Financováno Evropskou unií NextGenerationEU

NÁRODNÍ PLÁN OBNOVY

DIGITÁLNÍ A INFORMAČNÍ AGENTURA_

METODIKA POPISU DAT

vytvořeno v rámci projektu

Zajištění podmínek pro kvalitní správu datového fondu a zajištění řízeného přístupu k datům

Reg.č.: CZ.31.1.01/MV/23_62/0000062

DIGITÁLNÍ A INFORMAČNÍ AGENTURA
NA VÁPENCE 915/14, 130 00 PRAHA 3
POSTA@DIA.GOV.CZ

Historie verzí

Verze	Datum vydání	Popis
1	6. 1. 2025	První verze

1 O této metodice 5
1.1 Důvody pro vznik metodiky 5
1.2 Metodika a otevřená formální norma pro slovníky 6
1.3 Návaznost na existující metodiky a školení 6
2 Motivace popisu dat 7
2.1 Návaznost na správu dat 8
3 Využití popisu dat 9
3.1 Výkladový slovník 9
3.2 Popis datových sad 9
3.3 Evidence údajů agendy do Registru práv a povinností 9
3.4 Tvorba datových specifikací 9
4 Základy popisu dat 10
4.1 Pojmy11
4.2 Slovníky 12
5 Tvorba popisu dat 13
5.1 Úrovně popisu dat 13
5.2 Proces popisu dat 13
5.3 Výběr oblasti pro popis dat 16
5.4 Role potřebné pro popis dat 16
5.5 Zdroje popisu dat 16
5.5.1 Používání databázových modelů pro popis dat 16
5.6 Nástroje pro tvorbu popisu dat 17
5.6.1 Tabulka pro popis dat 18
5.6.2 Archi 18
5.6.3 Enterprise Architect 18
5.6.4 Výrobní linka 18
5.7 Uvedení příkladu 18
5.8 Výběr pojmů 19
5.9 Subjekty a objekty práva 19
5.10 Charakteristiky pojmu 20
5.11 Nadřazené pojmy 21
5.12 Nejednoznačně definované pojmy 21
5.13 Vlastnosti 22
5.14 Vztahy 23
5.15 Využití pojmů z jiných slovníků 25

5.16 Kompletní příklad slovníku 25
5.17 Pokročilý popis dat 27
5.18 Publikace slovníku 27

1 O TÉTO METODICE

Tato metodika je určena těm pracovníkům (manažerům, věcným znalcům daných oblastí, IT odborníkům a dalším), kteří ve veřejné správě chtějí zavést, spravovat, nebo využívat popis dat.

Metodika nad základy počítačové gramotnosti (ovládání webu a kancelářských aplikací) nepředpokládá další technické znalosti, ale zkušenost s prací s daty je výhodou.

Aktuální verze metodiky je společně s aktuálními informacemi a podpůrnými materiály k popisu dat umístěna na Portálu o datech:

https://data.gov.cz/popis-dat

Metodika se dělí na následující kapitoly:

Motivace popisu dat charakterizuje motivaci pro takové popisování dat, které je srozumitelné nejen pro lidi, ale i pro stroje. Zákonná motivace pro úřady pouze podtrhuje motivaci věcnou, která přináší užitek klientům i úředníkům veřejné správy. Kapitola zároveň informuje o souvislostech se správou dat, strategií a legislativou v této oblasti.

Využití popisu dat ukazuje, jak se dá vytvořený popis dat využít pro naplnění zákonných povinností a zefektivnění digitalizace úřadů způsoby popsanými v následujících kapitolách.

Základy popisu dat uvádí stěžejní principy popisu dat. Dozvíte se, s čím přesně se v popisu dat pracuje – slovníky a pojmy popisující subjekty a objekty práva, vlastnosti a vztahy. V této kapitole je také vysvětleno, jak zmíněné prvky spoluvytvářejí efektivní popis dat.

Tvorba popisu dat je hlavní část metodiky. Poskytuje návod pro popis dat, od sbírání zdrojů přes vytvoření slovníku až po jeho publikaci.

V dokumentu rozlišujeme doplňující poznámky a příklady dobré i špatné praxe:

Poznámky budou uváděny ve žlutých plných rámečcích.

Příklady dobré praxe budou uváděny v zelených přerušovaných rámečcích.

Příklady špatné praxe budou uváděny v červených tečkovaných rámečcích.

Pro tuto metodiku je doprovodným školením Popis dat veřejné správy, jehož záznam je dostupný ve vzdělávací sekci Portálu o datech.

1.1 Důvody pro vznik metodiky

Nová metodika vznikla nejen jako reakce na zpětnou vazbu na dosavadní procesy popisu dat (viz 1.3 Návaznost na existující metodiky a školení), ale také z důvodu zásadního vylepšení a zajištění souladu popisu dat s relevantními projekty a aktivitami. Konkrétně má metodika následující cíle:

- **Zjednodušení metodických pokynů:** Dosavadní metodický postup se zaměřoval na tvorbu úplného konceptuálního modelu a nezabýval se postupným vznikem popisu dat s ohledem na míru jeho detailu a rozsahu. Tato metodika je naopak postavena na „úrovních“ popisu z hlediska jeho podrobnosti. Podle nich se pak určují možnosti využití tohoto popisu. Tvůrci si tak mohou sami stanovit, jak chtějí popisu využít a podle toho vyplňovat podrobnosti pojmů/slovníků.
- **Vytvoření opory pro Strategii pro správu dat ve veřejné správě ČR a připravovaný zákon o správě dat a řízeném přístupu k datům:** Popis dat je v rámci projektu *Zajištění podmínek pro kvalitní správu datového fondu a zajištění řízeného přístupu k datům* (Reg. č.: CZ.31.1.01/MV/23_62/0000062, dále uváděno jako *projekt správy dat*) uveden ve Strategii pro

správu dat ve veřejné správě ČR (schválená vládou v usnesení č. 287 z 30. dubna 2024) jako opatření 1.3.1 a v připravovaném zákoně o správě dat jako nutný předpoklad pro kvalitní správu dat. V souvislosti s tímto projektem je tedy potřeba vytvoření metodiky, která přímo navazuje na relevantní opatření a aktivity projektu správy dat a je s nimi plně v souladu.

> V době vzniku této verze textu je zákon ve verzi odeslané Legislativní radě vlády. Veškeré reference se tak odkazují na zmíněnou verzi zákona. Zákon je sepsán na základě věcného záměru, schváleného vládou v usnesení č. 576 z 16. srpna 2023.

- **Podpora různých způsobů tvorby i využívání slovníků:** Dosavadní provedení popisu dat vedlo v podstatě k nutnosti využívat pro tvorbu popisů tzv. Výrobní linku – bylo sice možné používat také další nástroje, ale tyto možnosti byly kvůli technické implementaci složité. V této metodice chceme proto zohlednit využívání dalších nástrojů (tabulkové nástroje, Archi, Enterprise Architect). Dále bychom rádi podpořili vznik nástrojů určených nejen k tvorbě slovníků, ale také k jejich používání za pomoci zveřejněné otevřené formální normy pro slovníky (viz níže). Zároveň chceme objasnit, jaká data přesně požadujeme od tvůrců slovníku, aby se zlepšila interoperabilita, používaly se slovníky/pojmy napříč veřejnou správou a byly dodržovány FAIR principy.

1.2 Metodika a otevřená formální norma pro slovníky

Slovníky vytvářené podle této metodiky jsou v souladu s návrhem otevřené formální normy pro slovníky (OFN). Tato norma určuje obsah slovníků podle jednotlivých úrovní popisu (odvíjejících se podle možností využití takového popisu, blíže popsaných dále v kapitole 3 Využití popisu dat) a minimální standard, který musí popis splnit, aby se dal považovat za slovník. Datum formálního vydání OFN bude dáno zákonem o správě dat, nicméně již nyní doporučujeme slovníky tvořit podle metodiky a OFN.

> OFN určuje minimální standard, tj. je možné si OFN rozšířit o prvky, které ve vašem případě potřebujete. Myslete však na to, že již existující prvky OFN nelze ve vámi rozšířené verzi měnit kvůli zachování kompatibility. Přidané prvky pak nebudou oficiálně podporovány a mohou se tedy stát nekompatibilními s budoucími verzemi této OFN.

1.3 Návaznost na existující metodiky a školení

V průběhu předchozích aktivit zaměřených na představení popisu dat, především v rámci projektu Rozvoj datových politik v oblasti zlepšování kvality a interoperability dat veřejné správy (Reg. č.: CZ.03.4.74/0.0/0.0/15_025/0013983, dále jen projekt KODI), byly vytvořeny metodiky, nástroje a procesy pro popisování dat.

Tato metodika je však s předchozími výstupy kompatibilní. Nemusíte tedy mít obavu, že získané znalosti nebo vytvořené slovníky už nebudou platné. Konkrétní rozdíly mezi touto metodikou a ostatními jsou popsány níže:

- Metodika definice údajů vedených v agendě (výstup projektu KODI C2V11) je určena pro evidenci údajů agendy v Registru práv a povinností a z hlediska postupu tvorby slovníku je plně kompatibilní s touto metodikou. Nicméně, jak je zmíněno výše, vznikl požadavek na zjednodušení popisu dat a zohlednění dalších způsobů využití slovníků než jen k evidenci údajů agend. Všechna doporučení a příklady pro tvorbu popisu dat agendy obsažená v C2V11 jsou přímo použitelná i zde.
- Metodika tvorby a údržby konceptuálních datových modelů agend (výstup projektu KODI C5V4) se nahrazuje touto metodikou, která oproti výstupu C5V4 popis zjednodušuje a aktualizuje v souvislosti s vydáním otevřené formální normy pro slovníky. Kurz Modelování významu dat ve veřejné správě (výstup projektu KODI C4V5) ukazuje, jak se dá příkladový popis podle C5V4 vyjádřit pomocí prototypů Výrobní linky. Tento kurz je stále platný.

