při vyhledávání datových sad a získávání jejich metadat.
- (12) Poskytované služby ISGOD (viz [kapitola 6.3.2](#)), zajišťující:
- Čtení (stažení) datových sad s informacemi o vybrané agendě veřejné správy a jejich údajích z "Úložiště číselníků a evidence RPP".
 - Čtení (stažení) vybrané datové sady z „Úložiště AIS“.
 - Čtení (stažení) datové sady z "Úložiště číselníků a evidence RPP".
- (13) Poskytované služby propojeným datovým fondem PPDF (čtení dalších veřejných/neveřejných údajů poskytovaných PPDF) - nesouvisí přímo s VDF, uvedeno pouze pro úplnost kontextu.
- (14) Komunikace "Systému pro čtení údajů z VDF" s "Katalogem uživatelů dat" z důvodu registrace agendy v "Katalogu uživatelů dat" k zasílání notifikací při změnách v používaných datových sadách.
- (15) Komunikace „Notifikačního hubu“ a "Systému pro čtení údajů z VDF" (zasílání notifikací o změnách publikovaných veřejných údajů).
- (16) Přístup k datovým sadám prostřednictvím veřejného internetu:
- Čtení (stažení) datových sad z úložiště AIS.
 - Čtení (stažení) datových sad z "Úložiště číselníků a evidence RPP".

4.3 Provázání komponent VDF při publikaci údajů do VDF

Uvedeny jsou nejdůležitější vazby mezi komponentami VDF, které jsou nezbytné při ustavení a formování procesů pro publikaci veřejných údajů do VDF. Uvedené vazby nerozlišují, zda se jedná o interakci správce ISVS (manuální) s komponentami VDF, nebo zda se jedná o automatickou vazbu mezi komponentami navzájem.

Publikace údajů rozlišuje dva typy publikace:

- publikace veřejných údajů z datového kmene agendy (zde patří i publikace obsahu evidence RPP),
- publikace veřejných číselníků.

Publikace veřejných údajů z datového kmene agendy.

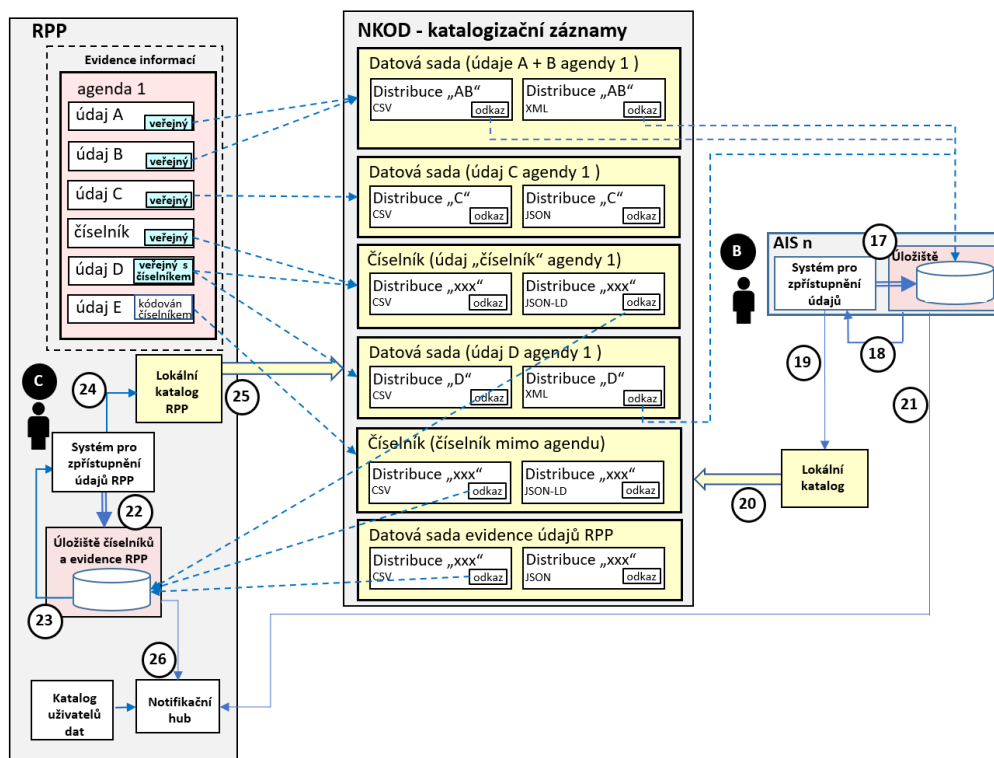
- Poskytovatelem veřejných údajů do VDF je správce ISVS, ve kterém jsou vedeny registrované veřejné údaje.
- Pro potřeby sdílení údajů ve VDF každý ISVS obsahuje "Systém pro zpřístupnění údajů prostřednictvím VDF".
- Pro ukládání publikovaného obsahu registrovaných veřejných údajů vedených v daném publikujícím AIS (ISVS) je určeno "Úložiště".

Publikace veřejných číselníků.

- Poskytovatelem veřejných číselníků do VDF je buď ohlašovatel agendy nebo ČSÚ.
- Ohlašovatel agendy zavádí veřejný číselník jen v případě, že je agendový údaj kódován číselníkem, který ještě není v RPP nikým zaveden.

- Pro potřeby sdílení veřejných číselníků ve VDF RPP (AIS působnostní) obsahuje "Systém pro zpřístupnění údajů prostřednictvím VDF", který zajišťuje publikaci obsahu veřejných číselníků do VDF.
- Pro ukládání obsahu veřejných číselníků a také obsahu veřejných údajů vedených v RPP je využito "Úložiště číselníků a evidence RPP", ve kterém je také uložen obsah RPP publikovaný jako otevřená data.

Schematické znázornění vazeb mezi komponentami při publikaci veřejných údajů.



Publikace veřejných údajů z datového kmene agend

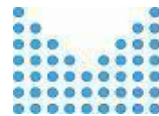
- (B) Symbolické znázornění přípravy a nastavení systému pro zpřístupnění údajů správcem AIS k publikaci veřejných údajů z datového kmene agendy do VDF.
- (17) Dávkové předávání připravených datových sad "Systémem pro zpřístupnění údajů" v podobě datových souborů (distribucí) v jednom základním formátu do "Úložiště".
- (18) Předávání metadat o všech vytvořených distribucích datových sad (včetně odpovídajících lokálních identifikátorů) "Systému pro zpřístupnění údajů".
- (19) Předávání katalogizačních záznamů datových sad a jejich distribucí včetně referenčních identifikátorů do LKOD.
- (20) Předávání katalogizačních záznamů z LKOD do NKOD.
- (21) Komunikace "Systému pro zpřístupnění údajů" s "Notifikačním hubem" (v případě změny publikovaného veřejného údaje, nebo provedené aktualizace stávající datové sady).

Publikace veřejných číselníků a obsahu veřejných údajů vedených v RPP

- (C) Symbolické znázornění přípravy a nastavení systému pro zpřístupnění údajů RPP a veřejných číselníků správcem AIS a poskytovatelem veřejných číselníků.



- (22) Předávání veřejných číselníků a obsahu veřejných údajů vedených v RPP v podobě definované otevřenou formální normou pro číselníky do “Úložiště číselníků a evidence RPP”.
- (23) Předávání metadat “Systému pro zpřístupnění údajů RPP” všech vytvořených distribucí datových sad:
 - s verzemi veřejných číselníků dle otevřené formální normy pro číselníky,
 - s obsahem veřejných údajů vedených v RPP.
- (24) Předávání katalogizačních záznamů datových sad a všech vytvořených distribucí “Lokálnímu katalogu RPP”.
- (25) Předávání katalogizačních záznamů z “Lokálního katalogu RPP” do NKOD.
- (26) Komunikace “Systému pro zpřístupnění údajů RPP” s “Notifikačním hubem” v případě změn:
 - publikovaného obsahu (údaje) RPP (proběhla aktualizace stávající datové sady),
 - publikovaného číselníku (proběhla aktualizace stávající verze číselníku).



5. Popis hlavních procesů VDF

5.1. Publikace veřejných údajů

Systém pro zpřístupnění údajů transformuje a exportuje obsah veřejných údajů v podobě datových sad do úložiště datových sad v takovém režimu, aby bylo zajištěno, že publikované veřejné údaje budou pravidelně aktualizované, platné a právně závazné. Stanovení režimu publikace a jeho zajištění je plně v zodpovědnosti publikujícího OVM (správce AIS), který musí při jeho stanovení vycházet z charakteru a povahy publikovaných dat, četnosti jejich změn a aktualizací, a také způsobů jejich použití.

Z úložiště datových sad jsou veřejná data následně dostupná v podobě otevřených dat pro výkon veřejné správy, pro vzájemné sdílení veřejných údajů mezi veřejnoprávními subjekty navzájem a pro sdílení veřejných údajů mezi veřejnoprávní a soukromoprávní sférou.

5.1.1 Pravidla pro evidenci veřejných údajů v RPP

Evidenci údajů v RPP a s ní související způsoby práce s RPP plně řeší "[Metodika evidence údajů v RPP](#)".

Dále v textu jsou uvedeny pouze vybrané informace, které jsou důležité pro dokreslení a pochopení kontextu zpřístupnění veřejných údajů prostřednictvím VDF.

K zajištění funkčnosti VDF je tedy nutné dodržet:

- Označení údajů vedených v agendách a registrech příznakem "veřejný údaj / neveřejný údaj", s uvedením odkazu na konkrétní legislativu, ze které vyplývá jejich případná neveřejnost.
- Uvedení IRI datových sad v NKOD, jejichž prostřednictvím je veřejný údaj publikován.
- U agendových údajů, které jsou kódované číselníky, jejich označení jako údaje "kódované konkrétní verzi veřejného číselníku", včetně uvedení IRI datové sady v NKOD, která obsahuje publikovanou příslušnou verzi veřejného číselníku.
- Evidenci (ohlášení) využívání jednotlivých veřejných údajů prostřednictvím VDF konkrétními agendami. Ohlašovatel agendy v ohlášení agendy uvede výčet veřejných údajů, které budou využívány pro výkon agendy.

5.1.2 Pravidla pro údaje zpřístupněné veřejným datovým fondem

- Prostřednictvím VDF jsou zpřístupňovány veřejné údaje spravované jednotlivými OVM.
- Veřejné registrované údaje jsou zpřístupněny pro výkon veřejné správy v datových sadách prostřednictvím referenčního rozhraní (ISGOD) pro potřeby čtenářů údajů – OVM a SPUÚ.
- Veřejné registrované údaje jsou také zpřístupněny veřejnosti prostřednictvím otevřeného přístupu (tj. jako otevřená data dle § 3 odst. 11 InfoZ) v totožné podobě (tj. s totožnou strukturou a sémantikou).
- Interoperabilita zpřístupněných veřejných údajů je zajištěna standardizací jejich publikace na základě Otevřených formálních norem (OFN) vydaných MV ČR a Datových specifikací vytvořených jednotlivými poskytovateli dat.
 - Datová specifikace popisuje strukturu a sémantiku konkrétní datové sady nebo datových sad. Obsahuje dokumentaci významu obsahu datové sady v podobě jejího konceptuálního modelu a jeho popisu a datová schémata určující datové struktury, ve kterých je obsah distribuován v různých formátech, např. JSON schéma a XSD schéma popisující JSON a XML distribuce obsahu datové sady.



