Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra geomatiky

Bakalářská práce

Geo-ontologie veřejné správy a samosprávy

Plzeň, 2021 Viktorie Sloupová

Prohlášení

Předkládám tímto k posouzení a obhajobě bakalářskou práci zpracovanou na závěr studia na Fakultě aplikovaných věd Západočeské univerzity v Plzni.

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně a výhradně s použitím odborné literatury a pramenů, jejichž úplný seznam je její součástí.

V Plzni dne 21. května 2021

.....

podpis

Poděkování

Mé poděkování patří doc. Ing. Mgr. Otakaru Čerbovi, Ph.D. za odborné vedení, trpělivost, ochotu a cenné rady, které mi v průběhu zpracování bakalářské práce věnoval. Poděkování patří také Bc. Petru Štruncovi za poskytnutí hardwarového vybavení a užitečných rad při řešení technických potíží. V neposlední řadě také děkuji své rodině za podporu, které se mi dostávalo během celého studia.

Abstrakt

Práce se zabývá vytvořením obecné geo-ontologie veřejné správy a samosprávy v České republice následně naplněné konkrétními daty pro Plzeňský kraj. Cílem navrhované geo-ontologie je vytvořit takový model skutečnosti, který plně vystihuje existující vazby mezi administrativním členěním státu a územní působností státních institucí a orgánů veřejné správy. Použitelnost výsledného obecného modelu je otestována doplněním jádrové struktury ontologie o reálná data z území Plzeňského kraje.

Klíčová slova: geo-ontologie, editor Protégé 5.5.0, administrativní členění státu, veřejná správa, státní správa, samospráva, územní působnost

Abstract

This thesis deals with the creation of a general geo-ontology of public administration and self-government in the Czech Republic, subsequently filled with specific data for the Pilsen region. The aim of the proposed geo-ontology is to create a model of reality that fully captures the existing links between the administrative division of the state and the territorial scope of state institutions and public administration bodies. The applicability of the resulting general model is tested by supplementing the core structure of the ontology with real data from the territory of the Pilsen Region.

Key words: geo-ontology, editor Protégé 5.5.0, administrative division of the state, public administration, state administration, self-government, territorial scope

Obsah

Použité zkratky **1**

Použité značení **2**

Úvod **4**

1 Rešerše literatury **6**
1.1 Použitá metodika . . . 6
1.2 Definice ontologie . . . 6
1.2.1 Ontologie jako filozofická disciplína . . . 6
1.2.2 Ontologie v informatice . . . 8
1.2.3 Dělení ontologií . . . 10
1.3 Definice geo-ontologie . . . 11
1.3.1 Přístupy k vytváření geografických konceptů . . . 12
1.3.2 Formální popis geo-ontologie . . . 14
1.4 Vybrané příklady geo-ontologií . . . 14
1.4.1 GeoNames Ontology . . . 14
1.4.2 FAO Geopolitical Ontology . . . 15
1.4.3 Projekt TOWNTOLOGY . . . 15
1.4.4 Europian Road Ontology EUROTL . . . 16
1.5 Deskripční logika a mereotopologie . . . 16
1.5.1 Deskripční logika . . . 17
1.5.2 Mereotopologie . . . 18

2 Návrh ontologie **19**
2.1 Struktura ontologie . . . 19
2.2 Jazyky pro návrh ontologie . . . 23
2.2.1 Resource Description Format RDF . . . 23
2.2.2 Web Ontology Language OWL . . . 25
2.3 Editor Protégé . . . 27
2.4 Metodika návrhu ontologie . . . 28

3 Geo-ontologie veřejné správy a samosprávy **31**
3.1 Specifikace geo-ontologie veřejné správy a samosprávy . . . 31
3.1.1 Účel a vymezení rozsahu . . . 31

3.1.2 Použitá data a jejich zdroje . . . 32
3.2 Vytvoření jádrové struktury . . . 32
3.2.1 Identifikace jednotlivých prvků ontologie . . . 33
3.2.2 Vytvoření taxonomie . . . 38
3.2.3 Podrobná deklarace tříd a vlastností . . . 45
3.3 Dokumentace ontologie . . . 52