2 MOTIVACE POPISU DAT

Kvalitní správa dat veřejné správy představuje základ pro to, aby je bylo možné sdílet a využít pro poskytování služeb veřejnosti, k naplňování klíčových zákonných povinností, plánování politik státu nebo k tvorbě analytických materiálů potřebných pro výkon veřejné správy. Data nelze kvalitně spravovat bez znalosti jejich obsahu a významu.

Představte si například, že máte za úkol zjistit „kolik vozidel je průměrně na silnicích Prahy 5 každé ráno“ k analýze toho, jak jsou využívány aktuální kapacity pozemních komunikací. Ta by měla sloužit k budoucímu plánování dopravní infrastruktury. Takový úkol zní na první pohled jednoduše – možná se vám i rychle podaří najít relevantní datové sady, ale poté se dozvíte, že se v těchto datech „vozidly“ myslí vozidla podle zákona č. 361/2000 Sb., ale vy jste chtěli informace o motorových vozidlech, a „Praha 5“ v datové sadě není městský obvod, ale znatelně větší městská část. Výsledkem by tak byly zavádějící podklady pro analýzu, což by při rozhodování mohlo vést k nepřiměřenému přidělování zdrojů na určité projekty nebo oblasti.

> Vozidlo je typickým příkladem slova, kdy se v běžné mluvě myslí jedna věc (motorová vozidla) ale v zákoně jiná – vozidlo je podle zákona č. 361/2000 Sb. i kolo. Řidič podle tohoto zákona může být i cyklista, ale typicky si pod tímto výrazem lidé představí řidiče motorového vozidla. Další zákony definují „vozidlo“ i jinak.

Problematika hledání správných dat pro své potřeby však není pouze zodpovědností uživatelů, kteří data přebírají, ale také poskytovatelů. Jenom oni dokáží popsat, o čem nebo o kom (např. o řidiči podle zákona/řidiči jako vlastníkovi řidičského průkazu) přesně datové sady jsou.

Základní součástí kvalitní správy dat je tedy jejich dobrá znalost ze strany úřadu, který tato data shromažďuje. Právě k tomu, aby měl o svých datech přehled, věděl, jaká data sbírá, k čemu je využívá a mohl je sdílet dle potřeb s ostatními úřady, je nezbytně nutné, aby si svá data kvalitně popsal (věděl, o kom/o čem přesně jsou). Při dodržení pravidel pro popis dat tak daný úřad výrazně sníží možnost, že by tato data byla interpretována nesprávným způsobem. Popis dat je klíčový také pro správnou evidenci údajů v Registru práv a povinností (RPP), kterou jsou úřady ze zákona povinny naplňovat a získají tím oprávnění pro čtení a zápis údajů taktéž evidovaných v RPP. Dalšími příklady zákonných povinností, pro které je popis dat důležitý, jsou:

- popisy publikovaných otevřených dat v Národním katalogu otevřených dat
- a sdílení dat v rámci Propojeného datového fondu (navazuje na evidenci do RPP).

Každý úřad vykonávající agendu si potřebuje pro své efektivní fungování udělat pořádek ve svých datech, pro digitalizaci agend to platí dvojnásob. Mimo jiné i proto je důležité, aby byl popis dat strojově čitelný. Pokud se takto popíšou data, která „tečou“ informačními systémy a rozhraními, pomůže se tím k zajištění dohledatelnosti a využitelnosti dat napříč systémy – i těch, které jsou ve správě třetích stran.

Popis dat má i další benefity, například:

- Zachová se tak co největší množství informací o agendě úřadu, procesech, správě dat apod. Úřad si tím navíc zajistí větší nezávislost na dodavatelích a zabrání tomu, aby dodavatelé měli nepřiměřeně velkou kontrolu nad daty a procesy agend.
- Interní znalosti úřadu má smysl s popisem dat udržovat i pro případ odchodu/pracovní neschopnosti klíčových zaměstnanců nebo pro zaškolení nových zaměstnanců.
- Mít popsaná data je důležité také při řízení kvality a rizik, která s nimi souvisí, například z hlediska bezpečnostního managementu a nakládání s osobními údaji – pokud nebudou mít úřady o svých datech přehled a nebudou je dobře znát, nebudou schopny tato rizika identifikovat a svá data efektivně chránit.

2.1 Návaznost na správu dat

V případě, že úřad naplňuje standard správy dat tak, jak jej uvádí Strategie pro správu dat ve veřejné správě ČR (2024—2030) (dále jen Strategie) v části 4.1.1, bude zaručena spolehlivost a důvěryhodnost jím spravovaných dat, která je nezbytným předpokladem pro další práci s nimi a také pro jejich sdílení. Oblasti popisu dat se věnuje Strategie podrobněji ve specifickém cíli 1.3 Zajištění dohledatelnosti, srozumitelnosti a kategorizace dat, respektive v opatření 1.3.1 Popsat data a definovat jejich význam.

Toto opatření cílí na to, aby úřad vytvořil a následně také udržoval aktuální popis struktury a významu dat v informačních systémech dle standardů a metodických doporučení DIA, a to alespoň ve všech prioritních oblastech v tzv. *datových* slovnících. Konkrétně by měl úřad vytvářet datové slovníky, které jsou katalogizovány v lokálním katalogu dat (viz opatření 1.3.2 Strategie) v podobě datové sady podle odpovídajícího standardu (otevřené formální normy pro slovníky). Tento katalog doplňuje k popisu dat další informace nutné pro jejich dohledatelnost a práci s nimi. Naplnění tohoto opatření bude podle navrhovaného zákona o správě dat a řízeném přístupu k datům povinné pro subjekty definované tímto zákonem.

Z pohledu správy dat jsou popis dat a jeho pravidelná aktualizace důležité i při udržování přehledu o datových potřebách, a to jak uvnitř samotného úřadu, tak i ze strany dalších úřadů (opatření 1.2.2 Strategie). Díky popisu a tím i dobré znalosti dat může úřad identifikovat příležitosti k získávání dat z jiných agend nebo úřadů a také zjistit, kde jsou možnosti využití dat pro lepší digitalizaci poskytovaných služeb veřejnosti nebo efektivnější výkon jeho procesů. Zároveň díky tomu, že budou mít úřady svá data popsána, mohou reagovat na požadavky na získání, využívání nebo sdílení dat, které k nim přicházejí ze strany dalších úřadů nebo odborné veřejnosti pro výzkumné a analytické účely (opatření 1.4.1). Nezanedbatelnou roli hraje popis dat také při řízení změn informačních systémů veřejné správy (této oblasti se věnuje opatření 1.4.2 Strategie), ve kterých se data shromažďují a uchovávají.

Pro popis dat má význam stanovit v úřadě také role, které jsou odpovědné za data (opatření 1.1.1 Stanovit odpovědnost za data úřadu a jejich správu Strategie). Klíčová je role datového architekta, který drží a rozvíjí celkový přehled o datech úřadu. Vytvoření a udržování popisu dat by proto mělo být jedním z jeho prvořadých zájmů, měl by jej koordinovat a vést po odborné stránce. Aby byl popis přesný a odpovídal tomu, jaký je význam popisovaných dat, je nutné zapojit i toho, kdo má věcnou znalost dané datové oblasti. Konkrétně jde o tyto osoby:

- zástupce vedení úřadu v roli, která je manažersky a věcně odpovědná za data v konkrétní věcné oblasti („vlastník dat“),
- zástupce úřadu v roli, která je věcně odpovědná za data v konkrétní věcné oblasti a na každodenní bázi vykonává činnosti související s jejich věcnou správou, např. kontrolu jejich správnosti a úplnosti („věcný správce dat“).

3 VYUŽITÍ POPISU DAT

Tvorba slovníků a jejich publikace (podle procesu v kapitole 5 Tvorba popisu dat) otevře vám a ostatním uživatelům vašich slovníků řadu možností, jak slovníky využívat. Slovníky se využívají především v rámci různých aplikací, které je podporují nebo jsou pro jejich potřeby přímo postaveny. Mohou to být například:

- nástroje pro vyhledávání pojmů (výkladové slovníky),
- nástroje pro tvorbu datových schémat (kontextů, výměnných formátů),
- katalogy (nástroje pro vyhledávání katalogizačních záznamů popsaných slovníky),
- nástroje pro tvorbu formulářů pro vkládání dat popsaných slovníky,
- nástroje pro koordinaci poptávky pro daná data daného úřadu popsaná slovníky,
- nástroje pro tvorbu slovníků.

Některá využití jsou blíže rozepsána níže.

3.1 Výkladový slovník

Podobně jako výkladové slovníky v papírové/textové podobě se dají vzniklé slovníky využít pro shromáždění, zachování a prezentaci klíčových pojmů a znalostí daných oblastí. Na rozdíl od těchto „klasických“, strojově nečitelných slovníků je možné ve slovnících vytvořených podle této metodiky hledat strojově podle různých kritérií (i napříč slovníky), libovolně organizovat jejich pojmy, upravovat jejich prezentaci apod.

3.2 Popis datových sad

S pojmy slovníků se dá popisovat, o kom nebo o čem vybraná data jsou. Taková informace je zvláště výhodná u datových sad publikovaných přes datová rozhraní (tj. katalogizační záznamy těchto datových sad). Naplňují podobnou funkci, jako „klíčová slova“ katalogizačního záznamu datové sady, ale na rozdíl od pouhých klíčových slov jsou pojmy rozšířeny právě o definici a obecný kontext. Potenciální uživatel tak dokáže rychle najít žádanou datovou sadu a ověřit, jaká data přesně (o kom, o čem) daná datová sada poskytuje a jaké jsou možné návaznosti na jiné sady či oblasti.