- Datovou specifikaci vydává poskytovatel konkrétních veřejných údajů a specifikuje pomocí ní konkrétní podobu datové sady nebo datových sad, ve kterých veřejné údaje zpřístupňuje.
- OFN jsou vymezeny § 3 odst. 9 zákona č. 106/1999 Sb. o svobodném přístupu k informacím a jsou dle § 4b odst. 1 závazné. OFN jsou Datové specifikace vydávané MV ČR a jsou určeny pro použití všemi poskytovateli, kteří publikují data odpovídající dané OFN. OFN jsou k dispozici v Repositáři OFN, který spravuje MV ČR. Navrhována jsou ve spolupráci s více poskytovateli k zachycení úplného kontextu publikovaných dat.
- Návrh Datové specifikace musí vycházet ze stávajících OFN dostupných v Repositáři OFN.
- Pro daný typ publikovaných údajů musí existovat, nebo být připravena Datová specifikace, případně OFN.
- Datové sady jsou popsány a katalogizovány v podobě katalogizačních záznamů (metadat) v NKOD.
- Fyzicky jsou datové sady dostupné v podobě distribucí, pro které platí:
 - různé distribuce stejné datové sady zpřístupňují její obsah v různých formátech (např. CSV, XML nebo JSON) a prostřednictvím různých přístupových mechanismů,
 - každá distribuce je zaznamenána v katalogizačním záznamu datové sady v NKOD.
- Způsoby fyzického zpřístupnění obsahu datové sady:
 - datový soubor s kompletním obsahem datové sady určen ke stažení,
 - prostřednictvím API, které umožňuje přistupovat ke kompletním údajům o každé jednotlivé entitě či konceptu, o němž jsou v datové sadě reprezentovány údaje, prostřednictvím dereference identifikátoru entity či konceptu, který je stanoven poskytovatelem údajů v podobě IRI (Internationalized Resource Identifier, viz [Otevřená formální norma pro propojená data](#))
 - prostřednictvím rozhraní API, které umožňuje dotazování nad obsahem datové sady s pomocí dotazovacího jazyka SPARQL.
- Informace o veřejnosti registrovaného údaje je zachycena v jeho evidenci v RPP označením údaje jako “veřejný údaj”.
- Zpřístupnění veřejného údaje prostřednictvím VDF je zachyceno v evidenci RPP:
 - uvedením IRI datové sady (nebo datových sad) v NKOD, v níž je obsah odpovídajícího údaje zpřístupněn a publikován jako otevřená data,
 - v případě údaje kódovaného číselníkem obsahuje evidence veřejného údaje také IRI datové sady v NKOD, v níž je číselník zpřístupněn a publikován jako otevřená data.
- Technická podoba dat určených pro sdílení prostřednictvím VDF je definována jejich datovou specifikací nebo [otevřenými formálními normami](#).

5.1.3 Popis procesu publikace veřejných údajů

Předpoklady:

- Řádně a úplně vedená [evidence údajů v RPP](#) ze strany ohlašovatelů agend a správců ISVS.

Postup:

- Poskytovatel údajů vybere údaje pro publikaci do VDF (viz. [Příprava datových sad k publikaci](#)).



- Systém pro zpřístupnění údajů připraví obsah datové sady v podobě datového souboru. Připravený datový soubor musí dodržovat připravenou Datovou specifikaci nebo příslušnou OFN.
- Systém pro zpřístupnění údajů zašle připravený obsah datové sady do úložiště datových sad.
- Úložiště datových sad provede technickou kontrolu zasláného obsahu:
 - Kontrola správného formátování (např. JSON, CSV nebo XML formátování).
 - Kontrola validity datové struktury vůči datovým schémátům definovaných Datovou specifikací nebo OFN.
 - V případě špatné syntaxe zašle zpět systému pro zpřístupnění údajů chybové hlášení a skončí.
- Úložiště datových sad zpřístupní všechny vytvořené distribuce jako datové soubory dostupné ke stažení prostřednictvím ISGOD a z veřejného internetu prostřednictvím URL pro stažení datového souboru (URL pro stažení datového souboru je stejné pro přístup prostřednictvím ISGOD i veřejného internetu).
- Úložiště datových sad zašle zpět systému pro zpřístupnění údajů potvrzení o úspěšném uložení distribucí.
 - Jako součást potvrzení zasílá metadata o vytvořených distribucích v podobě definované otevřenou formální normou pro rozhraní katalogů otevřených dat.
- Systém pro zpřístupnění údajů vytvoří kompletní katalogizační záznam o datové sadě, včetně doplnění metadat o distribucích vytvořených úložištěm datových sad a vše katalogizuje v lokálním katalogu.
- Systém pro zpřístupnění údajů zašle notifikačnímu hubu informaci o změně publikovaného veřejného údaje (obsahu datové sady).
- Národní katalog otevřených dat převezme katalogizační záznam prostřednictvím API lokálního katalogu a katalogizuje publikované datové sady.
- Ohlašovatel agendy ohlásí do RPP jako součást ohlášení agendy referenční IRI datové sady (datových sad) v NKOD, ve které (kterých) je veřejný údaj zpřístupněn. Ohlášení provede poté, co NKOD datovou sadu na základě zasláného katalogizačního záznamu katalogizuje (zpravidla do 1 dne).
- U agendových údajů, které jsou kódované číselníky, ohlašovatel agendy je označí jako údaje “kódované konkrétní verzi veřejného číselníku”, a uvede IRI datové sady v NKOD, která obsahuje publikovanou příslušnou verzi veřejného číselníku.
- Notifikační hub eviduje informaci o změně publikovaného veřejného údaje zaslanou systémem pro zpřístupnění údajů.

5.2 Publikace veřejných číselníků

5.2.1 Pravidla pro evidenci veřejných číselníků

K zajištění funkčnosti VDF je nutné dodržet:

- Uvedení IRI datových sad v NKOD, jejichž prostřednictvím je veřejný číselník publikován.
- Každá verze veřejného číselníku (1... n verzí) je spravována jako jednotlivá datová sada.
- Všechny datové sady reprezentující jednotlivé verze číselníku jsou seskupeny do zastřešující datové sady (tzv. datová série).



- Zastřešující datová sada i jednotlivé verze číselníků jsou evidovány prostřednictvím metadat datových sad dle [OFN pro rozhraní katalogů otevřených dat](#).
- Pro evidenci zastřešujících datových sad číselníků i datových sad jednotlivých verzí platí pro metadata následující úpravy:
 - vlastnost poskytovatel získává jiný význam (poskytovatelem datové sady u číselníků je vždy MV ČR),
 - eviduje se OVM, který zavádí veřejný číselník do RPP, ale jako “poskytovatel číselníku”, nikoliv jako poskytovatel datové sady
- Pro evidování datových sad reprezentujících jednotlivé verze číselníků je soubor údajů navíc rozšířen o evidenci vazeb mezi jednotlivými verzemi číselníků:
 - vazba “je verzí” - (reference na zastřešující datovou sadu),
 - vazba “má předchozí verzi” (reference na datovou sadu s předchozí verzí číselníku, pokud existuje).

5.2.2 Pravidla pro publikaci veřejných číselníků do VDF

- Publikace veřejného číselníku do VDF vyžaduje jeho zavedení do RPP dle § 51 odst. 8 ZoZR.
 - ČSÚ zavádí číselníky nezávislé na agendách,
 - ohlašovatel agendy zavádí číselník jenom v případě, kdy agendový údaj je kódován číselníkem, a ten není ještě v RPP zaveden.
- Veřejné číselníky mají statut referenčních údajů, jsou evidovány v RPP a jsou také z něho publikovány do VDF.
- Pro potřeby všech veřejných číselníků vystupuje AIS působnostní jako systém pro správu datového kmene s komponentami:
 - systém pro zpřístupnění údajů RPP, která zajišťuje publikaci obsahu veřejných číselníků do VDF,
 - úložiště číselníků a evidence RPP s uloženými publikovanými datovými sadami,
 - API poskytující katalogizační záznamy o jednotlivých datových sadách s veřejnými číselníky a s obsahem údajů vedených v RPP.
- Veřejné číselníky jsou publikovány a dostupné dle [OFN pro číselníky](#).
- Zavedení nového veřejného číselníku vyžaduje vždy nejprve zavedení zastřešující datové sady a až následně zavedení jednotlivých verzí číselníku.
- Zavedení číselníku a specifikace metadat se řídí následujícím postupem:
 - specifikace metadat pro zastřešující datovou sadu číselníku,
 - specifikace metadat pro datovou sadu s první verzí číselníku (možnost kopírování metadat ze zastřešující datové sady),
 - zavedení obsahu prvního číselníku v uživatelském rozhraní RPP pomocí nahrání připraveného souboru.
- Všechny tři uvedené kroky lze také realizovat automatizovaně načtením seznamu veřejných číselníků poskytovatele z URL, který musí být zpřístupněn dle otevřené formální normy pro rozhraní katalogů otevřených dat.
- V případě nahrazení číselníku zcela novým číselníkem, zůstane původní číselník evidován.



5.2.3 Popis procesu publikace číselníků

Předpoklady

- Řádně a úplně vedená evidence veřejných číselníků ze strany poskytovatelů číselníků.

Postup publikace

Postup publikace veřejných číselníků do VDF je v podstatě totožný s postupem publikace veřejných údajů pouze s úpravami uvedenými v pravidlech pro publikaci veřejných číselníků do VDF. Ty se týkají převážně interakce poskytovatele číselníku se systémem pro zpřístupnění údajů RPP prostřednictvím uživatelského rozhraní RPP.

Upravený postup:

- Poskytovatel číselníku prostřednictvím uživatelského rozhraní RPP specifikuje metadata pro zastřešující datovou sadu číselníku a specifikuje metadata pro datovou sadu s první verzí číselníku.
- Poskytovatel číselníku předá ručně obsah první verze číselníku v uživatelském rozhraní RPP nahráním připraveného souboru s obsahem první verze veřejného číselníku.
- Systém pro zpřístupnění údajů RPP připraví obsah datové sady číselníku v podobě datového souboru v jednom z formátů definovaným otevřenou formální normou pro číselníky.
- Systém pro zpřístupnění údajů RPP předá připravený obsah datové sady do úložiště číselníků a evidence RPP.
- Úložiště datových sad provede technickou kontrolu zaslaného obsahu:
 - Kontrola správného formátování (CSV, JSON-LD).
 - Kontrola validity datové struktury vůči datovým schématům definovaným otevřenou formální normou (CSV, JSON-LD).
 - V případě špatné syntaxe zašle zpět systému pro zpřístupnění údajů RPP chybové hlášení a skončí.
- Úložiště číselníků a evidence RPP vytvoří další distribuce obsahu datové sady jeho transformací do všech podob definovaných otevřenou formální normou.
- Úložiště číselníků a evidence RPP zpřístupní všechny vytvořené distribuce:
 - Jako datové soubory dostupné ke stažení prostřednictvím ISGOD a z veřejného internetu prostřednictvím URL pro stažení datového souboru (URL pro stažení datového souboru je stejné pro přístup prostřednictvím ISGOD i veřejného internetu, nutné správně nastavit DNS v prostředí KIVS/CMS a DNS v prostředí veřejného internetu).
 - Jako jednotlivé položky obsahu dle otevřené formální normy pro propojená data, každá položka má své referenční a lokální IRI dereferencovatelné prostřednictvím ISGOD a z veřejného internetu (referenční i lokální IRI položky jsou stejné pro přístup prostřednictvím ISGOD i veřejného internetu, nutné správně nastavit DNS v prostředí KIVS/CMS a DNS v prostředí veřejného internetu).
 - V podobě SPARQL endpointu prostřednictvím ISGOD i ve veřejném internetu (URL SPARQL endpointu je stejné pro přístup prostřednictvím ISGOD i veřejného internetu), nutné správně nastavit DNS v prostředí KIVS/CMS a DNS v prostředí veřejného internetu).
- Úložiště číselníků a evidence RPP zašle zpět systému pro zpřístupnění údajů RPP potvrzení o úspěšném uložení distribucí.
 - Jako součást potvrzení zasílá metadata o vytvořených distribucích v podobě definované otevřenou formální normou pro rozhraní katalogů otevřených dat.