4 Naplnění ontologie daty pro Plzeňský kraj 55
4.1 Obecné informace o Plzeňském kraji . . . 55
4.2 Použitá data a jejich zdroje . . . 57
4.3 Plnění ontologie daty . . . 59
4.4 Problémy při plnění ontologie konkrétními daty . . . 61
4.4.1 Řešení problémů . . . 62

Závěr 64

Zdroje 64

Seznam digitálních příloh 70

A Přílohy 71
A.1 Příloha 1 . . . 71
A.1.1 Úvod do matematické logiky a predikátové logiky prvního řádu . . . 71
A.1.2 Mereotopologické axiomy a definice . . . 72
A.1.3 Deskripční logika . . . 77
A.2 Příloha 2 . . . 81
A.2.1 Podtřídy třídy Územní jednotka . . . 81
A.2.2 Podtřídy třídy Instituce . . . 81

Použité zkratky

DAML+OIL = DARPA Agent Markup Language + Ontology Interface Layer
GIS = geografický informační systém
HTML = Hypertext Markup Language
JSON = JavaScript Object Notation
JVM = Java Virtual Machine
LOD = Linked Open Data
OCML = Optional Conceptual Modeling Language
OWL = Web Ontology Language
RDF = Resource Description Framework
RDFS = Resource Description Framework Schema
RIF = Rule Interchange Format
RSO = Registr sčítacích obvodů a budov
RÚIAN = Registr územní identifikace, adres a nemovitostí
SHOE = Simple HTML Ontology Extension
SPARQL = SPARQL Protocol and RDF Query Language
SWRL = Semantic Web Rule Language
URI = Uniform Resource Identifier
W3C = World Wide Web Consortium
XML = Extensible Markup Language

Použité značení

C	množina konceptů
R	množina vztahů
rel	funkce, která specifikuje relace v množině R
HC	hierarchie konceptů
A_0	množina axiomů
$\mathbf{I}$	universum (definiční obor)
Δ, $\top$	univerzální koncept, nejobecnější třída (*top class*)
$\emptyset$, $\bot$	prázdný koncept, nejspecifičtější třída (*bottom class*)
$x, y, z, atd.$	prvek z universa (definičního oboru), entita, element
$P, D, O, atd.$	predikát, výroková formule
σ, ι	termo-tvorný operátor
φ	podmínka

Značení predikátů

P	je částí	*is a part of*
IP	je vnitřní částí	*is an interior part of*
PP	je řádnou částí	*is a proper part of*
IPP	je řádnou vnitřní částí	*is a proper interior part of*
O	překrývá	*overlaps*
PO	řádně překrývá	*proper overlaps*
U	je překryto	*underlaps*
PU	je řádně překryto	*proper underlaps*
D	je oddělený od	*is discrete from*
Pt	je bod	*is a point*
X	protíná	*crosses*
OX	protíná shora	*over-crosses*
UX	je prot'ato zdola	*under-crosses*
St	překlenuje	*straddles*
B	je hranicí	*is a boundary of*
T	je tečnou	*is a tangent of*
cl	uzavření	*closure*
bdy	maximální ohraničení	*maximal boundary of*
int	vnitřní prostor	*interior of*
C	je propojený	*is connected*
EC	je propojený zvenčí	*is externally connected*