3.3 Evidence údajů agendy do Registru práv a povinností

Slovník se dá při naplnění určitých minimálních podmínek využít jako podklad pro evidenci do RPP. Celý proces evidence včetně požadavků a příkladů je podrobněji popsán na stránkách https://data.gov.cz/popis-dat.

Jak je uvedeno v sekci 1.3 Návaznost na existující metodiky a školení, tato metodika je kompatibilní s metodikou definice údajů vedených v agendě. V případě, že se rozhodnete vytvořit konceptuální model agendy, jsou doporučení z metodiky definice údajů přímo použitelná i pro tvorbu slovníků.

3.4 Tvorba datových specifikací

Slovníky mohou být použity pro tvorbu formátu, podle kterého si informační systémy vyměňují data, jak například dokazuje popis otevřených dat RPP a několik otevřených formálních norem. Po vytvoření slovníku obsahujícího všechny pojmy, které se ve formátu mají použít, je možné jej využít pro schéma mezi vašimi/spřízněnými informačními systémy. Příkladem mohou být:

- otevřené formální normy (viz https://ofn.gov.cz),

- kontexty pro sdílení dat v propojeném datovém fondu (PPDF); předpokladem pro zveřejnění kontextu je evidence do RPP, více viz průvodce na adrese https://pruvodcepripojenim.gov.cz/pruvodci/publikace-kontextu-na-isss.

4 ZÁKLADY POPISU DAT

Klíčové koncepty, které budeme dále používat, si přiblížíme na následujícím příkladu formuláře. V Evidenční kartě řidiče (v obrázku níže) nalezneme jméno, příjmení a další osobní údaje řidiče, spolu se seznamem jemu vydaných oprávnění, uvedením jejich druhu, data zkoušky a dalšími údaji.

EVIDENČNÍ KARTA ŘIDIČE

Obecní úřad obce s rozšířenou působností	Adresa:	
Příjmení:	Jméno:	Titul:
Datum a místo narození:	Okres:	
Rodné číslo:	/	
Adresa trvalého pobytu:		

VYDANÉ OPRÁVNĚNÍ

Druh oprávnění	Číslo a série řidičského průkazu	Číslo protokolu o zkoušce	Datum zkoušky	Pořadové č. knihy vyd. ŘP	Poznámky (omezení apod.)

Lékařská prohlídka platí do:	
Zvláštní záznamy:	
Datum vyplnění karty:	Podpis:

Pro každého řidiče je tato karta vyplněna jinak, zatímco formulář zůstává stejný. Např. políčko „Příjmení" může obsahovat hodnotu „Novák". Tato hodnota „Novák" je konkrétní příklad příjmení. Podobně „B" je příkladem druhu (řidičského) oprávnění.

> Nevyplněný formulář (šablona) obsahuje typy údajů (jméno, příjmení, rodné číslo). Vyplněný formulář obsahuje navíc konkrétní údaje (Jan, Novák, 7808243949).

Formuláře jsou názorné, ale jejich volnost umožňuje jejich tvůrcům dělat chyby, například

- překlepy,
- pojmenování významově stejných políček v různých formulářích různě, např. „Adresa trvalého pobytu", „Trvalá adresa",
- pojmenování významově různých políček v různých formulářích stejně, např. „Jméno" (řidiče), „Jméno" (vlastníka),
- reprezentovat v jediném políčku více informací nebo informaci jedinou, např. „Datum narození", „Místo narození" vs. „Datum a místo narození".

To komplikuje nejen údržbu těchto formulářů, ale i následné zpracování dat. Proto v dalších částech zavedeme prvky, které nám pomohou tyto problémy odstranit.

4.1 Pojmy

Ve formulářích se každé políčko vztahuje k „někomu“ nebo „něčemu“: příjmení *žadatele*, adresa *firmy*, evidenční číslo *faktury*. Častokrát jsou formuláře rozděleny na sekce podle těchto položek, aby se dalo snadno rozlišit mezi vyplňováním např. jména žadatele a jména dítěte.

Pro popis dat je důležité si všimnout následujícího:

- Ke stejným položkám (k „někomu“, „něčemu“) se vážou stejná políčka – tj. každý žadatel má jméno, příjmení, adresu bydliště apod., každé silniční vozidlo má VIN, značku atd., která se využívají napříč formuláři.
- Bez kontextu formuláře se informace o žadateli apod. nedají využít a není jasné, proč je potřeba je uchovávat (a kdo je má uchovávat). Jedná se o informace žadatele o řidičský průkaz, o pracovní pozici na Ministerstvu dopravy, o registraci silničního vozidla?

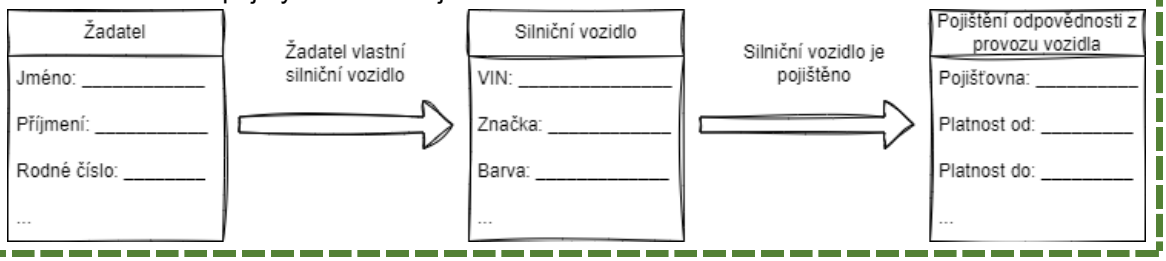
Seskupování dat do „kategorií“ reprezentujících osobu, věc nebo koncept z reálného světa je běžná praxe a jejich standardizovaná evidence s předpřipravenými údaji (políčky) a s propojeními mezi sebou pomocí vazeb je dobrý způsob, jak si v nich udělat pořádek.

Základním nástrojem pro evidenci těchto prvků je **pojem**. Pojem je slovo nebo sousloví opatřené definicí, zdrojem, příp. dalšími charakteristikami, které zpřesňují jeho význam tak, aby se omezila možnost špatně mu porozumět nebo jej špatně použít.

> Nejde tedy o samotné slovo nebo sousloví, ale právě o daný význam, který je tímto slovem nebo souslovím pojmenován.

Strukturováním dat do pojmů získáváme pomyslné šablony „kousků“ formulářů. Tyto „kousky“ se pak dají skládat podle toho, jaký vztah mají jednotlivé pojmy mezi sebou.

Formulář Žádost o zápis silničního vozidla do registru silničních vozidel toto dobře ilustruje: *Žadatel* se v rámci formulářů opakuje často, a často se od něj chtějí ty stejné údaje – což slouží jako perfektní příklad „šablony“. Pro posouzení registrace vozidla se samozřejmě zvažuje i vozidlo, které žadatel vlastní – *Silniční vozidlo*, které má také svá políčka. Obdobně pak i *Pojištění odpovědnosti z provozu vozidla*, jeho políčka a další části, které se poté pomocí vazeb s ostatními pojmy dostává do jednoho formuláře – viz obrázek.



Analogicky je tomu i u dat a jejich správy obecně – data jsou propojena vazbami, které dále doplňují kontext a umožňují podrobnější dotazování (do kdy je platný řidičský průkaz, který patří fyzické osobě Jan Novák?)

Každý pojem může popisovat buď:

- Subjekt práva (osobu, „někoho“) - Žadatel,
- Objekt práva (věc, koncept, „něco“) – Silniční vozidlo,
- Vlastnost subjektu/objektu práva („políčko formuláře“) – Příjmení žadatele,
- Vztah mezi subjekty/objekty práva (věcné propojení) – Žadatel vlastní silniční vozidlo.

Určité pojmy („šablony“) se mohou používat nejen v rámci jedné oblasti nebo jedné organizace, ale i v jiných: například Řidič se hodí nejen pro resort dopravy (registr řidičů), ale také pro policii (delikty při řízení/provozování vozidla) a pojišťovny (pojištění odpovědnosti z provozu vozidla). Zároveň se může stát, že si každý „Řidiče“ vykládá trochu jinak (tj. mají pro něj jinou definici), jak bylo zmíněno v kapitole 2 Motivace popisu dat.

Pro tyto případy je důležité jednak určit tvůrce pojmu, který vlastní i samotná data a také kontext (věcná oblast), ve kterém je daný pojem s danou definicí používán (Řidič jako řidič motorových vozidel a řidič podle definice zákona o silničním provozu). Organizaci pojmů do „balíčků“ podle oblasti a gestora zajišťují **slovníky**.

4.2 Slovníky

Slovníkem rozumíme lidsky i strojově čitelnou jednotku správy pojmů – popisující subjekty práva a objekty práva s jejich vlastnostmi, vztahy a dalšími charakteristikami. Slovník je vytvářen pro určitou oblast. Tou může být například odborná oblast (slovník územního plánování), normativní dokument (slovník zákona č. 283/2021 Sb.), agenda (slovník agendy A1046 Agenda řidičů), nebo třeba datová sada (slovník datové sady eSbírka). Konkrétnímu úřadu (resp. oddělení nebo osobě v ní působící) by měla být přiřazena zodpovědnost za jeho kvalitu, přesnost, a úplnost (vzhledem k dané doméně).

Každý slovník se dá popsat těmito charakteristikami:

- Název (povinný) – označující, jakou oblast slovník popisuje.
- Popis (volitelný) – doplňující poznámka, upřesňující např. vybrané zdroje nebo účely použití slovníku.

V rámci Strategie se v této metodice konkrétně mluví o *datových slovnících* (popsaných mj. v opatření 1.3.1 Strategie), což je takový slovník, který popisuje data vedená v informačních systémech veřejné správy. Pro to, aby slovník splňoval minimální standardy pro popis dat, stačí, aby obsahoval subjekty a objekty práva. Od míry podrobnosti uvedených informací se ale odvíjí využitelnost slovníků pro aktivity zmíněné v kapitole 2 Motivace popisu dat – čím více detailu do slovníku vložíme, tím širší použitelnost od něj můžeme očekávat (další informace viz podkapitola 5.1 Úrovně popisu dat).