- Systém pro zpřístupnění údajů RPP doplní katalogizační záznamy zadané poskytovatelem číselníku o metadata poslané úložištěm číselníků a evidence RPP, a vše katalogizuje v lokálním katalogu RPP.
- Systém pro zpřístupnění údajů RPP zašle notifikačnímu hubu informaci o změně publikovaného číselníku.
- Národní katalog otevřených dat převezme katalogizační záznam prostřednictvím API lokálního katalogu RPP a katalogizuje publikované datové sady s číselníky.
- Poskytovatel číselníku zavede do RPP referenční IRI datové sady (datových sad) v NKOD, ve které (kterých) je veřejný číselník zpřístupněn. Ohlášení provede poté, co NKOD datovou sadu na základě zaslaného katalogizačního záznamu katalogizuje (zpravidla do 1 dne).
- Notifikační hub eviduje informaci o změně publikovaného číselníku zaslanou systémem pro zpřístupnění údajů RPP.

5.3 Čtení veřejných údajů a veřejných číselníků z VDF

Postupy čtení jsou shodné pro veřejné údaje i pro veřejné číselníky.

Systém pro čtení údajů provádí čtení datového souboru za účelem aktualizace vlastní kopie údajů přebíraných z VDF. Aktualizaci typicky provádí v pravidelných intervalech nebo na základě notifikací o změnách zasílaných notifikačním hubem na základě registrace v katalogu uživatelů.

Předpoklady pro čtení:

- Systém pro čtení údajů má k dispozici referenční IRI datové sady v NKOD, která obsahuje požadovaný veřejný údaj.
- Možnosti získání IRI datové sady:
 - nalezení v evidenci agendových údajů v RPP,
 - prostřednictvím uživatelského rozhraní RPP,
 - prostřednictvím datových sad s publikovaným obsahem RPP,
 - vyhledání datové sady s publikovanými údaji přímo v NKOD.

Postup čtení:

- Systém pro čtení údajů přistoupí k referenčnímu IRI datové sady v NKOD a přečte si její metadata.
- Systém pro čtení údajů zvolí vhodnou distribuci datové sady dle potřebného formátu a prostřednictvím její URL přistoupí prostřednictvím služeb ISGOD ke stažení obsahu vybrané distribuce.
- Úložiště datových sad zprostředkuje systému pro čtení údajů obsah datového souboru na daném URL.



6. Specifikace jednotlivých komponent VDF

Kapitola je zaměřena na popis funkcí a požadovaných činností jednotlivých logických komponent VDF bez rozlišení, zda jsou zajišťovány informačními systémy VS a jejich doplňky, nebo činnostmi pracovníků VS. Míra automatizace uvedených funkcí/činností je plně závislá na rozhodnutí každého konkrétního OVM (nebo správce AIS) při jejich implementaci.

Kompletní realizace VDF bude vyžadovat delší časové období a odvíjí se od současného stavu implementace jeho společných komponent. Z tohoto důvodu jsou funkce a činnosti u některých komponent rozděleny do dvou skupin:

- První skupina **“Základní funkce”** obsahuje funkce a činnosti odpovídající aktuálnímu stávajícímu stavu společných komponent VDF. Jsou nutné a nezbytné pro zahájení implementace VDF v prostředí veřejné správy.
- Druhá skupina **“Rozšiřující funkce”** obsahuje zbývající funkce a činnosti kompletující finální podobu VDF popsanou v globální analýze VDF, jejich realizace proběhne až v navazujícím období po implementaci základní funkcionality VDF.

6.1 Společné komponenty VDF

Společné komponenty jsou zaměřeny především na správu a organizaci zpřístupňovaných údajů a souvisejících procesů, a v podstatě vytvářejí základní společně sdílenou “infrastrukturu” veřejného datového fondu.

Implementaci všech společných komponent zajišťuje a také je následně provozuje MV ČR.

6.1.1 RPP registr práv a povinností

- [Registr práv a povinností](#) (AIS působnostní) spravuje Ministerstvo vnitra a obsahuje informace pro řízení přístupu k údajům základních registrů a k údajům v agendových informačních systémech, nutných pro výkon veřejné správy.
- Pro potřeby VDF byl RPP rozšířen o evidenci přístupnosti registrovaných údajů veřejnosti (veřejnost údajů), o evidenci veřejných číselníků a evidenci IRI datových sad v NKOD, obsahujících jednotlivé údaje ve formátu otevřených dat.
- Dále byl RPP rozšířen o evidenci informací o způsobu kódování veřejných údajů číselníky s uvedením IRI datových sad v NKOD, obsahujících odpovídající publikované číselníky ve formátu otevřených dat.

Nutné předpoklady a podmínky:

- Základním nutným předpokladem a podmínkou je zodpovědný přístup všech subjektů veřejné správy k řádné, úplné a bezchybné evidenci všech evidovaných údajů v RPP.

6.1.1.1 Evidence údajů agend

Evidence údajů v RPP:

- Seznam agend veřejné správy a jejich povinností.
- Seznam Orgánů veřejné moci a soukromoprávních uživatelů údajů ze základních registrů.
- Mapa působnosti orgánů veřejné moci v rámci agendového modelu.
- Údaje o údajích vedených v agendách a o jejich poskytování a využívání.
- Údaje o oprávněních orgánů veřejné moci a soukromoprávních uživatelů k přístupu k údajům ze základních registrů a agendových informačních systémů.



- Rozhodnutí, na jehož základě dochází ke změnám referenčních údajů v Registru obyvatel a Registru osob.
- Seznam informačních systémů veřejné správy a jejich vazeb na agendy a údaje v nich vedených.
- Označení údajů vedených v agendách a registrech příznakem “veřejný údaj / neveřejný údaj”, s uvedením odkazu na konkrétní legislativu, ze které vyplývá jejich případná neveřejnost.
- Uvedení IRI datových sad v NKOD, jejichž prostřednictvím je veřejný údaj publikován.
- Označení agendových údajů, které jsou kódované číselníky, jako údaje “kódované konkrétní verzi veřejného číselníku”, a uvedení IRI datové sady v NKOD, která obsahuje publikovanou příslušnou verzi veřejného číselníku.
- Evidence (ohlášení) využívání jednotlivých veřejných údajů konkrétními agendami prostřednictvím VDF.

Základní funkce:

- Označování “veřejnosti / neveřejnosti” údajů jednotlivých agend.
- Evidování odkazů na příslušnou legislativu v případě “neveřejnosti” údajů, ze které vyplývá, že se jedná o neveřejné údaje.
- Evidování IRI datových sad v NKOD, prostřednictvím kterých je veřejný údaj publikován.
- Označování údajů jako údaje kódované konkrétní verzi veřejného číselníku.
- Evidování IRI datové sady v NKOD, prostřednictvím které je konkrétní verze veřejného číselníku publikována.
- Evidování agend s výčtem veřejných registrovaných údajů, které jsou využívány pro výkon agendy.
- Vyhledávání potřebných informací o agendách a jejich údajích v evidenci údajů RPP.

6.1.1.2 Evidence veřejných číselníků

Evidence údajů o číselnících v RPP:

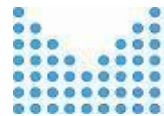
- Identifikátor číselníku,
- název číselníku,
- verze číselníku (doplněné automaticky, nový číselník = 1. verze),
- popis číselníku,
- OVM, který zodpovídá za číselník (doplněno podle OVM přihlášeného uživatele),
- platnost od,
- platnost do.

Funkce:

- Funkce přístupné prostřednictvím uživatelského rozhraní RPP pro práci s číselníky
 - Zavedení nového veřejného číselníku poskytovatelem číselníku.
 - Vytvoření nové verze již zavedeného veřejného číselníku poskytovatelem číselníku.
 - Editování rozpracované verze číselníku.
 - Zobrazení číselníku.
- Automatizované načtení seznamu veřejných číselníků poskytovatele ze zadaného URL.

6.1.2 Systém pro zpřístupňování údajů z RPP

Nutné předpoklady a podmínky:



- Vybraný obsah veřejných údajů evidovaných v RPP k publikaci, nebo připravený číselník k publikaci.

Funkce:

- Uživatelské rozhraní pro zavádění číselníků:
 - specifikace metadat pro zastřešující datovou sadu číselníku,
 - specifikace metadat pro datovou sadu s první verzí číselníku,
 - zavedení obsahu prvního číselníku v uživatelském rozhraní RPP pomocí nahrání připraveného souboru.
- Automatizované načtení seznamu veřejných číselníků z URL zadaným poskytovatelem číselníků.
- Příprava obsahu datové sady číselníku v podobě datového souboru v jednom z formátů definovaných otevřenou formální normou pro číselníky.
- Příprava obsahu datové sady s obsahem veřejných údajů evidovaných v RPP získaných z interní databáze RPP a AIS působnostní, v podobě datového souboru v jednom z formátů definovaných otevřenou formální normou (formáty CSV, JSON-LD a SPARQL endpoint).
- Předání připraveného obsahu datové sady do úložiště číselníků a evidence RPP.
- Kompletace katalogizačních záznamů (zadaných ručně poskytovatelem číselníku) o metadata poslané úložištěm číselníků a evidence RPP.
- Záznam katalogizačních záznamů do LKOD.
- Odeslání informace o změně publikovaného obsahu nebo veřejného číselníku notifikačnímu hubu

6.1.3 Úložiště číselníků a evidence RPP a datových sad RPP

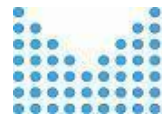
Obsahuje a zpřístupňuje datové sady publikovaných veřejných číselníků a obsahu RPP ve formátu otevřených dat.

Nutné předpoklady a podmínky:

- Systém pro zpřístupnění údajů RPP předá připravený obsah datové sady úložišti číselníků a evidence RPP.

Funkce:

- Technická kontrola datové sady:
 - kontrola správného formátování (CSV, JSON, JSON-LD),
 - kontrola validity datové struktury vůči datovým schématům definovaných otevřenou formální normou (CSV, JSON, JSON-LD),
 - odeslání systému pro zpřístupnění údajů RPP chybové hlášení v případě špatné syntaxe a ukončení činnosti.
- Vytvoření dalších distribucí datových sad s verzemi veřejných číselníků evidovaných v RPP dle otevřené formální normy pro číselníky (formáty CSV, JSON-LD).
- Vytvoření dalších distribucí datových sad s obsahem veřejných údajů evidovaných v RPP, které jsou získané z interní databáze RPP a AIS působnostní dle otevřené formální normy (formáty CSV, JSON-LD).
- Umístění a zpřístupnění evidovaných veřejných číselníků a datových sad RPP ve formě publikovaných datových sad.
- Umístění a zpřístupnění dalších publikovaných veřejných údajů evidovaných v RPP, které jsou získané z interní databáze RPP a AIS působnostní ve formě publikovaných datových sad.



- Odeslání potvrzení o úspěšném uložení distribucí systému pro zpřístupnění údajů RPP s metadaty o vytvořených distribucích, v podobě definované otevřenou formální normou pro rozhraní katalogů otevřených dat.
- Zajištění dostupnosti publikovaných datových sad z veřejného internetu.
- Zajištění dostupnosti publikovaných datových sad prostřednictvím referenčního rozhraní (ISGOD) s garancí jejich dostupností pro výkon veřejné správy.

6.1.4 Lokální katalog RPP

Lokální katalog otevřených dat RPP obsahuje katalogizační záznamy pro:

- datové sady zastřešující verze číselníků,
- datové sady s verzemi číselníků,
- datové sady veřejných údajů vedených v RPP.

Metadata pro datové sady číselníku vychází z údajů evidovaných ve správě číselníku a jsou doplněny o automatizovaně vytvářená data z publikace datových sad s číselníky v úložišti.

Nutné předpoklady a podmínky:

- Implementované rozhraní katalogu dle OFN [Rozhraní katalogů otevřených dat](#).