Úvod

V současné době je velmi aktuálním tématem digitalizace státní správy a obecně usnadnění přístupu občanů ke státní správě a zjednodušení jednání s jednotlivými institucemi. Tato bakalářská práce navíc vznikala v době probíhající pandemie koronavirového onemocnění Covid-19, kdy docházelo k omezování osobního kontaktu na úplné minimum včetně uzavírání kontaktních míst institucí veřejné správy a drtivá většina jednání s úřady byla přesunuta do online prostředí. O to více se začaly projevovat a diskutovat dlouhotrvající problémy spojené s poměrně zastaralým způsobem fungování státní správy v České republice, jako je nedostatečný postup elektronizace jednotlivých úřadů, často chybějící možnost vyřizovat úřední záležitosti dálkově, složité a zdlouhavé úřední postupy v jednání, nutnost jednat ohledně jedné záležitosti s více dotčenými orgány či úřad navštěvovat opakovaně.
V posledních dvaceti letech dochází k digitalizaci nejrůznějších národních registrů a všechny úřady již mají vlastní webové stránky, kde bývají dostupné informace ohledně vykonávané agendy spolu s příslušnými dokumenty a doporučenými postupy, jak řešit různé životní situace. V rámci rozvoje eGovernmentu byl také vytvořen Portál veřejné správy, který spolu s Portálem občana představuje obrovský krok kupředu v elektronizaci komunikace mezi občanem a státem a funguje již od začátku roku 2018. Nicméně podle údajů z července 2020 z kontroly provedené Nejvyšším kontrolním úřadem, se k elektronické identifikaci v rámci Portálu občana přihlásilo pouhé půl procento obyvatel České republiky [1]. Otázkou tedy zůstává, co za nezájmem občanů o využívání nabízených digitálních služeb státu stojí a zda se vůbec provádí digitalizace správným způsobem, ale to není tématem této práce.
Obecně je česká státní správa poměrně složitá a neflexibilní soustava sestávající z velkého množství vzájemně různě provázaných institucí a je velmi obtížné se v tomto celku zorientovat. Cílem mojí bakalářské práce je proto vytvořit co nejpřehlednější a zároveň co nejobecnější, a přesto výstižný model této komplikované struktury. Výsledný model by měl být navíc v takové podobě, aby byl strojově čitelný, přenositelný mezi uživateli a různými platformami a použitelný v dalších aplikacích. Uspokojit všechny tyto požadavky umožňují ontologie, a proto jsem se zaměřila na vytvoření geo-ontologie veřejné správy a samosprávy, která by kromě modelu struktury státní správy řešila i vazby mezi územní působností jednotlivých institucí a administrativním členěním státu. Pro ověření použitelnosti navrženého modelu byla použita data týkající se územního členění a působnosti jednotlivých institucí v Plzeňském kraji.
Text bakalářské práce je rozdělen do čtyř kapitol. První kapitola obsahuje rešerši literatury a seznámení s problematikou ontologií a ontologického inženýrství. Ve druhé kapitole pak podrobněji popisuji obecný postup návrhu ontologie a následně ve třetí kapitole se zaměřuji

na specifikaci mnou vytvářené Geo-ontologie veřejné správy a samosprávy. V závěrečné kapitole se pak věnuji naplnění ontologie konkrétními daty pro Plzeňský kraj. V závěru práce jsou stručně shrnuty všechny výstupy bakalářské práce včetně potíží, se kterými jsem se během vypracování setkala. Součástí bakalářské práce jsou také dvě vložené přílohy - první se týká deskripční logiky a mereotopologie, druhá pak specifikace jednotlivých tříd ontologie. Digitální přílohy bakalářské práce zahrnují kód samotné geo-ontologie, dokumentaci, transformační pravidla a tabulky se zdrojovými daty. Všechny tyto soubory včetně textu bakalářské práce jsou uloženy na cloudovém úložišti Google Drive pod doménou *gapps.zcu.cz*.

1 Rešerše literatury

1.1 Použitá metodika

Rešerše literatury byla provedena metodou sněhové koule, která je založena na získávání dalších zdrojů informací pomocí referencí a citací v základní doporučené literatuře. Další použitou metodou pro získání informací k danému tématu bylo vyhledávání pomocí klíčových slov či jejich vhodnou kombinací. Vzhledem k faktu, že zpracovávané téma je poměrně mladé a rychle se rozvíjející, neexistuje mnoho tištěných publikací, které by danou problematiku zpracovávaly, a proto jsem vycházela převážně z odborných článků nebo sborníků vědeckých prací publikovaných online. Všechny použité prameny byly nalezeny na internetu v databázích odborných článků (Google Scholar, Research Gate, Academia) nebo na webových stránkách univerzit, soukromých subjektů a organizací zabývajících se vývojem ontologií. Použité zdroje jsou uvedené v závěru bakalářské práce.