5 TVORBA POPISU DAT

V této kapitole se dozvíte, jak tvořit slovníky. Proces této tvorby je navržen tak, aby procházel jednotlivými úrovněmi detailu podle možného využití výsledného slovníku (viz níže). Proces popisu dat ukážeme na komentovaném příkladu. Příklad si projdeme od začátku (sběr zdrojů) do konce (publikace vzniklého slovníku).

V prvé řadě je nutné uvést pro ty, co se popisu dat obávají (buď z hlediska kapacit nebo správnosti), že popis dat je průběžná činnost. Je jasné, že se legislativa a podpůrné dokumenty časem mění – kromě toho je pochopitelné, že není reálné provést vše do největšího detailu ihned bez chyby. Proto vám doporučujeme, abyste si nejdříve popis dat zkusili na malé oblasti (resp. její malé části) v malém detailu, který pak můžete společně se získanými zkušenostmi rozvíjet.

> Nedokončený popis dat je mnohem lepší, než žádný popis dat! Již částečný popis dat totiž umožňuje sdílet jejich význam.

5.1 Úrovně popisu dat

Vytvořením pojmů popisujících subjekty a objekty práva z dané oblasti získáte slovník v jeho minimální podobě. Taková podoba se dá využít jako Výkladový slovník nebo pro základní Popis datových sad. Dodáním vlastností k subjektům a objektům práva se využitelnost rozšiřuje dále na evidenci do RPP (viz 3.3 Evidence do Registru práv a povinností). Dodáním vztahů mezi subjekty a objekty práva vzniká plnohodnotný popis, který je potom využitelný mj. na datové specifikace (viz 3.4 Tvorba datových specifikací).

5.2 Proces popisu dat

Proces je stručně uveden v následující tabulce, která sleduje posloupnost následujících podkapitol. Vedle každého kroku je odkaz na příslušnou část metodiky, kde je tento krok rozebírán ve větším detailu.

> Tabulka popisuje jednu *iteraci* popisu – jak je zmíněno výše, nenalháme na vytvoření kompletního popisu v jediné iteraci.

Proces je v tabulce rozdělen na 4 fáze, přičemž je možné některé kroky provádět souběžně (například evidence subjektů a objektů práva společně s vlastnostmi a vztahy) nebo v jiném pořadí (například výběr nástroje před výběrem oblasti).

Příprava iterace	Vyberte oblast (nebo její část), kterou budete v rámci slovníku popisovat a určete osobu/osoby, která bude naplňovat role důležité pro popis dat (především znalec dané oblasti a tvůrce samotného popisu).	5.3 Výběr oblasti pro popis dat, 5.4 Role potřebné pro popis dat
	V rámci této oblasti vyberte autoritativní zdroje definic pojmů (zákony, dokumentace, metodiky, směrnice apod.).	5.5 Zdroje popisu dat
	Vyberte nástroj, ve kterém budete popis dat vytvářet. Doporučujeme podporované nástroje (viz podkapitola 5.6 Nástroje pro tvorbu popisu dat).	5.6 Nástroje pro tvorbu popisu dat
Základní slovník	Ve zdrojích identifikujte subjekty práva (osoby) a objekty práva (věci, koncepty, skutečnosti), které jsou v oblasti důležité a popisují data evidovaná vaší organizací (kromě provozních údajů) v rámci informačních systémů veřejné správy.	5.8 Výběr pojmů, 5.9 Subjekty a objekty práva
	Ke každému subjektu a objektu práva, pokud je to možné, doplňte následující informace: a) Název: jednoznačný, přesný, bez zkratek a nespisovného jazyka, v prvním pádě jednotného čísla. Píše se v lidsky čitelné podobě; nejde o názvy databázových tabulek. b) Definice: text ze zdroje určující význam pojmu. Pokud zdroj definici nijak neuvádí (pojem pouze zmiňuje), lze definici vynechat a vytvořit popis (bod d). c) Zdroj: nejlépe ve formě strukturovaného, strojově čitelného odkazu na veřejně dostupný web, kde si každý může definici zkontrolovat. Pokud však žádný odkaz neexistuje, lze zdroj uvádět v podobě textu - např. název materiálu se stranou. d) Popis: dále upřesňuje význam pojmu nad rámec definice nebo vysvětluje definici jinými slovy. Může to být vlastní popis tvůrce slovníku, tj. nemusí pocházet ze zdroje.	5.10 Charakteristiky pojmu
	Kde je to vhodné, doplňte nadřazené pojmy. Nadřazený pojem je obecnější vůči pojmům jemu podřazeným, které ho rozdělují na různé „druhy“ (například „Vozidlo“ a „Motorové vozidlo“, „Nemotorové vozidlo“).	5.11 Nadřazené pojmy,
	Splněna úroveň základního slovníku – následující kroky nejsou pro tvorbu slovníku povinné, ale jejich vyplnění rozšiřuje možnosti využití slovníku (viz podkapitola 5.1 Úrovně popisu dat). Můžete se tedy rozhodnout, jestli pokračovat s popisem vlastností a vztahů, nebo tuto iteraci popisu ukončit publikací slovníku podle instrukcí níže v části „Konec iterace“.	5.1 Úrovně popisu dat

Vlastnosti a vztahy (volitelné)	Dále ze zdrojů vyberte vlastnosti; vlastnost se vždy váže na konkrétní subjekt nebo objekt práva. V názvu vlastnosti je i subjekt nebo objekt práva, ke kterému se vlastnost vztahuje – například *Jméno řidiče* namísto *Jméno*. Stejně, jako u subjektů a objektů práva, se i k vlastnostem doplňují výše zmíněné charakteristiky (název, definice, zdroj, popis). ▪ Každá vlastnost popisuje dále nedělitelnou informaci daného pojmu. Vlastnost nesmí kombinovat více informací najednou, tj. nelze vytvořit vlastnost „Jméno a příjmení žadatele o řidičské oprávnění“. Pro kontrolu byste měli ke každé vlastnosti dokázat vymyslet příkladovou hodnotu (například „Jméno“ – „Jan“). ▪ Vlastnostmi nejsou provozní data informačních systémů – identifikátory, databázové klíče apod.	5.13 Vlastnosti
	Dále ze zdrojů vyberte vztahy; vztahy se vždy vážou na dva subjekty nebo objekty práva. Stejně, jako u subjektů a objektů práva, se i ke vztahům doplňují výše zmíněné charakteristiky (název, definice, zdroj, popis). ▪ V názvu vztahu je i alespoň jeden ze dvou pojmů, ke kterému se vztah váže – například *řídí vozidlo* namísto *řídí*. ▪ Vztahy je vhodné využít pro kontrolu rozsahu slovníku. Pokud považujete slovník za hotový (resp. jste přesvědčeni, že zachycuje celou zvolenou oblast/část oblasti) a máte v něm pojem, který není s žádným jiným navázaný (vztahem ani nadřazeností), tak je s velkou pravděpodobností pojem přebytečný, nebo jsou vztahy ve slovníku chybně vyjádřeny.	5.14 Vztahy
Konec iterace	Vytvořte *název* a *popis* slovníku. Z názvu by mělo být jasné, jakou oblast slovník popisuje. Popis obsahuje doplňující poznámky k vybraným zdrojům a/nebo účelům použití.	5.18 Publikace slovníku
	Slovník zaregistrujte do lokálního katalogu dat, aby byl katalogizovaný i v Národním katalogu dat a následně využívejte podle dosažené úrovně popisu.	

5.3 Výběr oblasti pro popis dat

Velikost oblastí se může řídit tím, jaké zdroje jsou pro ni k dispozici. Máte volbu členit oblasti například podle agend, procesů, nebo jednotlivých informačních systémů. Problematika vymezení věcných oblastí se více rozebírá v podpůrném materiálu správy dat 1.1-1: Vymezení a prioritizace věcných oblastí dat. Tyto oblasti doporučujeme dále rozdělit na jednotlivé části – abyste si mohli popis rozvrhnout a publikovat postupně podle vašich schopností a kapacit.

Pro komentovaný příklad jsme si vybrali oblast řidičů, kde budeme popisovat zákon č. 361/2000 Sb. o provozu na pozemních komunikacích a o změnách některých zákonů.

5.4 Role potřebné pro popis dat

Tvorbě slovníku by se měly přinejmenším účastnit role:

- věcného znalce vybrané oblasti,
- samotného tvůrce slovníku – za pomoci zdrojů a konzultací se znalcem vytváří slovník ve zvoleném nástroji.

Obě role může naplňovat tatáž osoba. Popisu dat celého úřadu by se měly také věnovat role stanovené v sekci 2.1 Návaznost na správu dat, které koordinují, zprostředkovávají a zodpovídají za dodržení pravidel kvalitní správy dat.

5.5 Zdroje popisu dat

Hlavním předpokladem pro popis dat je sběr autoritativních zdrojů, ze kterých se pojmy slovníků budou vytvářet. Pojmy vždy mají nějaký zdroj – viz sekce Charakteristiky pojmu. Mezi hlavní druhy zdrojů patří:

- **Ustanovení** zákona, vyhlášky, nebo usnesení. Ta přestavují oblíbený startovní bod pro popis dat, díky kterému je možné určit nejdůležitější pojmy a vztahy dané oblasti. Popis vlastností je však typicky poměrně omezený a někdy jsou ustanovení psána takovým způsobem, který umožňuje popis jen těch nejzákladnějších pojmů.
- **Formuláře**, které jsou velmi nápomocné pro určení vlastností i vztahů. Před jejich evidencí do datových slovníků by měl být potvrzen soulad s daty v informačních systémech veřejné správy.
- **Metodické pokyny, směrnice a řády**, které jsou vhodné pro získání dalšího detailu z praktického využití zákona.
- **Dokumentace informačních systémů.** Dávejte si však pozor, abyste místo popisu věcných subjektů/objektů práva, vlastností a vztahů nepopisovali striktně provozní data anebo obecně aspekty databázového modelu navržené pro konkrétní informační systém – tato specifika do slovníků nepatří.