Funkce:

- Správa katalogizačních záznamů datových sad publikovaných číselníků a publikovaného obsahu RPP, publikovaných v úložišti veřejných číselníků a datových sad RPP.
- Zpřístupnění katalogizačních záznamů k jejich katalogizaci v NKOD.
- Zprovozněné API poskytující katalogizační záznamy o jednotlivých datových sadách s veřejnými číselníky a s obsahem údajů vedených v RPP.

6.1.5 Katalog uživatelů

- Katalog uživatelů dat eviduje, jaké veřejné údaje registrované v RPP čerpají konkrétní OVM a SPUÚ a pro výkon jakých agend ve smyslu § 54 odst. 1 písm. c) zákona č. 111/2009 Sb. o základních registrech.
- Registraci čerpání veřejných údajů provádí OVM či SPUÚ za účelem získávání notifikací o změnách v datových sadách, ve kterých jsou veřejné údaje publikovány.
- Správcem katalogu uživatelů dat je MV ČR.
- Registrace je prováděna prostřednictvím uživatelského rozhraní AIS Působnostní v rámci ohlášení agendy.

Nutné předpoklady a podmínky:

- Správná a úplná evidence veřejných údajů v RPP.

Funkce:

- Evidence registrací čerpání veřejných údajů jednotlivými OVM a SPUÚ.

6.1.6 Notifikační hub

- Nástroj zajišťující s pomocí katalogu uživatelů dat notifikační službu, která informuje subjekty registrované v katalogu uživatelů dat z VDF o změnách veřejných údajů, vedoucích následně i ke změnám v souvisejících publikovaných datových sadách.
- Notifikace probíhá automaticky při ohlášení změny veřejného údaje v publikovaných datových sadách ze strany poskytujícího ISVS.



- Notifikační hub je implementován na bázi mezinárodního standardu [W3C Recommendation WebSub](#).
- Správcem Notifikačního hubu je MV ČR.

Nutné předpoklady a podmínky:

- Funkcionalita a naplnění katalogu uživatelů.

Funkce:

- Evidence informací o změnách veřejných údajů zasílaných systémy pro zpřístupňování údajů.
- Odeslání notifikace o změnách veřejných údajů a číselníků systémům pro čtení údajů z VDF, registrovaným v katalogu uživatelů dat z VDF.

6.1.7 NKOD – Národní katalog otevřených dat

- V NKOD jsou evidovány katalogizační záznamy o všech datových sadách dostupných ve VDF.
- Za katalogizaci datových sad je zodpovědný publikující ISVS, který u svého LKOD poskytuje API dle otevřené formální normy rozhraní katalogů otevřených dat. Prostřednictvím API poskytuje katalogizační záznamy o svých publikovaných datových sadách evidované v LKOD.
- Metadata v katalogizačních záznamech obsahují informace (odkazy) k přístupu k distribucím datových sad.
- Správcem NKOD je MV ČR.

Nutné předpoklady a podmínky:

- Registrace všech lokálních katalogů v NKOD.

Funkce:

- Správa katalogizačních záznamů všech datových sad dostupných ve VDF.
- Zpřístupnění katalogizačních záznamů datových sad dostupných ve VDF pomocí metadat popisujících datové sady.
- Pravidelné načítání katalogizačních záznamů ze všech lokálních katalogů VDF prostřednictvím API dle [otevřené formální normy rozhraní katalogů otevřených dat](#).
- Uživatelské rozhraní pro vyhledávání katalogizačních záznamů publikovaných datových sad ve VDF a s pomocí metadat nalezených datových sad umožnění přístupu k jejich distribucím.
- Rozhraní pro automatickou komunikaci se "Systémem pro čtení údajů z VDF":
 - stažení souboru s obsahem NKOD v podobě datové sady s IRI [https://data.gov.cz/zdroj/datové-sady/00007064/fecee27b4a44fa89cbe8cc491fe5540c](https://data.gov.cz/zdroj/datove-sady/00007064/fecee27b4a44fa89cbe8cc491fe5540c),
 - čtení prostřednictvím SPARQL endpoint <https://data.gov.cz/sparql>,
 - zpřístupnění dat především pro webové aplikace prostřednictvím rozhraní GraphQL, způsob použití uveden v článku: [https://data.gov.cz/clanky/rozhraní-graphql-národního-katalogu-otevřených-dat](https://data.gov.cz/clanky/rozhrani-graphql-narodniho-katalogu-otevrenych-dat).

6.2 Komponenty jednotlivých agend (AIS)

Komponenty jednotlivých agend jsou zaměřeny na postupné zpřístupňování a užívání veřejných údajů jednotlivých agend, jsou zajišťovány a spravovány jednotlivými OVM, správci agend, správci AIS.

Způsob fyzické realizace logických komponent (zajištění požadovaných funkcí a činností) závisí pouze na rozhodnutí jednotlivých OVM.



6.2.1 Systém pro zpřístupnění údajů

- Aktivně v pravidelných intervalech definovaných správcem AIS exportuje obsah veřejných údajů do podoby datových sad a dávkově je předává do úložiště datových sad.
- Systém pro zpřístupnění údajů je logickou součástí každého AIS, prakticky se může jednat o samostatný systém nebo to může být modul v rámci existujícího systému.

Nutné předpoklady, podmínky:

- Příslušná agenda musí být ohlášena v RPP včetně všech jejích údajů v potřebné úrovni granularity.

Prvotní (přípravné) funkce a činnosti:

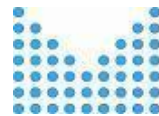
- Výběr údajů pro publikaci do VDF.
- Získání obsahu veřejných údajů z datového kmene poskytovatele.
- Rozdělení obsahu veřejných údajů do vhodných datových sad.
- Vytvoření kompletních katalogizačních záznamů o datových sadách, včetně doplnění metadat o distribucích vytvořených úložištích datových sad
- Katalogizace katalogizačních záznamů v LKOD (případně v NKOD), přenos údajů z LKOD do NKOD probíhá automatizovaně.
- Ohlášení referenční IRI datové sady (datových sad) v NKOD k příslušné agendě v RPP, ve které (kterých) je veřejný údaj zpřístupněn.
- Agendové údaje, které jsou kódované číselníky, ohlašovatel agendy označí jako údaje “kódované konkrétní verzí veřejného číselníku”, a uvede IRI datové sady v NKOD, která obsahuje publikovanou příslušnou verzi veřejného číselníku.

Základní (automatické) funkce:

- Rozdělení obsahu veřejných údajů do vhodných datových sad.
- Převod údajů do navržených distribucí datových sad odpovídajících datovým schémátům uvedeným v Datové specifikaci nebo v OFN:
 - OFN jsou uvedeny v repositáři OFN,
 - v případě, že pro publikované údaje zatím Datová specifikace neexistuje, musí ji správce AIS pro zpřístupnění údajů nejprve vytvořit.
- Dávkové předání datových sad do úložiště datových sad.
- Vytvoření kompletních katalogizačních záznamů o datových sadách, včetně doplnění metadat o distribucích vytvořených úložištích datových sad
- Katalogizace metadatových záznamů v LKOD (případně v NKOD), přenos údajů z LKOD do NKOD probíhá automatizovaně.

Rozšiřující (automatické) funkce:

- Odeslání informace o změně obsahu datové sady notifikačnímu hubu (neplatí pro první publikaci datové sady).



6.2.2 Systém pro čtení veřejných údajů z VDF

- Systém pro čtení údajů čte veřejné údaje z VDF jako otevřená data prostřednictvím ISGOD v podobě distribucí datových sad v různých formátech definovaných Datovými specifikacemi a OFN.
- Systém pro čtení veřejných údajů z VDF je logickou součástí každého AIS, prakticky se může jednat o samostatný systém nebo to může být modul v rámci existujícího systému.
- Cílem čtení obsahu datových sad v podobě datového souboru je typicky aktualizace vlastní kopie údajů přebíraných z VDF v datovém kmene agentury.
- Aktualizaci provádí v pravidelných intervalech definovaných správcem AIS nebo na základě notifikací o změnách obsahu datových sad, které zasílá notifikační hub na základě registrace v katalogu uživatelů.

Prvotní (přípravné) funkce a činnosti:

- Nalezení datových sad s publikovaným veřejným údajem a výběr nejvhodnější datové sady a distribuce.

Základní (automatické) funkce:

- Zjištění IRI datové sady s požadovaným veřejným údajem (IRI datové sady je uvedeno jako součást metadata o veřejném údaji v RPP).
- V případě údaje kódovaného číselníkem zjištění IRI datové sady s požadovaným veřejným číselníkem (IRI datové sady je uvedeno jako součást metadata o veřejném údaji v RPP).
- Čtení katalogizačního záznamu z NKOD pro zadané IRI datové sady obsahující požadované údaje.
- Výběr vhodné distribuce datové sady a odpovídajícího identifikátoru IRI.
- Stažení datové sady s požadovanými veřejnými údaji prostřednictvím ISGOD na základě zjištěných IRI distribucí z NKOD – povinná varianta.
- Aktualizace údajů v příslušné části datového kmene agentury.
- Zahájení čtení veřejných údajů z VDF prostřednictvím IS GOD na základě stanoveného intervalu aktualizace kopie údajů.

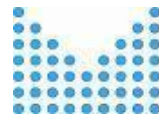
Rozšiřující (automatické) funkce:

- Přístup k jednotlivým položkám datových sad voláním služeb ISGOD umožňujících přistoupit k položce na základě jejího referenčního IRI. - (volitelné)
- Dotazování nad položkami datových sad voláním dotazovacích služeb ISGOD. - (volitelné)
- Zahájení čtení veřejných údajů z VDF na základě přijaté informace od notifikačního hubu o změně datové sady obsahující veřejný údaj.

6.2.3 Úložiště datových sad

- Úložiště datových sad slouží k ukládání publikovaného obsahu registrovaných veřejných údajů vedených v daném publikujícím AIS (ISVS) v podobě distribucí datových sad.
- Každý datový soubor je dostupný na jednom URL, které je stejné pro VDF i veřejný internet.
- Úložiště datových sad daného publikujícího ISVS není novým ISVS, ale je součástí publikujícího ISVS (AIS).
- Každý AIS má své vlastní úložiště a jeho správcem je správce ISVS.

Nutné předpoklady a podmínky:



- Pro daný typ publikovaných údajů musí existovat, nebo být připravena Datová specifikace, která je vytvořena na základě prvků repozitáře OFN, bez takové Datové specifikace není možné údaje prostřednictvím VDF zpřístupnit.

Základní funkce:

- Získání příslušné Datové specifikace publikovaných údajů z příslušného ISVS nebo použité OFN z repozitáře OFN.
- Kontrola správného formátování předaných datových sad "Systémem pro zpřístupnění údajů".
- Kontrola validity datových struktur předaných datových sad vůči datovým schématům definovaných Datovou specifikací nebo OFN.
- Odeslání chybového hlášení systému pro zpřístupnění údajů v případě špatné syntaxe a ukončení práce s obdrženými datovými sadami.
- Zpřístupnění distribucí datových sad k využití v podobě datových souborů, které jsou dostupné ke stažení prostřednictvím rozhraní ISGOD a také z prostředí veřejného internetu – jedná se o povinný způsob zpřístupnění,
 - URL pro stažení datového souboru je shodné pro přístup prostřednictvím ISGOD i pro přístup pomocí veřejného internetu.
- Odeslání systému pro zpřístupnění údajů potvrzení o úspěšném uložení distribucí.
- Odeslání metadat o vytvořených distribucích ve struktuře a podobě definované otevřenou formální normou pro rozhraní katalogů otevřených dat, jako odkazy na umístění datových sad jsou uvedeny lokální identifikátory v podobě IRI.