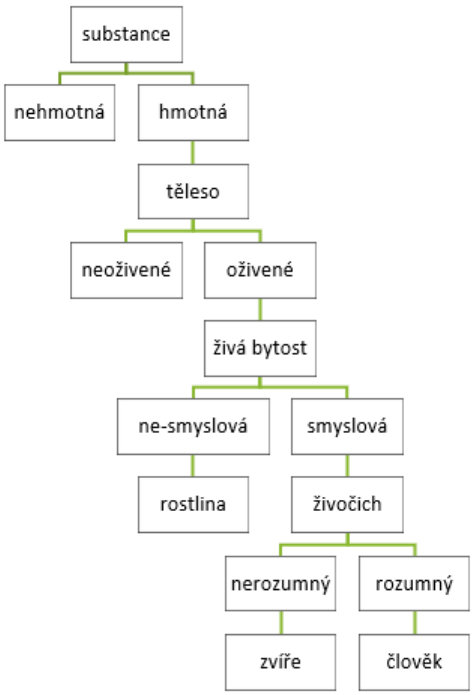
1.2 Definice ontologie

1.2.1 Ontologie jako filozofická disciplína

Slovo *ontologie* pochází z řečtiny (vzniklo složením slov *on, ontos = jsoucí* a *logos = slovo, řeč, smysl*). Poprvé jej použili v 17. století němečtí filosofové Rudolf Göckel a Jacob Lorhard zabývající se metafyzikou a od té doby je ontologie braná jako jedna ze základních teoretických filozofických disciplín. Ovšem kořeny ontologie sahají mnohem hlouběji do minulosti až do doby antiky, kdy se řečtí myslitelé snažili pochopit a popsat veškeré jsoucno a podstatu existence skutečností. Již Aristotelés se v 3. století př. n. l. zabýval ve svých dílech *Kategorie* a *Metafyzika* tím, jak popsat vše existující (*jsoucí*) a jak rozpoznat základní vlastnosti jsoucna. Aristotelova nauka zabývající se podstatou jsoucna a nejobecnějšími otázkami bytí je nazývána jako *první filozofie* [1] a je dodnes součástí metafyziky. Pokusil se rozčlenit veškerou skutečnost do 10 základních kategorií, kdy první kategorií je *substance* představující samotné bytí, a dalších devět kategorií jsou různé *akcidenty*: kvantita, kvalita, relace, kde, kdy, poloha, schopnost mít, činnost a trpnost. Tyto akcidenty mohou existovat jen v nějakém

[1] Aristotelés definoval ve svém spisu *Metafyzika* první filozofii jako "... druh vědy, jež zkoumá jsoucno jako jsoucno a to, co mu o sobě náleží. Tato věda není totožná se žádnou tak zvanou vědou zvláštní. Neboť žádná jiná věda nepojednává obecně o jsoucnu jako jsoucnu, nýbrž každá si z něho vybere určitou část a zkoumá určení, jež jí náležejí, jako například vědy matematické" [4].

vztahu k určité substanci. Například určitá kvalita vyjádřená slovem "červená" může existovat a mít smysl jen v souvislosti s určitou věcí, například s jablkem, ale nikdy ne samostatně. Jeho žáci a následovníci pak v jeho myšlenkách pokračovali dále a snažili se je dále rozvíjet například tím, že se pokoušeli najít mezi kategoriemi hierarchické vztahy, jak dokládá tzv. Porfyriův strom ze spisu *Isagoge* z 3. století n. l. znázorněný na obrázku 1.1 [2].



Obrázek 1.1: Porfyriův strom - první doložená kategorizace jsoucna. Vlastní zpracování podle [3]