5.5.1 Používání databázových modelů pro popis dat

Jelikož tvorba slovníků (obzvlášť ve specializovaných nástrojích) připomíná tvorbu logických modelů nebo diagramů tříd a cílem datových slovníků je popsat data ISVS, je možné pro usnadnění popisu dat postupovat „odspodu“ a využít pojmy z existujících databázových schémat – např. názvy a definice tabulek a „sloupců“.

Je ovšem důležité vzít v potaz následující:

- Aby byl popis dat využitelný pro porozumění datům, jejich prezentaci netechnickému publiku a integraci s daty ostatních, musí být pojmy srozumitelné a řádně definované s pomocí zdrojů. Databázová schémata toto většinou nesplňují, navíc mohou obsahovat technická a provozní data,

která nejsou relevantní pro datový popis, případně i duplicity. Je tak velmi pravděpodobné, že budete muset potřebné informace dále hledat v materiálech, které jsou zmíněny v podkapitole výše.

- Záměrem slovníků je i to, aby se mohly do jejich tvorby a využívání zapojit i věcné (netechnicky založené) osoby. Tyto osoby mohou mít lepší věcný vhled do popisované oblasti (resp. dokáží lépe určit zdroje) a vědí, jak věcně navazují na jiné oblasti a jakým způsobem se dají využívat. Technický podklad může pro tyto osoby do popisu představit překážku k efektivní účasti na popisu.
- *Přímá* vazba na konkrétní návrh databáze není žádoucí, jelikož se může často měnit, a to nejen v závislosti na změně věcné podstaty uchovávaných dat - např. aktualizaci/změnách IS. Může podléhat také určitým podmínkám zamezujícím jeho zveřejnění (např. z důvodu ochrany duševního vlastnictví, kybernetické bezpečnosti), nebo může neoddělitelným způsobem zachycovat více nesourodých oblastí.

5.6 Nástroje pro tvorbu popisu dat

Aby byly slovníky strojově čitelné, je potřeba je tvořit v jednotném formátu. Pro slovníky byla proto vytvořena otevřená formální norma. Otevřené formální normy (OFN) jsou definovány v zákoně č. 106/1996 Sb. o svobodném přístupu k informacím. Jedná se o technická doporučení zaměřená na vybrané datové sady (slovníky, turistické cíle, úřední desky apod.), jejichž následování zajišťuje, že se stejná data publikovaná různými poskytovateli mohou jednoduše využívat nezávisle na tom, kdo je poskytl. Více informací naleznete na https://data.gov.cz/ofn/. Nejnovější verze otevřené formální normy pro slovníky (v době vydání této verze metodiky ve verzi draftu) se nachází na https://ofn.gov.cz/slovníky/. Zveřejňování slovníků podle této OFN uloží jako povinnost připravovaný zákon o správě dat.

Tvůrci slovníků ale OFN nemusí nijak studovat. Připravili jsme několik způsobů a nástrojů, ve kterých je možné slovníky tvořit, aniž by bylo potřeba se v OFN orientovat nebo se příliš starat o technickou stránku popisu dat. Jde o:

- tabulkový popis (např. v Excelu),
- nástroj Archi,
- nástroj Enterprise Architect,
- Výrobní linku.

> Jsme si vědomi nedávných inovací v oblasti umělé inteligence (a konkrétně dosažených milníků ve vývoji velkých jazykových modelů - LLM) a zkoumáme možnosti jejího využití pro asistenci při popisu dat. Mezi studované možnosti patří vytvoření podkladů pro slovníky od určitých zdrojů, hledání souvislostí mezi pojmy, ověření souladu výstupů s touto metodikou atd. Při jakémkoliv využití LLM je však nutné mít na paměti, že se může mýlit, ať už kvůli halucinacím (tj. „vymýšlení si“), špatnému zpracování materiálů, přesažení kontextového „okna“ apod. Je tak velmi důležité každý výstup AI před použitím kontrolovat. Případná doporučení pro použití AI budeme sdílet ve znalostní bázi na stránkách popisu dat.

Vedle námi poskytnutých a podporovaných nástrojů lze využít i jiné. Jediný požadavek, který musí nástroje splnit je, že umožňují vytvořit soubor se slovníkem, který je kompatibilní s otevřenou formální normou pro slovníky.

> Aktuální informace, včetně šablon, návodů a doporučení pro jednotlivé nástroje naleznete na domovské stránce popisu dat: https://data.gov.cz/popis-dat.

5.6.1 Tabulka pro popis dat

Tabulkové řešení jsme připravili především pro ty, kteří s popisem dat začínají a v úřadě nemají zkušenosti s nástroji na tvorbu popisu dat v grafické podobě (neboli s diagramy/pohledy). Tabulka je nejjednodušší řešení pro základní popis dat. Nevyžaduje vynaložení speciálních kapacit kromě znalce oblastí, které jste zvolili pro tvorbu slovníků. Berte však na vědomí, že s růstem počtu pojmů, vlastností a vztahů ale roste komplikovanost a klesá srozumitelnost tabulky. Šablony jsou připraveny jednak ve formátu XLSX (pro Microsoft Excel) a CSV.

5.6.2 Archi

Archi je nástroj na modelování v jazyku ArchiMate. Nástroj je zdarma ke stažení i k jakémukoliv využití (podle permisivní licence MIT) a dostupný je pro Windows, Linux i macOS. Jazyk ArchiMate se používá pro vyjádření architektury podniku. Mimo jiné i kvůli jeho otevřenosti a nezávislosti jej využívá odbor Hlavního architekta eGovernmentu (OHA) v žádostech o stanovisko projektů eGovernmentu – více informací na https://archi.gov.cz/uvod_schvalovani. Slovníky tak můžete tvořit i ve vazbě s vaší podnikovou architekturou.

> V žádostech o stanovisko OHA je potřeba uvádět i subjekty a objekty, jejichž data se zpracovávají v plánovaném informačním systému.

Pro Archi jsme vytvořili šablonu pro popis dat, se kterou můžete začít nový projekt, nebo ji naimportujete do vašeho existujícího projektu.

5.6.3 Enterprise Architect

Sparx Enterprise Architect je oproti Archi komerční nástroj, který podporuje několik modelovacích jazyků, včetně ArchiMate a UML. Připravili jsme šablonu a MDG technologii pro popis dat, pomocí které můžete tvořit slovníky.

5.6.4 Výrobní linka

> Výrobní linka je sada *prototypů* a není tedy momentálně možné garantovat stálou dostupnost těchto nástrojů.

Výrobní linka je balíček softwarových nástrojů, vytvořený jako výstup projektu KODI (dokumentováno ve výstupech cíle 5 projektu), který slouží k vytváření slovníků a jejich následné publikaci. Jeho součástí jsou

- nástroje na tvorbu slovníků TermIT a OntoGrapher,
- prohlížecí nástroj ShowIT,
- nástroj pro tvorbu datových specifikací Dataspecer.

Ukázku těchto nástrojů naleznete ve školení Modelování popisu dat ve veřejné správě.

5.7 Uvedení příkladu

Po výběru oblasti, zdrojů a nástroje můžeme začít se samotným popisem dat.

> Cílem příkladu je ukázat popis dat v praxi – negarantujeme soulad výsledného slovníku s aktuálními zákony a jinými zdroji z této oblasti.

Zdroj jsme pro náš příklad vybrali jeden: Zákon o provozu na pozemních komunikacích a o změnách některých zákonů (č. 361/2000 Sb.). Z tohoto zákona budeme postupně vytahovat subjekty a objekty práva, poté vlastnosti a na konec vztahy.

5.8 Výběr pojmů

Podíváme-li se do zákona o silničním provozu, nalezneme §2, Vymezení základních pojmů. Podobný paragraf se vyskytuje v několika zákonech a představuje dobrý začátek pro vybrání těch nejdůležitějších pojmů. Například máme v tomto paragrafu písmeno d):

> *řidič je účastník provozu na pozemních komunikacích, který řídí motorové nebo nemotorové vozidlo anebo tramvaj; řidičem je i jezdec na zvířeti,*

z čehož můžeme založit první pojem. Váš výběr si můžete zkontrolovat několika hrubými pravidly:

- **Sbíráte v rámci vaší oblasti o daném pojmu data?** Pokud se pojem objevuje v informačním systému, formuláři, výpisu, či jiných záznamech, je odpověď jednoznačně ano.
- **Je pojem součástí nějaké hierarchie pojmů?** Jinými slovy, pomohlo by klasifikaci pojmů do jednotlivých kategorií, kdybychom pojem do slovníku zahrnuli? Více informací o hierarchickém řazení pojmů naleznete v podkapitole 5.11 Nadřazené pojmy.
- **Je pojem ve vztahu s nějakým jiným pojmem?** Tento požadavek slouží především jako kontrola správného rozsahu slovníku. V kapitole 4 Základy popisu dat uvádíme, že pojmy vstupují s ostatními pojmy do věcných vztahů (žadatel *vlastní* silniční vozidlo, silniční vozidlo *je pojištěno* pojištěním). I když není nutné tyto vztahy formalizovat do minimálního slovníku, měli byste být schopni nějaký (alespoň základní) vztah určit (resp. říct, *proč* daný pojem patří do dané oblasti).

V této metodice se zaměřujeme na tvorbu slovníků určených pro evidenci dat v ISVS z určité oblasti (*datové slovníky* dle Strategie). Existují však i další použití slovníků, např. jako výkladový slovník (neboli tezaurus), kde jsou kritéria pro zahrnutí pojmu do slovníku mnohem volnější. Slovníky popsané touto metodikou jsou vhodné pro oba případy.