Rozšiřující funkce:

- Transformace předané datové sady do všech dalších formátů (distribucí) definovaných Datovou specifikací nebo použitými OFN.
- Uložení výsledku transformací do souborového úložiště, v případě RDF distribucí také do triplestore dle OFN (triplestore = databázový systém pro ukládání RDF dat v podobě trojic).
- Zpřístupnění distribucí datových sad k využití následujícími způsoby:
 - prostřednictvím API umožňující přistoupit ke každé jednotlivé položce datové sady s uvedením jejich referenčních a lokálních IRI (dereferencování identifikátorů zajišťuje směrovací služba pro ISGOD i veřejný internet), specifikace je uvedena v [OFN pro propojená data](#),
 - prostřednictvím API umožňující dotazování nad obsahem uložených datových sad pomocí dotazovacího jazyka SPARQL a zpřístupnění SPARQL endpointu.
- Odeslání oznámení notifikačnímu hubu, že došlo ke změně obsahu datové sady.

6.2.4 LKOD – Lokální katalog otevřených dat

Nutné předpoklady a podmínky:

- Implementované rozhraní katalogu dle otevřené formální normy rozhraní katalogů otevřených dat.

Přípravné (ruční) funkce a činnosti:

- Registrace API LKOD správcem ISVS v NKOD.

Základní funkce:

- Katalogizace katalogizačních záznamů o publikovaných datových sadách.
- Zveřejnění API dle otevřené formální normy rozhraní katalogů otevřených dat.

6.3 Technické zajištění přístupu k VDF pro výkon VS

6.3.1 Referenční rozhraní veřejné správy

Referenčním rozhraním se v souladu s jeho definicí zakotvenou zejména v zákoně č. 365/2000 Sb., o informačních systémech veřejné správy a zákoně č. 111/2009 Sb., o základních registrech fakticky rozumí rozhraní pro uskutečňování vazeb mezi informačními systémy veřejné správy, a to především při realizaci propojeného datového fondu sdílením údajů mezi jednotlivými agendovými informačními systémy formou sdílených služeb. Referenční rozhraní je tedy komunikačním rozhraním pro poskytování a využívání sdílených služeb jednotlivých správců informačních systémů veřejné správy.

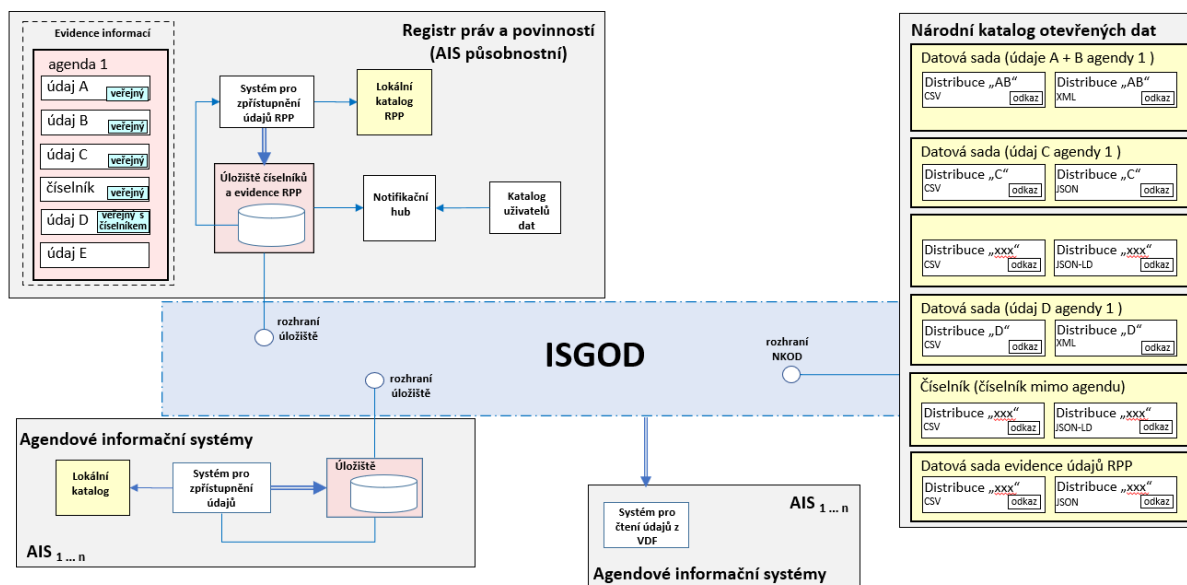
6.3.2 ISGOD – součást referenčního rozhraní

V souladu s výše uvedenou definicí referenčního rozhraní veřejné správy se v případě VDF jedná o rozhraní (ISGOD) zajišťující propojení funkcí jednotlivých komponent VDF (případně AIS, pokud jsou komponenty jejich součástí) za účelem realizace samotného VDF a poskytování souvisejících základních služeb nad veřejně přístupnými údaji organizací veřejné správy.

Jelikož publikované údaje jsou uloženy distribuovaným způsobem v úložištích jednotlivých agendových systémů, je nutné, aby rozhraní zajišťující přístup k uloženým údajům bylo distribuováno stejným způsobem. Z tohoto důvodu jsou komponenty (úložiště jednotlivých AIS, NKOD) poskytující publikované údaje vybavené k tomu potřebnými funkcemi (viz [kapitola 6](#)).

Rozhraní VDF pro přístup k publikovaným údajům je tedy distribuované. Přístupy jsou zajišťovány rozhraními (funkcemi) úložišť AIS a NKOD a rozhraní ISGOD je pouze jejich logickým zastřešením.

ISGOD – referenční rozhraní VDF



Základní služby zajišťované rozhraním ISGOD:

- zpřístupnění veřejných údajů AIS,
- zpřístupnění veřejných číselníků,
- čtení veřejných údajů,
- čtení veřejných číselníků,
- zpřístupnění obsahu RPP,
- čtení obsahu RPP,



- čtení obsahu NKOD,
- dotazovací služby.

Každá služba vyžaduje specifické propojení funkcí jednotlivých komponent VDF. Propojení mezi komponentami je tvořeno komunikací využívající webové standardy.

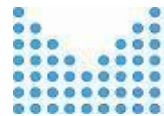
ISGOD umožňuje 3 základní druhy přístupu ke zpřístupněným údajům:

- čtení publikovaného datového souboru (distribuce) s obsahem datové sady,
- čtení publikované položky datové sady na základě jejího referenčního IRI,
- dotazování nad obsahem datové sady voláním dotazovacích služeb.

Služby jsou realizovány jako webové služby postavené na principech REST, které jsou poskytovány jednotlivými komponentami VDF:

- REST služby NKOD umožňují číst metadata o datových sadách a jejich distribucích.
- REST služby úložiště datových sad umožňují číst obsah uložených datových sad v úložišti následujícími způsoby:
 - stahování datových souborů s obsahem uložených datových sad (povinný způsob),
 - přístupu k IRI jednotlivých položek obsahu uložených datových sad (volitelný způsob),
 - SPARQL dotazy nad obsahem uložených datových sad (volitelný způsob).

ISGOD je logickým zastřešením vybraných komponent VDF v závislosti na typu poskytovaných služeb.



7. Příprava datových sad k publikaci

Příprava datových sad k publikaci se skládá z několika na sebe navazujících kroků.

- Legislativní vymezení oblastí a typů informací k publikaci do VDF v podobě otevřených dat, v souladu s možnostmi a povinnostmi vyplývajícími ze stávající legislativy. Ta se samozřejmě v průběhu času vyvíjí a rozšiřuje okruh informací k veřejnému užití.
- Identifikovat ve vymezených oblastech konkrétní agendy a konkrétní množiny údajů k publikaci do VDF v podobě otevřených dat.
- Seskupení vybraných (veřejných) dat k publikaci do smysluplných ucelených datových sad.
- Pro zformované datové sady použít otevřené formální normy k návrhu struktury datové sady a specifikaci jednotlivých datových položek.
- Připravit obsah datové sady ve vybraném základním otevřeném formátu na základě navrženého datového schématu a tuto předat systému pro zpřístupnění údajů.

7.1 Legislativní vymezení oblastí a typů informací k publikaci do VDF

Posláním VDF je zajistit sdílení veřejných informací veřejné správy mezi veřejnoprávními subjekty navzájem a sdílení veřejných údajů mezi veřejnoprávní a soukromoprávní sférou v ČR na základě publikovaných veřejných údajů veřejné správy v podobě otevřených dat. Za veřejné údaje jsou považovány údaje vedené, nebo vytvářené, v rámci agend registrovaných v RPP, a pro které platí, že jsou přístupné veřejnosti.

Výběr těchto údajů aktuálně upravuje a usnadňuje novelizace zákona o svobodném přístupu k informacím č. 106/1999 Sb.

Nový [§ 5a odstavec 2](#) zavádí za účelem následného efektivního využití zveřejněných informací povinnost poskytovat každému dostupné informace obsažené v registrech jako otevřená data.

Pojem „registr“ je legislativní zkratkou založenou v [§ 5a odst. 1](#), která označuje registry, evidence, seznamy nebo rejstříky obsahující informace, které jsou na základě zákona každému přístupné, přičemž se jedná jen o takové registry, které jsou vedené nebo spravované na základě zákona, tedy nikoli dobrovolně vedené registry (které si povinné subjekty vedou samy o svém rozhodnutí).

Jinými slovy, [§ 5a odst. 2](#) zakládá povinnost poskytovat informace, které jsou na základě zákona veřejné, a které jsou vedené v zákonem založených nebo předvídaných registrech také jako otevřená data.

Jedná se o projev zásady „standardního zpřístupnění dokumentů ve formě otevřených dat“ stanovené bodem 16 preambule a čl. 5 odst. 2 [Směrnice 2019/1024](#) (tzv. „Open Data směrnice“).

[§ 5a odst. 2](#) nahradil předchozí právní úpravu povinných otevřených dat spočívající v nařízení vlády č. [425/2016 Sb.](#), které provádělo § 4b odst. 2 zákona č. 106/1999 Sb. ve znění pozdějších předpisů v účinnosti do 24. 7. 2021. Nové ustanovení věcně nepředstavuje velký rozdíl, protože i do nařízení vlády č. [425/2016 Sb.](#) mohly být vybírány pouze informace, které již byly na základě zákona veřejně dostupné a byly vedeny v zákonem založených předvídaných registrech.

Z toho důvodu platí i nadále povinnost, která se vztahovala na informace uvedené nařízení vlády č. [425/2016 Sb.](#) Povinné subjekty, které spravují informace, na které tato povinnost od července 2021 dopadla nově mají na jejich zveřejnění lhůtu do 31. 12. 2023 a to na základě [Čl. II zákona č. 261/2021 Sb.](#), kterým se mění některé zákony v souvislosti s další elektronizací postupů orgánů veřejné moci.