5.9 Subjekty a objekty práva

Ve výkonu veřejné správy se běžně pracuje s daty nějakého subjektu práva a s ním souvisejícími objekty práva. Subjekty a objekty práva jako typy pojmů jsou hlavním předmětem popisu – další typy (vlastnosti a vztahy) je poté upřesňují a dodávají jim bližší kontext.

Subjekt práva je osoba, která se účastní právních vztahů. Subjekt práva můžeme také definovat jako nositele práv a povinností. Tyto osoby působí na ostatní subjekty nebo objekty práva (viz níže) – činí, vykonávají svoji vůli. Příkladem jsou:

- lidé – žadatelé, studenti, děti, řidiči, cizinci,
- organizace a soukromé subjekty – firmy, orgány veřejné moci, odbory,
- role – provozovatelé, správci, dodavatelé.

Objekt práva je příčinou vstupu subjektu do právního vztahu. Každá „věc“, která je předmětem nějakého právního vztahu, je objektem práva. Působí na ni subjekty práva – objekty z hlediska veřejné správy nemají svoji vůli a nemohou ji tedy vykonávat. Příkladem jsou:

- aktiva a hmotné předměty – vozidla, nemovitosti, zvířata, informační systémy, duševní vlastnictví, energie, voda, suroviny, průkazy,
- právní koncepty – agendy, osvědčení, práva, povinnosti, přestupky,
- další abstraktní koncepty – služby, audity, pracoviště, databáze, informace.

Ve výše uvedeném příkladu řidiče z §2, písmena d) zákona o silničním provozu se jedná o subjekt práva (řidič je fyzická *osoba*, která se *účastní* provozu).

5.10 Charakteristiky pojmu

Každý pojem, jak je naznačeno v kapitole 4 Základy popisu dat, musí být zasazen do správného kontextu a řádně vysvětlen – jinak jde o pouhé pojmenování bez jasného významu. Pro toto upřesnění slouží vyplnění následujících charakteristik k danému pojmu:

- **Název:** Jedná se o jednoznačný název pojmu, je spisovný, bez použití zkratek a v prvním pádu jednotného čísla: „*Řidič*“.
- **Zdroj:** Dokument a jeho část, v níž je pojem definován. Odstavec zákona, strana metodiky apod. Uvedení zdroje je důležité nejen pro dohledání významu, ale i pro zdůvodnění některých skutečností (např.: Z jakého důvodu data popsaná pojmem evidujete?) Pokud možno se uvádí jako odkaz URL na konkrétní dokument, nejlépe přímo na sekci dokumentu/stránky: *§ 2 písm. d) zákona č. 361/2000 Sb. o silničním provozu* (odkaz na e-Sbírku).
- **Definice:** Text přímo vyňatý ze zdroje, který stanovuje přesný význam pojmu – „*účastník provozu na pozemních komunikacích, který řídí motorové nebo nemotorové vozidlo anebo tramvaj; řidičem je i jezdec na zvířeti*“.
- **Popis:** „Zjednodušený“ (relativně k definici) výklad nebo doplňující poznámky, který je vytvořený přímo pro pojem (tj. píše ho tvůrce pojmu a nemusí tedy pocházet z nějakého zdroje): *Řidič je tedy i cyklista, nejen řidič motorových vozidel – neváže se na něj nutně řidičské oprávnění.*

Výsledný pojem „Řidič“ by tedy vypadal následovně:

Řidič
zdroj: § 2 písm. d) zákona č. 361/2000 Sb. o silničním provozu
definice: účastník provozu na pozemních komunikacích, který řídí motorové nebo nemotorové vozidlo anebo tramvaj; řidičem je i jezdec na zvířeti
popis: Řidič je tedy i cyklista, nejen řidič motorových vozidel – neváže se na něj nutně řidičské oprávnění.
typ: Subjekt práva

Vždy rozvádějte název pojmu a nedávejte do názvu zkratky. Pište „Fyzická osoba“, ne FO, „Identifikační číslo osoby“, ne IČO. Pamatujte, že vaše pojmy mají být dohledatelné a použitelné napříč oblastmi, agendami, informačními systémy atd. I v na první pohled jednoznačných pojmech může totiž vzniknout nejasnost a nepořádek, obzvlášť když se stejným/podobným názvem označuje větší množství pojmů z různých oblastí.

Dále si vybereme ve stejném paragrafu písmeno f):

vozidlo je motorové vozidlo, nemotorové vozidlo nebo tramvaj,

Protože vozidlo samo o sobě z právního hlediska nevykonává svou vůli (nepůsobí, nýbrž je na něj působeno), jde o objekt práva. Pojem do slovníku zavedeme stejně, jako „Řidič“:

Vozidlo
zdroj: § 2 písm. f) zákona č. 361/2000 Sb.
definice: motorové vozidlo, nemotorové vozidlo nebo tramvaj
typ: Objekt práva

Příklad pojmů, které nebudeme uvádět do našeho příkladového slovníku, jsou konkrétní práva a povinnosti, situace, nebo příliš obecné pojmy (snížená viditelnost, pokyn zastavit, boční sezení na motocyklu, krajnice apod.). Jejich uvedení není chybou, ale v našem příkladě to není něco, o čem bychom měli nebo chtěli mít data (např. nevedeme evidenci všech jednotlivých zastavení jednotlivých vozidel).

Pro účely příkladu popisu zavedeme ještě pojem „Řidičský průkaz“ z paragrafu 103:

Řidičský průkaz
zdroj: § 103 odst. 1 zákona č. 361/2000 Sb.
definice: Řidičský průkaz je veřejná listina, která osvědčuje řidičské oprávnění držitele a jeho rozsah a kterou držitel prokazuje své jméno, příjmení a podobu, jakož i další údaje v ní zapsané podle tohoto zákona.
typ: Objekt práva

5.11 Nadřazené pojmy

V zákoně si můžete všimnout, že se některé pojmy používají v definici dalších pojmů, například:

*řidič je **účastník provozu na pozemních komunikacích**, […]*
*Řidičský průkaz je **veřejná listina**, […]*

Je vidět, že se pojmy řadí do určité hierarchie. Podle zákona existuje několik druhů účastníků provozu na pozemních komunikacích – jedním příkladem je řidič (každý řidič je účastník provozu, ale každý účastník provozu není řidič). Pro takovéto situace existuje pojmová charakteristika *nadřazenosti*. Pojem A je nadřazený pojmu B, pokud každý příklad spadající pod pojem A zároveň spadá pod pojem B, ale pojem B dále upřesňuje pojem A nebo jej rozděluje na jednotlivé podtypy.

Pro pojem „Řidič“ je nadřazený pojem „Účastník provozu na pozemních komunikacích“ z paragrafu 2, písmena a): Každý řidič je zároveň účastníkem provozu, ale opačně to neplatí – každý účastník provozu není vždy řidič.

V definici pojmu „Vozidlo“ z předchozí podkapitoly si však můžete všimnout, že se tu hovoří o možných „variantách“: motorové vozidlo, nemotorové vozidlo, nebo tramvaj. Oproti řidiči a účastníku provozu se tu uvádí obrácená vazba, tj. že tyto tři pojmy jsou podřazené pojmu „Vozidlo“ (každé motorové vozidlo, nemotorové vozidlo a tramvaj je vozidlo).

5.12 Nejednoznačně definované pojmy

Občas se zdroje potýkají s pojmy, které nejsou přímo definovány. V zákoně o silničním provozu je to například řidič evidovaný v registru řidičů, zmíněný v §119 odst. 2:

Registr řidičů obsahuje a) osobní údaje o řidiči motorového vozidla uvedené v řidičském průkazu a v mezinárodním řidičském průkazu, včetně digitalizované fotografie a digitalizovaného podpisu řidiče,

Kromě registru řidičů se v tomto odstavci hovoří o řidičích motorového vozidla – přesněji těch, kteří mají (mezinárodní) řidičský průkaz a jsou v evidenci registru řidičů. To je však dostatečně identifikovatelná skupina, pro kterou již lze založit pojem. Charakteristika pojmu *popis* v této situaci hraje důležitou roli – v každém případě by mělo být jasné, co pod pojem „Řidič evidovaný v registru řidičů“ spadá:

Řidič evidovaný v registru řidičů
zdroj: § 119 odst. 2 písm. a) zákona č. 361/2000 Sb.
popis: Registr řidičů obsahuje osobní údaje o řidičích motorových vozidel uvedené v řidičských průkazech.
nadřazený pojem: Řidič
typ: Subjekt práva

Pokud existuje stejné označení (název „Řidič“) pro různé definice („Řidič“ jako řidič motorových vozidel a „Řidič“ jako řidič podle definice zákona), tak se jedná o více pojmů! Více pojmů se stejným názvem nemůže být v jednom slovníku – podobně jako nedává smysl ve stejné oblasti mít stejný název pro více typů „věcí“.

Podobně se například v § 105 odst. 1 písmeno i) mluví o úřadě vydávajícím řidičské průkazy:

i) název a sídlo úřadu, který řidičský průkaz vydal,

Pokud zákon budeme pročítat dále, uvidíme, že se jedná o obecní úřad obce s rozšířenou působností. Ten v zákoně není přesně definován, ale v paragrafech je dostatek informací pro vytvoření pojmu:

Obecní úřad obce s rozšířenou působností
zdroj: § 109 odst. 3 zákona č. 361/2000 Sb.
popis: Úřad vydávající řidičský průkaz je vždy obecním úřadem obce s rozšířenou působností.
typ: Subjekt práva

5.13 Vlastnosti

Vlastnosti (neboli *údaje*) jsou, stejně jako jednotlivá políčka ve formuláři, zachycením typů informací o daném subjektu nebo objektu práva. Vlastnosti určují, co všechno o daném subjektu nebo objektu práva vedeme (resp. co za políčka máme) a jak (jak se má políčko vyplňovat). I u vlastností chceme název a ideálně i zdroj, definici a popis. U vlastností je popis obzvlášť důležitý, protože také potřebujeme uvést, jak se „políčko“ má vyplňovat.

Každá vlastnost se musí vázat na jeden konkrétní subjekt nebo objekt práva. To znamená, že pokud máte vlastnosti, které evidujete/potřebujete evidovat a nemáte pro ně subjekt nebo objekt práva, měli byste založit nový a k němu vlastnost přiřadit.

Dávejte si pozor, abyste vlastností nepopisovali samostatný subjekt/objekt práva. Například „Provozovatel vozidla“ není vlastností „Vozidla“ – „Provozovatel“ je samostatný subjekt práva.

Vlastností by se měla popisovat základní jednotka dat (individuální datumy, názvy, příznaky apod.). Důležitá forma sebekontroly evidence vlastností je představit si konkrétní hodnotu takové vlastnosti (kdyby byla vlastnost vyjádřená jako políčko ve formuláři, jak by vypadal vyplněný vzor?).

Rodné číslo je vlastnost pro subjekt práva „Fyzická osoba“, což je (jedno) desetimístné číslo. 7808243949 je konkrétní rodné číslo pro konkrétní subjekt práva Jan Novák.

Pokud jste uvedli k subjektu nebo objektu práva nadřazený pojem, tak se vlastnosti z tohoto nadřazeného pojmu automaticky přebírají z nadřazeného pojmu. Je to jeden z primárních nástrojů pro zajištění, abyste se ve slovnících opakovali co nejméně.

Z rozpracovaného zákona o silničním provozu můžeme vybrat dvě vlastnosti pro „Řidičský průkaz“:

Jméno držitele řidičského průkazu
zdroj: § 105 odst. 1 písm. a) zákona č. 361/2000 Sb.
popis: Řidičský průkaz obsahuje mj. jméno držitele.
typ: Vlastnost pojmu Řidičský průkaz

Příjmení držitele řidičského průkazu
zdroj: § 105 odst. 1 písm. a) zákona č. 361/2000 Sb.
popis: Řidičský průkaz obsahuje mj. příjmení držitele.
typ: Vlastnost pojmu Řidičský průkaz

Vlastnost nesmí kombinovat více jednotek informací najednou, tj. nelze vytvořit vlastnost „Jméno a příjmení žadatele o řidičské oprávnění“.

V názvu vlastnosti je i pojem, ke kterému se vlastnost vztahuje. Zajišťujeme tak splnění požadavku na neopakující se názvy pojmů ve slovníku a jednoznačněji určujeme, k čemu vlastnost patří.

Vlastnostmi nejsou provozní data informačních systémů – provozní identifikátory, databázové klíče apod.

Paragraf 105 odst. 1 písmeno i) uvádí, že mezi informacemi na řidičském průkazu je také:

i) název a sídlo úřadu, který řidičský průkaz vydal,

z čehož nám zdánlivě vznikají dvě vlastnosti. První je název úřadu. Protože z předchozích pojmů víme, že se zde mluví přesně o obecním úřadu obce s rozšířenou působností, můžeme přiřadit název přímo k danému pojmu:

Název obecního úřadu obce s rozšířenou působností
zdroj: § 105 odst. 1 písm. i) zákona č. 361/2000 Sb.
popis: Řidičský průkaz obsahuje název úřadu, který řidičský průkaz vydal (což je vždy obecní úřad obce s rozšířenou působností).
typ: Vlastnost pojmu Obecní úřad obce s rozšířenou působností

Dále, adresa sídla se skládá z více prvků – pokud se podíváme na formulář žádosti o řidičský průkaz, adresa má hned několik dílčích vlastností: ulice, číslo orientační, číslo popisné, poštovní směrovací číslo atd. Jak bylo zmíněno výše, **vlastnosti nemohou mít své dílčí vlastnosti** – měly by zachycovat co nejmenší možné „políčko“ formuláře. Proto „Adresa“ není vlastnost, ale objekt práva, pod který spadají výše zmíněné „části“ adresy jako vlastnosti.

Díky evidenci adresy jako objektu práva navíc dovolíme, aby ostatní pojmy, které nějakým způsobem využívají adresu, se na ní navázaly také. Díky tomu se nemusí duplicitně uvádět všechny položky adresy u každého subjektu/objektu práva.

Navázání této adresy na jiné subjekty nebo objekty práva zprostředkovávají vztahy.

5.14 Vztahy

Vztah věcně propojuje jeden subjekt nebo objekt práva s jiným subjektem nebo objektem práva.

Pojem „řídí vozidlo“ je vztah mezi subjektem práva „Řidič“ a objektem práva „Vozidlo“. Řízení vozidla Škoda 120L Janem Novákem je konkrétní příklad vztahu mezi Janem Novákem (příklad subjektu práva „Řidič“) a Škoda 120L (příklad objektu práva „Vozidlo“).

Vyhýbejte se vztahům typu „má vlastnost“ – vlastnosti se k subjektům a objektům práva v podporovaných nástrojích připisují jinými způsoby.

I u vztahů chceme název a ideálně i zdroj, definici a popis. Tyto charakteristiky vztahů jsou nezbytné, neboť dva pojmy mohou být souběžně ve vztazích různých typů (je manželem, partnerem, rodičem apod.).

Vhodná pomůcka pro tvoření vztahů je, že si představíte větu ve formátu [pojem] [vztah] [pojem], například „Řidič“ *řídí vozidlo* „Vozidlo“. V názvu vztahu je i pojem, na který se vztah váže. Zajišťujeme tak, že se názvy pojmů ve slovníku neopakují a jednoznačně určujeme, k čemu vztah patří. Doporučujeme názvy psát v rodu činném, nebo přinejmenším rod činný a trpný v rámci slovníku nemíchat.

V rámci našeho příkladu jsme evidovali několik subjektů a objektů práva, které už podle vyňatých definic jsou v nějakém věcném vztahu:

řídí vozidlo
zdroj: § 2 písm. d) zákona č. 361/2000 Sb. o silničním provozu
definice: řidič je účastník provozu na pozemních komunikacích, který řídí motorové nebo nemotorové vozidlo anebo tramvaj; řidičem je i jezdec na zvířeti
typ: Vztah mezi Řidič a Vozidlo

drží řidičský průkaz
zdroj: § 119 odst. 2 písm. a) zákona č. 361/2000 Sb.
definice: Registr řidičů obsahuje a) osobní údaje o řidiči motorového vozidla uvedené v řidičském průkazu a v mezinárodním řidičském průkazu, včetně digitalizované fotografie a digitalizovaného podpisu řidiče,
popis: Z definice vyplývá, že řidiči, kteří jsou evidováni v registru řidičů, jsou držiteli řidičského průkazu.
typ: Vztah mezi Řidič evidovaný v registru řidičů a Řidičský průkaz

Daný vztah nemusí platit pro všechna konkrétní data. Kdybychom udělali stejný vztah *drží řidičský průkaz* mezi obecným „Řidičem“ jako účastníkem provozu a „Řidičským průkazem“, tak tím říkáme, že každý řidič může (ale nemusí) mít řidičský průkaz. Podle toho vytvářejte názvy vztahů – není potřeba uvádět *„může“ vlastnit řidičský průkaz* – *může* se implikuje. Jelikož je ale tento vztah mezi „Řidič evidovaný v registru řidičů“ a „Řidičský průkaz“ jasnější, uvádíme ho v podobě výše.

Pomocí vztahů si můžete ověřit naplnění požadovaného rozsahu popisu dat. Pokud považujete slovník za hotový (resp. jste přesvědčeni, že zachycuje celou zvolenou oblast/část oblasti) a máte v něm pojem, který není s žádným jiným provázaný (vztahem ani nadřazeností), je tento pojem s velkou pravděpodobností přebytečný, nebo jsou vztahy ve slovníku chybně vyjádřeny. S tím bychom si mohli uvědomit, že jsme zapomněli zavést vztah mezi „Obecní úřad obce s rozšířenou působností“ a „Řidičský průkaz“:

vydává řidičský průkaz
zdroj: § 109 odst. 3 zákona č. 361/2000 Sb.
definice: Řidičský průkaz vydá příslušný obecní úřad obce s rozšířenou působností na žádost držitele řidičského oprávnění.
popis: Z definice vyplývá, že řidiči, kteří jsou evidovaní v registru řidičů, mohou být držiteli řidičského průkazu.
typ: Vztah mezi Obecní úřad obce s rozšířenou působností a Řidičský průkaz

Při pojmenovávání vztahů dbejte na to, abyste nechali co nejmenší prostor pro různé interpretace významu vztahu.

Nevytvářejte nespecificky pojmenované vztahy („je součástí“, „je“, „se skládá z“, „tvoří“, „řeší“ apod.) – je to relativně nejasné vyjádření významu („je součástí“ proč? V jakém slova smyslu?). Místo toho používejte specifické názvy – například „provozuje“, „podává zprávu“, „vydává“, „schvaluje“ apod.

V podkapitole 5.13 Vlastnosti jsme diskutovali potenciální vztah mezi adresou a „Úřadem vydávajícím řidičské průkazy“. Pojem „Adresa“ je již definovaný v § 29 odst. 1 písm. h) zákona č. 111/2009 Sb., o základních registrech, který je vzhledem k jeho navázání na základní registry (jejichž data jsou velkou skupinou informačních systémů využívány) nanejvýše relevantní. Pokud je tento zákon popsaný jiným slovníkem, je nejjednodušší použít již definovaný pojem.