Při analýze, které informace je potřeba v souladu s [§ 5a odst. 2](#) poskytovat jako otevřená data je možné vyjít z jednoduchého schématu dělení informací, kde lze definovat pomocí následujících 5 kategorií:



- **Kategorie 1** - Informace, které již byly součástí nařízení vlády [č. 425/2016 Sb.](#), o seznamu informací zveřejňovaných jako otevřená data, a existuje tedy povinnost jejich poskytování v podobě otevřených dat. Tato povinnost platí i nadále a okamžitě po nabytí účinnosti změnového zákona [č. 261/2021 Sb.](#) Přechodné ustanovení se totiž týká pouze odkladu zveřejnění, ale pokud již informace zveřejněná byla, tato povinnost je naplněna a bylo by proti smyslu zákona, kdyby přechodné ustanovení bylo vykládáno jako možnost odstranění těchto informací a pak jako následná povinnost jejich opětovného zveřejnění.
- **Kategorie 2** – Informace, které povinný subjekt musí poskytovat jako otevřená data podle [§ 5a odst. 2](#) (přechodné ustanovení zakládá lhůtu na splnění této povinnosti do 31. 12. 2023). Jedná se o takové informace, u nichž již existuje primární povinnost jejich poskytování povinnými subjekty a které jsou zároveň vedené v zákonem založených nebo předvídaných registrech. Jinými slovy, informace vedené v zákonem založených nebo předvídaných registrech, které má již povinný subjekt povinnost poskytovat, budou povinně poskytována rovněž jako otevřená data. Zároveň je třeba nezapomenout na nutné provedení pseudonymizace v souladu s výjimkou uvedenou v [§ 5a odst. 2](#).
- **Kategorie 3** – Informace, které je možné poskytovat jako otevřená data na základě diskrece povinného subjektu (v souladu s ustanovením [§ 5 odst. 5 zákona č. 106/1999 Sb.](#), o svobodném přístupu k informacím). Jde tedy o informace, u kterých neexistuje právní povinnost jejich zveřejnění, nebo o informace, které jsou sice zveřejňovány na základě právní povinnosti, ale nejsou vedeny v zákonem založeném, nebo předvídaném registru. Takové informace může povinný subjekt zveřejnit, pokud se tak na základě vlastní diskrece rozhodne. Tyto informace je však nezbytné poskytovat v kvalitě, kterou předepisuje [§ 4b](#), která se otevřeným datům přibližuje.
- **Kategorie 4** – Informace, které není možné poskytovat jako otevřená data proto, že je jejich poskytování vyloučenou výjimkou uvedenou v [§ 8a](#) nebo v [§ 9-11](#) zákona [č. 106/1999 Sb.](#), případně jiným zvláštním zákonem.
- **Kategorie 5** – Informace, které jsou zcela vyňaty z působnosti zákona [106/1999 Sb.](#), o svobodném přístupu k informacím. Právní úprava otevřených dat podle tohoto zákona se tak na ně přímo nevztahuje a pro jejich poskytování je třeba se řídit jejich zvláštními právními předpisy. Jde například o informace spadající pod působnost zákona [č. 123/1998 Sb.](#), nebo o informace katastru nemovitostí.

7.2 Identifikace konkrétních agend a konkrétních množin údajů

K přesnému vymezení veřejných údajů konkrétních agend, aktuálně odpovídajících novelizaci zákona o svobodném přístupu k informacím č. 106/1999 Sb. (viz. [7.1. Nasměrování identifikace veřejných informací k publikaci](#)), je nutné znát podrobnější informace o samotných agendách a jejich údajích.

K získání potřebných informací se nabízí několik způsobů.

- Poskytovatel otevřených dat má dokonalou znalost provozovaných agend a spravovaných údajů agendou, proto dokáže snadno identifikovat všechny veřejné údaje splňující požadavky legislativy (viz. [7.1](#)).
- Poskytovatel otevřených dat prostřednictvím uživatelského rozhraní RPP může v evidenci RPP dohledat všechny potřebné údaje o jím provozovaných agendách a spravovaných veřejných údajích. Nutnou podmínkou je ale řádná a úplná evidence údajů v RPP, která ale není v současnosti splněna. Tuto možnost mohou tedy využít pouze ti poskytovatelé, jejichž agendy mají evidenci v RPP v pořádku. Komplikací může být skutečnost, že v RPP zatím schází informace uvedené u [rozšířené evidence údajů](#) pro potřebu VDF. Z tohoto pohledu by bylo žádoucí, aby ohlašovatelé agend tuto rozšiřující evidenci doplnili jako první krok při přípravě



publikace. V opačném případě je nutné posuzovat všechny údaje požadavky legislativy ([viz. 7.1](#)).

- Poskytovatel otevřených dat může dohledat všechny potřebné údaje o jím provozovaných agendách a spravovaných veřejných údajích prostřednictvím publikovaných datových sad s obsahem evidence RPP. Katalogizační záznamy jsou k dispozici v NKOD s využitím klíčového slova "[RPP](#)", datové sady jsou k dispozici v různých formátech včetně uvedení SPARQL Endpointu. I v tomto případě je ale nutnou podmínkou řádná a úplná evidence údajů v RPP, jako u předchozího způsobu.
- Poslední možností je pracné zjišťování a hledání informací o provozovaných agendách u správce ISVS a v legislativě definující agendu.

7.3 Rozdělení a seskupení dat k publikaci do ucelených datových sad.

Nalezené veřejné údaje agendy, které budou předmětem publikace, je nutné seskupit do ucelených a smysluplných datových sad.

Při návrhu a vytváření datových sad (a souvisejících Datových specifikací) je vhodné brát v úvahu:

- počet entit s atributy v údajích agend a jejich vzájemné vazby – složité datové sady vedou ke komplikovaným Datovým specifikacím (datovým schémátům a konceptuálním modelům), a také k pracnému použití prvků repozitáře OFN,
- definice a striktní vymezení tematických oblastí zákonem ustavujícím příslušnou agendu - např. zákon může přesně vymezovat (oddělovat) vedené rejstříky, seznamy a evidence,
- smysluplný kontext – i samostatná datová sada by měla být cíleně použitelná a měla by dávat smysl (nabízí se použití konceptuálního modelování, případně využití již hotových konceptuálních modelů),
- cíl a použití publikované datové sady – vymezuje kořenovou entitu a její atributy, předurčuje jednoduchost / složitost její struktury,
- množství publikovaných atributů – hodně atributů vzniká v důsledku vazeb na jiné objekty (entity), v takových případech je vhodnější údaje rozdělit do několika datových sad (například podle jednotlivých entit) a tím zmenšit datové struktury, a přitom publikovat všechny atributy,
- množství publikovaných dat – obtížnější práce s velkými datovými soubory.

Jednoduchá ukázka využití publikovaných datových sad s obsahem RPP

- Pro vyhledání detailů o agendách je možné využít datovou sadu "Agendy" (anotace datové sady: "Agendy evidované v Registru práv a povinností ve smyslu § 51 zákona č. 111/2009 Sb. o základních registrech.")
- Katalogizační záznam datové sady je dostupný v NKOD prostřednictvím [odkazu](#).



OTEVŘENÁ DATA

Vzdělávání Články Datová kvalita Datové sady Poskytovatelé Klíčová slova Další

Agendy

Ministerstvo vnitra

Agendy evidované v Registru práv a povinností ve smyslu § 51 zákona č. 111/2009 Sb. o základních registrech.

registr práv a povinností agenda činnost rpp

Téma Vláda a veřejný sektor	Související geografické území Česká republika	Dokumentace Zobrazit dokumentaci	Periodičita aktualizace denní
Klasifikace dle EuroVoc výkonná moc místní orgány státní správy správa věcí veřejných veřejná správa orgány veřejné správy veřejná služba územní samospráva regionální orgány státní správy samospráva výkonný orgán právní moc výkonné moci právní moc instituce		Kontaktní bod Help desk správy základních registrů (SZR) Specifikace OPN: RPP - agendy	

Distribuce datové sady

JSON-LD

Podmínky užití distribuce

Neobsahuje Autorské dílo

Neobsahuje Originální datábase

Není chráněna Zřízením právo pořizovatele datábase

Neobsahuje Osobní údaje

Soubor ke stažení

Stáhnout

Schéma

application/ld+json

Typ média

JSON

Podmínky užití distribuce

Neobsahuje Autorské dílo

Neobsahuje Originální datábase

Není chráněna Zřízením právo pořizovatele datábase

Neobsahuje Osobní údaje

Soubor ke stažení

Stáhnout

Schéma

application/json

Typ média

SPARQL Endpoint

Podmínky užití distribuce

Neobsahuje Autorské dílo

Neobsahuje Originální datábase

Není chráněna Zřízením právo pořizovatele datábase

Neobsahuje Osobní údaje

Datová služba

Popis endpointu

Endpoint

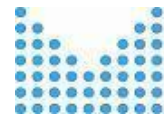
SPARQL dotazování

Specifikace

- Pro ukázkou práce s datovými sadami jsou použity jejich distribuce ve formátu JSON a jejich vizualizace je provedena v nástroji MS Excel.
- Zobrazení obsahu staženého souboru s datovou sadou a ukázkou výběru některých informací například k agendě A 1041 Vnitrozemská plavba:

název.cs	hlavní-ustanovení.označení	ustanovení.označení	vykonávající-ovm	ohlašovatel	vznik	platnost-od	poslední-změna
Vnitrozemská plavba	zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě, ve znění pozdějších předpisů	zákona č. 42/2015 Sb. - Vyhláška o způsobilosti osob k vedení a obsluze plavidel	orgán-veřejné-moci/66003008	orgán-veřejné-moci/66003008	2012-04-17	2019-10-10	2019-10-09
Vnitrozemská plavba	zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě, ve znění pozdějších předpisů	zákona č. 42/2015 Sb. - Vyhláška o způsobilosti osob k vedení a obsluze plavidel	orgán-veřejné-moci/00003352	orgán-veřejné-moci/66003008	2012-04-17	2019-10-10	2019-10-09
Vnitrozemská plavba	zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě, ve znění pozdějších předpisů	zákona č. 222/1995 Sb., o vodních cestách, plavebním provozu v přístavech, společné havárii a dopravě nebezpečných věcí, ve znění pozdějších předpisů	orgán-veřejné-moci/66003008	orgán-veřejné-moci/66003008	2012-04-17	2019-10-10	2019-10-09
Vnitrozemská plavba	zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě, ve znění pozdějších předpisů	zákona č. 222/1995 Sb. - Vyhláška č. 222/1995 Sb., o vodních cestách, plavebním provozu v přístavech, společné havárii a dopravě nebezpečných věcí, ve znění pozdějších předpisů	orgán-veřejné-moci/00003352	orgán-veřejné-moci/66003008	2012-04-17	2019-10-10	2019-10-09
Vnitrozemská plavba	zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě, ve znění pozdějších předpisů	zákona č. 223/1995 Sb. - Vyhláška č. 223/1995 Sb., o způsobilosti plavidel k provozu na vnitrozemských vodních cestách, ve znění pozdějších předpisů	orgán-veřejné-moci/66003008	orgán-veřejné-moci/66003008	2012-04-17	2019-10-10	2019-10-09
Vnitrozemská plavba	zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě, ve znění pozdějších předpisů	zákona č. 223/1995 Sb. - Vyhláška č. 223/1995 Sb., o způsobilosti plavidel k provozu na vnitrozemských vodních cestách, ve znění pozdějších předpisů	orgán-veřejné-moci/00003352	orgán-veřejné-moci/66003008	2012-04-17	2019-10-10	2019-10-09
Vnitrozemská plavba	zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě, ve znění pozdějších předpisů	zákona č. 67/2015 Sb. - Vyhláška o pravidlech plavebního provozu (pravidla plavebního provozu)	orgán-veřejné-moci/66003008	orgán-veřejné-moci/66003008	2012-04-17	2019-10-10	2019-10-09
Vnitrozemská plavba	zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě, ve znění pozdějších předpisů	zákona č. 67/2015 Sb. - Vyhláška o pravidlech plavebního provozu (pravidla plavebního provozu)	orgán-veřejné-moci/00003352	orgán-veřejné-moci/66003008	2012-04-17	2019-10-10	2019-10-09
Vnitrozemská plavba	zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě, ve znění pozdějších předpisů	zákona č. 46/2015 Sb. - Vyhláška o stanovení vodních nádrží a vodních toků, na kterých je zakázána plavba plavidel se spalovacími motory, a o rozsahu a podmínkách užívání povrchových vod k plavbě	orgán-veřejné-moci/66003008	orgán-veřejné-moci/66003008	2012-04-17	2019-10-10	2019-10-09

- Ohlašovatel agendy:
 - orgán-veřejné-moci, IČO 66003008 - Ministerstvo dopravy.
- Vykonavatel agendy:
 - orgán-veřejné-moci, IČO 66003008 - Ministerstvo dopravy.
 - orgán-veřejné-moci, IČO 00003352 - Státní plavební správa.