5.15 Využití pojmů z jiných slovníků

Používání pojmů z jiných slovníků je jedna z nejdůležitějších možností, které slovníky přinášejí. Valná většina oblastí veřejné správy je propojena s jinou oblastí. Dále celá veřejná správa pracuje s několika základními pojmy (například „Fyzická osoba“, „Adresa“, „Dokument“ apod.). Proto se můžete při tvorbě slovníků odkazovat na pojmy mimo váš slovník – ať už se nacházejí na jednom konci vztahu nebo jsou součástí nadřazenosti.

Každý **publikovaný** pojem (více viz 5.18 Publikace slovníku) má svůj vlastní jednoznačný identifikátor – IRI (podobné webové adrese s diakritikou). Díky IRI se může pojem využívat v ostatních slovnících, aniž by se musel kopírovat do cílového slovníku a také nedojde ke ztrátě kontextu a informací o daném pojmu. Kroky k získání a použití IRI jsou popsány na https://data.gov.cz/popis-dat.

Příkladů z oblasti řidičů je několik. Zmínili jsme „Adresu“, která je definována ve slovníku zákona o základních registrech, kde je i napojena na adresní místo RÚIAN.

Výsledný vztah mezi „Obecním úřadem obce s rozšířenou působností“ a „Adresou“ by tedy vypadal následovně:

sídlí na adrese
zdroj: § 105 odst. 1 písm. i) zákona č. 361/2000 Sb.
popis: Sídlo úřadu vydávajícího průkaz (obecním úřadem ORP) se nachází na nějaké adrese. Adresa je obecně definována v zákoně o základních registrech.
typ: Vztah mezi Obecní úřad obce s rozšířenou působností a Adresa ze slovníku zákona č. 111/2009 Sb., o základních registrech

DIA spravuje a publikuje slovník obecných pojmů veřejného sektoru, ve kterém jsou obecné pojmy, které jsou často používány napříč agendami veřejné správy (například Osoba, Dokument, Subjekt práva). Tento slovník se vyskytuje ve všech šablonách pro nástroje pro tvorbu popisu dat a je určen právě k použití v rámci jiných slovníků. Pojmy ze slovníku obecných pojmů veřejného sektoru slouží jako pojmy *nadřazené* vámi vytvořeným pojmům.

5.16 Kompletní příklad slovníku

Zde jsou uvedeny pojmy celé iterace slovníku, který byl vytvářen v předchozích sekcích. U všech zdrojů je uveden i odkaz ve formátu ELI na příslušnou část zákona v e-Sbírce. Příklady vytvoření tohoto slovníku v podporovaných nástrojích jsou uvedeny v návodech pro popis dat v těchto nástrojích.

Řidič
zdroj: § 2 písm. d) zákona č. 361/2000 Sb. o silničním provozu
definice: účastník provozu na pozemních komunikacích, který řídí motorové nebo nemotorové vozidlo anebo tramvaj; řidičem je i jezdec na zvířeti

popis: Řidič je tedy i cyklista, nejen řidič motorových vozidel – neváže se na něj nutně řidičské oprávnění.
nadřazený pojem: Účastník provozu na pozemních komunikacích
typ: Subjekt práva

Vozidlo
zdroj: § 2 písm. f) zákona č. 361/2000 Sb.
definice: motorové vozidlo, nemotorové vozidlo nebo tramvaj
typ: Objekt práva

Účastník provozu na pozemních komunikacích
zdroj: § 2 písm. a) zákona č. 361/2000 Sb. o silničním provozu
definice: každý, kdo se přímým způsobem účastní provozu na pozemních komunikacích
typ: Subjekt práva

Řidič evidovaný v registru řidičů
zdroj: § 119 odst. 2 písm. a) zákona č. 361/2000 Sb.
definice: Registr řidičů obsahuje a) osobní údaje o řidiči motorového vozidla uvedené v řidičském průkazu a v mezinárodním řidičském průkazu, včetně digitalizované fotografie a digitalizovaného podpisu řidiče,
nadřazený pojem: Řidič
typ: Subjekt práva

Řidičský průkaz
zdroj: § 103 odst. 1 zákona č. 361/2000 Sb.
definice: Řidičský průkaz je veřejná listina, která osvědčuje řidičské oprávnění držitele a jeho rozsah a kterou držitel prokazuje své jméno, příjmení a podobu, jakož i další údaje v ní zapsané podle tohoto zákona.
typ: Objekt práva

Obecní úřad obce s rozšířenou působností
zdroj: § 109 odst. 3 zákona č. 361/2000 Sb.
popis: Úřad vydávající řidičský průkaz je vždy obecním úřadem obce s rozšířenou působností.
typ: Subjekt práva

Jméno držitele řidičského průkazu
zdroj: § 105 odst. 1 písm. a) zákona č. 361/2000 Sb.
popis: Řidičský průkaz obsahuje mj. jméno a příjmení držitele.
typ: Vlastnost pojmu Řidičský průkaz

Příjmení držitele řidičského průkazu
zdroj: § 105 odst. 1 písm. a) zákona č. 361/2000 Sb.
popis: Řidičský průkaz obsahuje mj. jméno a příjmení držitele.
typ: Vlastnost pojmu Řidičský průkaz

Název obecního úřadu obce s rozšířenou působností
zdroj: § 105 odst. 1 písm. i) zákona č. 361/2000 Sb.
popis: Řidičský průkaz obsahuje název úřadu, který řidičský průkaz vydal (což je vždy obecní úřad obce s rozšířenou působností).
typ: Vlastnost pojmu Obecní úřad obce s rozšířenou působností

drží řidičský průkaz
zdroj: § 119 odst. 2 písm. a) zákona č. 361/2000 Sb.
definice: Registr řidičů obsahuje a) osobní údaje o řidiči motorového vozidla uvedené v řidičském průkazu a v mezinárodním řidičském průkazu, včetně digitalizované fotografie a digitalizovaného podpisu řidiče,
popis: Z definice vyplývá, že řidiči, kteří jsou evidováni v registru řidičů, mohou být držiteli řidičského průkazu.
typ: Vztah mezi Řidič evidovaný v registru řidičů a Řidičský průkaz

řídí vozidlo
zdroj: § 2 písm. d) zákona č. 361/2000 Sb. o silničním provozu
definice: řidič je účastník provozu na pozemních komunikacích, který řídí motorové nebo nemotorové vozidlo anebo tramvaj; řidičem je i jezdec na zvířeti
typ: Vztah mezi Řidič a Vozidlo

vydává řidičský průkaz
zdroj: § 109 odst. 3 zákona č. 361/2000 Sb.
definice: Řidičský průkaz vydá příslušný obecní úřad obce s rozšířenou působností na žádost držitele řidičského oprávnění.
popis: Z definice vyplývá, že řidiči, kteří jsou evidováni v registru řidičů, mohou být držiteli řidičského průkazu.
typ: Vztah mezi Řidič evidovaný v registru řidičů a Řidičský průkaz

sídlí na adrese
zdroj: § 105 odst. 1 písm. i) zákona č. 361/2000 Sb.
popis: Sídlo úřadu vydávajícího průkaz (obecním úřadem ORP) se nachází na nějaké adrese. Adresa je obecně definována v zákoně o základních registrech.
typ: Vztah mezi Obecní úřad obce s rozšířenou působností a Adresa ze slovníku zákona č. 111/2009 Sb., o základních registrech

5.17 Pokročilý popis dat

Při popisování dat můžete narazit na specifické problémy nebo potřeby vyjádřit specifické skutečnosti. Tyto scénáře a jejich řešení jsou blíže popsány na https://data.gov.cz/popis-dat.

5.18 Publikace slovníku

Před jakýmkoliv využitím popisu dat je potřeba jej publikovat tak, aby byl katalogizován v Národním katalogu dat. Slovníky jsou otevřenými daty, a bez tohoto kroku nemohou podle být zákona č. 106/1999 Sb. považována za otevřená data. Není to však jen o povinnostech, publikace vám poskytne několik výhod:

- Slovníky budou snadno dostupné přes internet vám i ostatním. Publikací získává každý pojem a slovník jednoznačný identifikátor (IRI), pomocí kterého se na něj dá odkazovat.
- Díky katalogizaci bude hledání slovníků a pojmů mnohem jednodušší.
- Během přípravy slovníku se potvrzuje kompatibilita s otevřenou formální normou pro slovníky – aby bylo zajištěno, že s vaším slovníkem mohou pracovat všichni.

V prvé řadě je nutné určit název slovníku, který jste vytvořili. Název se odvíjí od oblasti, kterou slovníkem popisujete (nebo zdroje, pokud pro slovník používáte právě jeden).

> Dobré příklady názvu slovníku jsou: Slovník zákona č. 361/2000 Sb., Slovník agendy A998, Slovník veřejných zakázek, Slovník badatelského řádu apod. Slovník z našeho příkladu jsme nazvali Slovník *zákona č. 361/2000 Sb.* a přidáme popis: Příkladový slovník, který sbírá pojmy ze zákona o silničním provozu.

Poté vezměte svůj slovník (resp. výstup z nástrojů), a pokud již není ve formátu OFN pro slovníky, převeďte jej do tohoto formátu. Návody pro získání výstupů z jednotlivých nástrojů jsou obsaženy v návodech pro jednotlivé nástroje. Tyto výstupy pak nahrajte na stránkách https://slovník.gov.cz/popisujeme. Zde se nachází webová služba určená pro převod do formátu dle OFN a základní kontrolu. Následně můžete soubor se slovníkem dle OFN nahrát na veřejně dostupnou adresu a založit katalogizační záznam do svého lokálního katalogu dat.

> Slovník je datová sada, jako jakákoliv jiná v Národním katalogu dat. Aby však byla tato datová sada rozpoznána jako slovník, je nutné pro ni uvést OFN slovníky jako specifikaci této sady.

Pokud se slovník poté katalogizuje v Národním katalogu dat jako slovník, máte publikaci a tím i celou iteraci hotovou. Pokud postupujete podle opatření Strategie, iterace v rámci vašich možností opakujte, dokud nemáte popsaná všechna data vašich informačních systémů veřejné správy.