- Informace o údajích spravovaných agendou lze například zjistit v datové sadě “[Subjekty a objekty údajů a jejich údaje](#)”, katalogizační záznam v NKOD:

OTEVŘENÁ DATA

Vzdělávání Články Datová kvalita Datové sady Poskytovatelé Klíčová slova Další ▼

Subjekty a objekty údajů a jejich údaje

Ministerstvo vnitra

Typy subjektů a objektů práva a typy údajů o nich vedených v agendách evidované v Registru práv a povinností ve smyslu § 51 zákona č. 111/2009 Sb. o základních registrech.

subjekt práva rpp registru práv a povinností údaj objekt práva

Téma
Vláda a veřejný sektor
Klasifikace dle EuroVoc
výkonný orgán
právní moc
právní instituce
výkonná moc
místní orgány státní správy
správa věcí veřejných
veřejná správa
orgány veřejné správy
veřejná služba
územní samospráva
regionální orgány státní správy
samospráva

Související geografické území
Česká republika

Dokumentace
Zobrazit dokumentaci
Kontaktní bod
Help desk správy základních registrů (SZR)
Specifikace
OFN: RPP - údaje

Periodicita aktualizace
denní

Distribuce datové sady

JSON

Podmínky užití distribuce
Soubor ke stažení
Stáhnout
Schéma
application/json
Typ média

JSON-LD

Podmínky užití distribuce
Soubor ke stažení
Stáhnout
Schéma
application/ld+json
Typ média

SPARQL Endpoint

Podmínky užití distribuce
Soubor ke stažení
Stáhnout
Schéma
application/ld+json
Typ média

Datová služba
Popis endpointu
Endpoint
SPARQL dotazování
Specifikace

- Ukázka zobrazení obsahu staženého souboru a filtrace údajů pro agendu A1041:

type	id	kód	cs	agenda	id.1	kód-údaje	název údaje	komentář	typ-úaje	ustanovení-právního-předpisu
objekt-subjekt	objekt-subjekt/1041-1	1041-1	Plavidla jiná než malá	agenda/A1041	údaj/1041-1-1	1041-1-1	Vlastník plavidla	Fyzická, právnická osoba vlastníci plavidlo	Nereferenční	§ 14 odst. 2 písm. a zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-1	1041-1	Plavidla jiná než malá	agenda/A1041	údaj/1041-1-2	1041-1-2	Provozovatel plavidla	Fyzická, právnická osoba provozující plavidlo	Nereferenční	§ 14 odst. 2 písm. b zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-1	1041-1	Plavidla jiná než malá	agenda/A1041	údaj/1041-1-3	1041-1-3	Poznávací znaky	Poznávací znaky plavidla stanovené prováděcím právním předpisem	Nereferenční	§ 14 odst. 2 písm. c zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-1	1041-1	Plavidla jiná než malá	agenda/A1041	údaj/1041-1-4	1041-1-4	Jméno loď	Jméno loď	Nereferenční	§ 14 odst. 2 písm. d zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-1	1041-1	Plavidla jiná než malá	agenda/A1041	údaj/1041-1-5	1041-1-5	Druh plavidla, oblast jeho provozu na vodní cestě a účel, pro který je plavidlo určeno	Druh plavidla, oblast jeho provozu a účel, pro který je určeno	Nereferenční	§ 14 odst. 2 písm. e zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-1	1041-1	Plavidla jiná než malá	agenda/A1041	údaj/1041-1-6	1041-1-6	Nosnost, dovolený počet cestujících a ostatní základní technické údaje	Nosnost, dovolený počet cestujících a ostatní základní technické údaje	Nereferenční	§ 14 odst. 2 písm. f zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-1	1041-1	Plavidla jiná než malá	agenda/A1041	údaj/1041-1-7	1041-1-7	Zástavní právo k plavidlu	Zástavní právo	Nereferenční	§ 14 odst. 2 písm. g zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-1	1041-1	Plavidla jiná než malá	agenda/A1041	údaj/1041-1-8	1041-1-8	Datum zápisu a výmazu z plavebního rejstříku	Datum zápisu a výmazu z plavebního rejstříku	Nereferenční	§ 14 odst. 2 písm. h zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-2	1041-2	Malá plavidla	agenda/A1041	údaj/1041-2-1	1041-2-1	Vlastník plavidla	Fyzická, právnická osoba vlastníci plavidlo	Nereferenční	§ 15a odst. 2 zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-2	1041-2	Malá plavidla	agenda/A1041	údaj/1041-2-2	1041-2-2	Provozovatel plavidla	Fyzická, právnická osoba provozující plavidlo	Nereferenční	§ 15a odst. 2 zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-2	1041-2	Malá plavidla	agenda/A1041	údaj/1041-2-3	1041-2-3	Poznávací znaky	Poznávací znaky plavidla stanovené prováděcím právním předpisem	Nereferenční	§ 15a odst. 2 zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-2	1041-2	Malá plavidla	agenda/A1041	údaj/1041-2-5	1041-2-5	Druh plavidla, oblast jeho provozu na vodní cestě a účel, pro který je plavidlo určeno	Druh plavidla, oblast jeho provozu a účel, pro který je určeno	Nereferenční	§ 14 odst. 2 písm. e zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-2	1041-2	Malá plavidla	agenda/A1041	údaj/1041-2-6	1041-2-6	Nosnost, dovolený počet cestujících a ostatní základní technické údaje	Nosnost, dovolený počet cestujících a ostatní základní technické údaje	Nereferenční	§ 15a odst. 2 zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-2	1041-2	Malá plavidla	agenda/A1041	údaj/1041-2-8	1041-2-8	Datum zápisu a výmazu z plavebního rejstříku	Datum zápisu a výmazu z plavebního rejstříku	Nereferenční	§ 14 odst. 2 písm. h zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě
objekt-subjekt	objekt-subjekt/1041-3	1041-3	Vodní cesty a provoz na vodních cestách	agenda/A1041	údaj/1041-3-1	1041-3-1	Informace vedené v systému říčních informačních služeb	Informace zahrnují informace o vodní cestě, provozu na vodní cestě, mimořádných událostech na vodní cestě, plavidlech a přepravovaném nákladu a o přístavních poplatcích.	Nereferenční	§ 32a odst. 1 zákona č. 114/1995 Sb. - Zákon o vnitrozemské plavbě

V uvedeném příkladu je zřejmé, že se jedná o jednoduchý rejstřík vedený na základě [Zákona č. 114/1995 Sb.](#) o vnitrozemské plavbě, přičemž objekt 1 “Plavidla jiná než malá” je evidován na základě [§ 14](#) zákona, objekt 2 “Malá plavidla” je evidován dle [§ 15a](#).

Současně je v odstavci [§ 14 odst. 2](#) uvedeno, že plavební rejstřík je veřejný seznam.



Na první pohled je tedy evidentní, že údaje se týkají dvou entit, a tak se nabízí i údaje rozdělit do dvou datových sad:

Datová sada 1 “Plavidla jiná než malá”.

- Kořenová entita:
 - Plavidlo jiné než malé.
- Atributy:
 - Vlastník plavidla.
 - Provozovatel plavidla.
 - Poznávací znaky.
 - Jméno lodě.
 - Druh plavidla, oblast jeho provozu na vodní cestě a účel, pro který je plavidlo určeno.
 - Nosnost, dovolený počet cestujících a ostatní základní technické údaje.
 - Zástavní právo k plavidlu.
 - Datum zápisu a výmazu z plavebního rejstříku.

Datová sada 2 “Malá plavidla”.

- Kořenová entita:
 - Plavidlo malé.
- Atributy:
 - Vlastník plavidla.
 - Provozovatel plavidla.
 - Poznávací znaky.
 - Druh plavidla, oblast jeho provozu na vodní cestě a účel, pro který je plavidlo určeno.
 - Nosnost, dovolený počet cestujících a ostatní základní technické údaje.
 - Datum zápisu a výmazu z plavebního rejstříku.

Bylo by samozřejmě možné publikovat všechny údaje v jedné datové sadě “Plavidla” s jednou kořenovou entitou “plavidlo” a přidáním nového atributu pro rozlišení plavidel, ale mohlo by to zkomplikovat práci s daty. Navíc v tomto případě by to nebylo ani v souladu se [Zákonem č. 114/1995 Sb.](#), neboť první datová sada vychází z definice [§ 14](#) (Plavební rejstřík), druhá datová sada z [§ 15a](#) (Rejstřík malých plavidel).

Při návrhu datových sad je vždy nutné přistupovat ke každé tematické oblasti individuálně a brát v úvahu již výše uvedená [poznámky](#).

7.4 Příprava a použití prvků repozitáře otevřených formálních norem pro publikaci údajů

Nutnou podmínkou pro publikaci veřejných údajů do VDF je používání a dodržování konceptu otevřených formálních norem pro každou publikovanou datovou sadu. Prakticky to znamená přípravu Datové specifikace publikované sady s využitím standardizovaných prvků datových schémat dostupných v [repozitáři OFN](#).

Základním východiskem pro návrh a přípravu Datové specifikace je výsledek rozdělení a seskupení dat k publikaci do ucelených datových sad. Tento výsledek předurčuje složitost a pracnost tvorby Datové specifikace publikované datové sady.

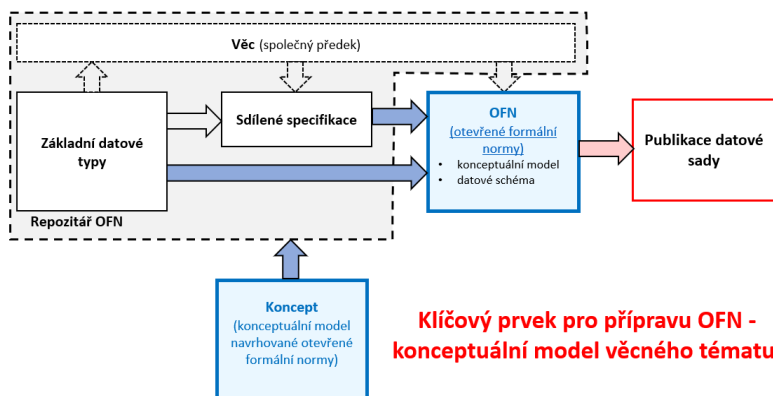
Příprava Datové specifikace ve zkratce:

- Základní princip vychází z principu tvorby OFN a spočívá v transformaci konceptu (konceptuálního modelu) do datového modelu pomocí standardizovaných stavebních prvků (sdílených specifikací a základních datových typů) obsažených v repozitáři OFN.
- Výchozím stavebním prvkem jsou základní datové typy, pomocí kterých je vytvořen společný předek (věc) a také jsou základem vytvořených sdílených specifikací. Tyto sdílené specifikace

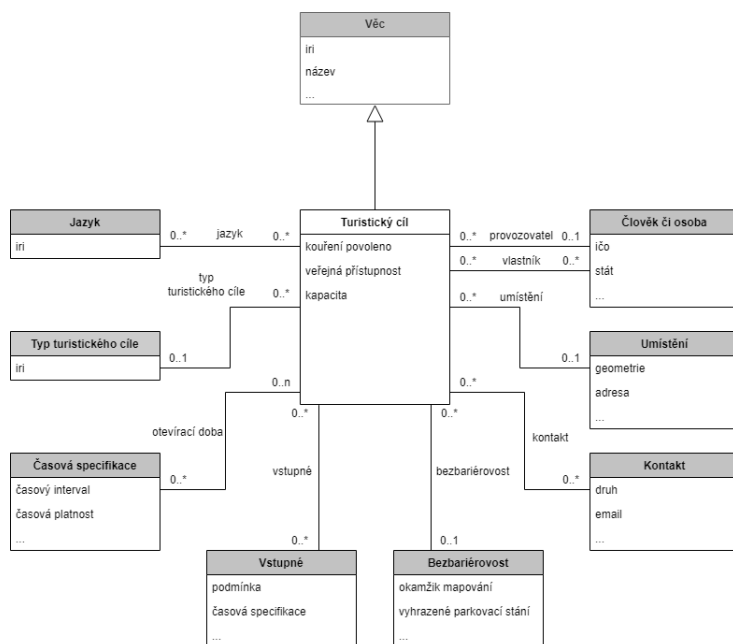
také přebírají vlastnosti svého předka (Věc) a všechny vytvořené stavební prvky datových schémát jsou obsahem [repositáře OFN](#).

- Sdílené specifikace standardizují podobu často se opakujících částí dat.

Přípravu otevřených formálních norem principiálně zachycuje následující schéma, a téměř totožný princip je použitý i pro přípravu Datové specifikace konkrétní datové sady:



Při tvorbě konkrétní OFN i Datové specifikace konkrétní datové sady sehrává klíčovou roli téma oblasti publikovaných veřejných údajů (nebo publikované datové sady) a jeho významový kontext. Příklad zachycení významového kontextu datové sady:



V případě publikace dat do VDF je příprava významového kontextu publikovaných údajů specifická důvodu, neboť je omezena dostupnými veřejnými údaji agentury a způsobem rozdělení a seskupení dat určených k publikaci do datových sad, viz. [7.3 Rozdělení a seskupení dat k publikaci do ucelených datových sad](#).

Z uvedených důvodů je pro publikaci dat v prvním období implementace VDF doporučený zjednodušený postup pro přípravu datových sad k publikaci založený na:

- jednoduchém vytvoření významového konceptu a
- využití stávajícího repositáře OFN.

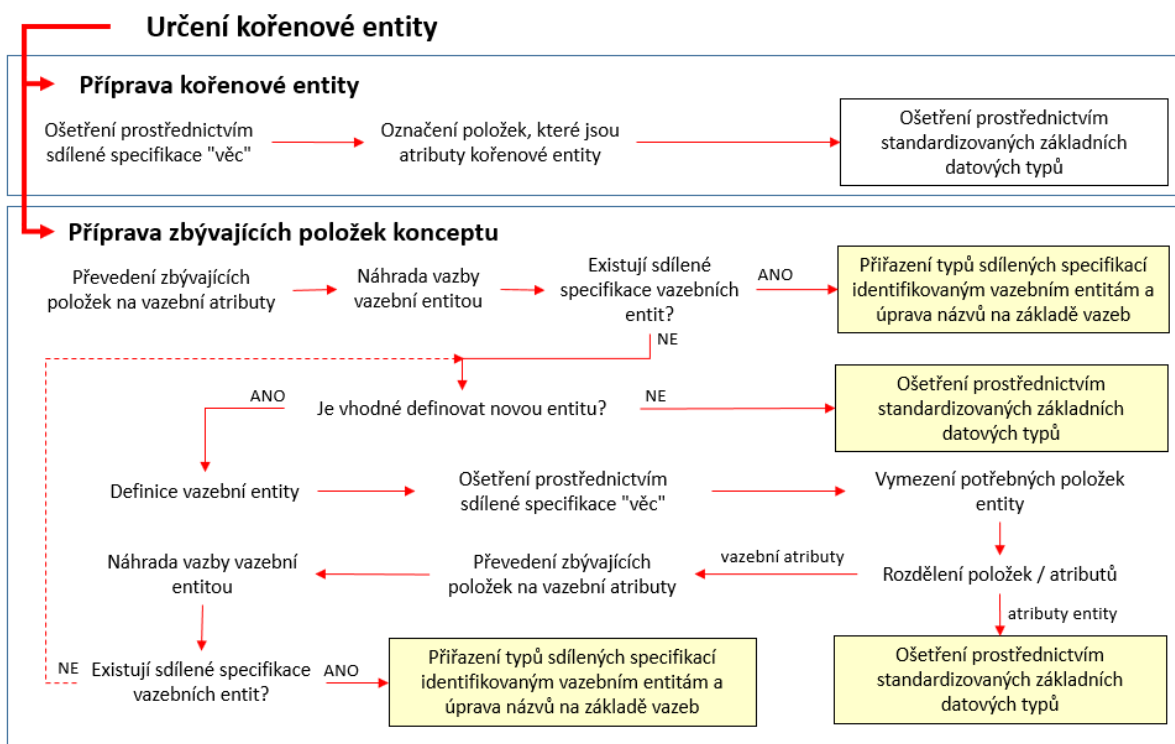
Zjednodušený postup vytvoření významového konceptu z identifikovaných dat k publikaci:

- Identifikace a příprava kořenné entity:

- určení kořenové entity připraveného konceptu (schématu) - určuje téma datové sady,
- kořenová entita bude většinou vycházet ze subjektů nebo objektů práva agendy, jejíž veřejná data budou publikována,
- ošetření kořenové entity prostřednictvím sdílené specifikace "věc",
- označení položek, které jsou atributy kořenové entity.
- Příprava zbývajících položek konceptu:
 - převedení zbývajících položek na vazební atributy a pojmenování vazby,
 - specifikace entity, na kterou vazba směřuje.
- Ošetření entit navázaných vazbami na kořenovou entitu:
 - nahrazení navázaných entit sdílenými specifikacemi (pokud jsou k dispozici) a úprava jejich názvů na základě vazeb.
- Ošetření entit navázaných vazbami na kořenovou entitu a neodpovídajících sdíleným specifikacím:
 - ošetření navázané entity prostřednictvím sdílené specifikace "věc",
 - označení položek, které jsou atributy entity,
 - převedení zbývajících položek na vazební atributy a pojmenování vazby,
 - specifikace entity, na kterou vazba směřuje.
 - nahrazení navázaných entit sdílenými specifikacemi (pokud jsou k dispozici) a úprava jejich názvů na základě vazeb.

Vytvořený významový koncept publikované datové sady a jeho následné zpracování pomocí repositéře OFN až do detailní specifikace jeho jednotlivých prvků, je výchozím obsahovým základem dokumentace datové sady, a také její Datové specifikace.

Schematický pohled na uvedenou tvorbu významového konceptu a detailní specifikace jeho jednotlivých prvků zachycuje následující diagram:





Pokračování praktické ukázky příkladu z kapitoly 7.3

Pokračování ukázky příkladu z kapitoly [7.3 Rozdělení a seskupení dat k publikaci do ucelených datových sad](#) v případě navržené datové sady 1.

Datová sada 1 “Plavidla jiná než malé”.

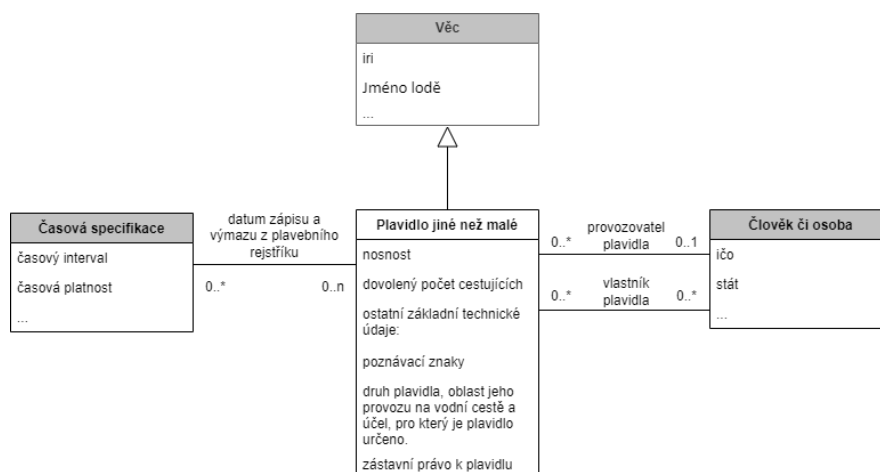
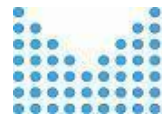
- Kořenová entita:
 - Plavidlo jiné než malé.
- Atributy:
 - Vlastník plavidla - [§ 14a odst. 1c.](#)
 - Provozovatel plavidla - [§ 14a odst. 1d.](#)
 - Poznávací znaky - [§ 18.](#)
 - Jméno lodě - [§ 14a odst. 1a.](#)
 - Druh plavidla, oblast jeho provozu na vodní cestě a účel, pro který je plavidlo určeno - [§ 14 odst. 2e, § 10d odst. 1e.](#)
 - Nosnost, dovolený počet cestujících a ostatní základní technické údaje - [§ 10d odst. 1f, § 14 odst. 2f.](#)
 - Zástavní právo k plavidlu - [§ 14 odst. 2g.](#)
 - Datum zápisu a výmazu z plavebního rejstříku - [§ 14 odst. 2h.](#)

K přesnému posouzení typu jednotlivých údajů, například zda se jedná skutečně pouze o vlastnosti kořenové entity, nebo zda jde spíše o formu nějakého vztahu na jiné entity je nutná dobrá znalost spravované agendy i zákona, který agendu ustavuje (z tohoto důvodu jsou u jednotlivých údajů uvedeny i konkrétní odkazy na příslušný zákon).

Názorným příkladem složitosti může být údaj “Druh plavidla, oblast jeho provozu na vodní cestě a účel, pro který je plavidlo určeno”. Bez znalosti způsobu a rozsahu dalších vedených informací v odpovídajícím ISVS, je vytvoření významového konceptu problematické. Z příslušného zákona lze vyčíst, že například oblast provozu plavidla na vodní cestě je závislá na vydaném osvědčení plavidla, které se odvíjí od technického stavu plavidla, dále že existuje evidence těchto osvědčení a také vazby na evidenci osvědčení EU apod. Bez podrobnější znalosti celé problematiky je ale orientace v údajích složitější a pouhá dostupnost veřejných údajů k zachycení úplného významového kontextu zcela určitě nebude stačit.

Využití konceptu OFN je proto především zaměřeno na poskytnutí dostatečného kontextu pro zachycení významu zejména publikovaných dat a neklade si za cíl úplné zachycení celé věcné problematiky.

Z tohoto důvodu je i následující ukázka finálního konceptu zjednodušená a určena pouze pro demonstraci přístupu k využití konceptu OFN na základě identifikovaných údajů k publikaci a přípravě Datové specifikaci.



7.5 Příprava obsahu datové sady ve vybraném základním otevřeném formátu

Na základě připravené Datové specifikace (prakticky to znamená významového konceptu a navrženého schématu datové sady), je nutné:

- fyzicky identifikovat údaje v ISVS
- připravit skripty pro transformaci údajů do navržené datové struktury
- validovat výslednou datovou sadu vůči strojově čitelnému schématu datové sady.



8. Externí odkazy a standardy

8.1 Metodiky, pravidla, existující OFN

Metodika evidence údajů v RPP – metodika bude zpřístupněna až bude dokončena.

[Otevřená formální norma pro propojená data.](#)

[Otevřené formální normy](#) - repositář OFN.

[Otevřená formální norma pro číselníky.](#)

[Otevřené formální norma pro rozhraní katalogů otevřených dat.](#)

[Pravidla pro tvorbu IRI.](#)

8.2 Komunikační standardy komponent

Komunikace Notifikační hub – Úložiště datových sad (AIS)

- <https://www.w3.org/TR/websub/>

Komunikace Katalog uživatelů – Systém pro čtení údajů z VDF (AIS)

- <https://ofn.gov.cz/rozhraní-katalogů-otevřených-dat/>

Komunikace NKOD – Systém pro čtení údajů z VDF (AIS)

- <https://ofn.gov.cz/rozhraní-katalogů-otevřených-dat/>

9. Zdroje a terminologie

- ZoZR - [Zákon č. 111/2009 Sb., o základních registrech](#), ve znění zákona č. 100/2010 Sb. a zákona č. 424/2010 Sb.
- InfoZ - [Zákon č. 106/1999 Sb. o svobodném přístupu k informacím](#)
- ZoISVS - [Zákon č. 365/2000 Sb. o informačních systémech veřejné správy](#) a o změně některých dalších zákonů.
- [Slovník pojmů eGovernmentu](#)