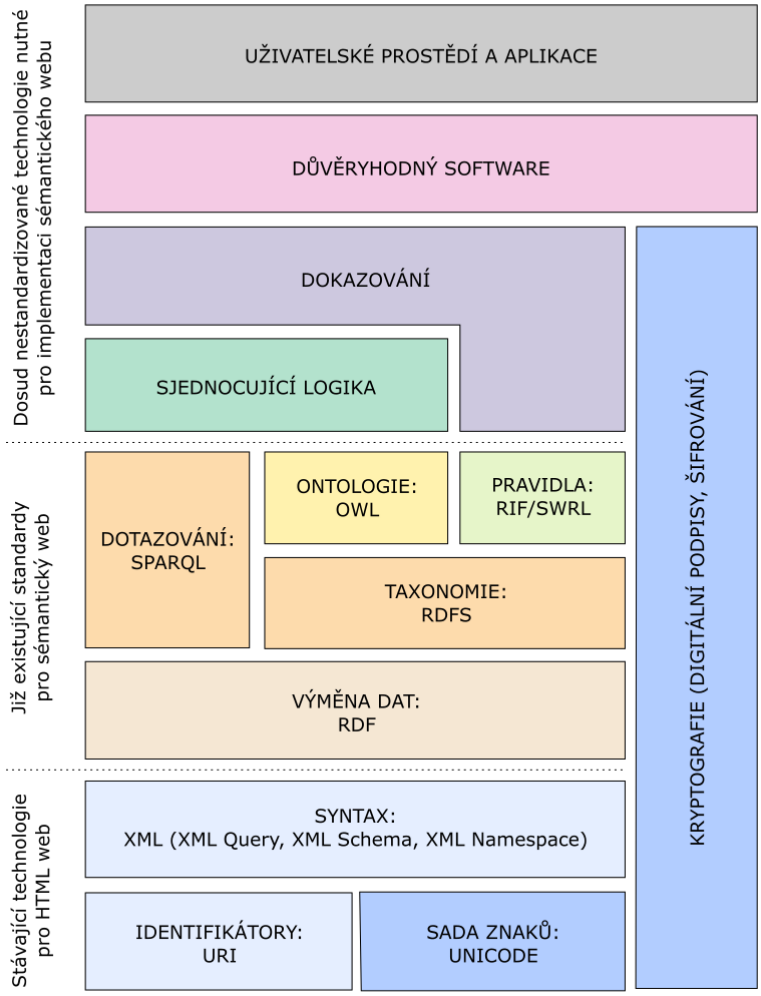
Porfyriova nauka o kategoriích a jejich hierarchii se stala základem středověké *starší logiky (logica vetus)*, kterou reprezentuje dílo myslitele Ramona Llulla *Ars Magna* z počátku 14. století přinášející propracovanější logickou klasifikaci veškerého poznání nezávislou na jazyce, náboženství a kulturním pozadí. Za skutečného otce ontologie bývá označován osvícenský filozof Christian Wolff, který byl proslulý svojí systematičností a snahou o utřídění veškerého tehdejšího vědění a poznatků včetně přesných popisů pomocí axiomů a logických formulací. Tím vznikla v první polovině 18. století formální ontologie, někdy také označovaná jako deskriptivní ontologie. Té se v moderní době věnoval například Sir Peter F. Strawson v díle *Individuals*, kde popisuje problematiku klasifikace jsoucna a sítí sdílených významů, tedy toho, jak lidé vnímají skutečnost a jak o ní přemýšlí [2].

1.2.2 Ontologie v informatice

V současné době se pojem ontologie velmi často spojuje s informatikou a je nutné zdůraznit, že s původním filozofickým pojetím ontologie nejsou zcela totožné. Společná je pro oba přístupy snaha o popis všeho, co existuje (tj. veškerého jsoucna) nezávisle na lidském usuzování o něm, a o vytvoření univerzální soustavy znalostí o objektech, jevech a zákonitostech světa. Informatické pojetí ontologie se pak uplatňuje v ontologickém inženýrství od 90. let 20. století jako možnost pro formální reprezentaci znalostí hlavně v souvislosti s iniciativou *sémantického webu* a rozvoje znalostních systémů nutných pro implementaci umělé inteligence [5]. Jako jeden z prvních definoval ontologii v kontextu informatiky Thomas Gruber jako "... explicitní specifikaci konceptualizace" [6]. Tuto definici dále rozpracovávali a upřesňovali další experti, například v [7] můžeme najít modifikaci Gruberovy definice ontologie jako "... formální specifikace sdílené konceptualizace". V [8] se autoři pokouší základní definici ontologie podrobněji vysvětlit specifikováním jednotlivých termínů v kontextu ontologií:

- **konceptualizaci** popisují jako "vytváření abstraktních modelů libovolných jevů z našeho světa pomocí relevantní koncepce těchto jevů"
- **explicitní specifikace** znamená, že "jednotlivé koncepce a jejich vymezení je přesně, výslovně definováno pomocí pojmů - nejedná se tedy pouze o myšlenkový konstrukt, ale skutečné explicitní vyjádření"
- **formální** v souvislosti s ontologiemi označuje vlastnost, že je ontologie strojově čitelná
- **sdílená** konceptualizace značí konceptualizaci "obecných znalostí uznávaných celou skupinou lidí, nikoliv pouze jenjednotlivcem"

Novější definice pak vnímají ontologie již více v kontextu umělé inteligence a znalostních systémů, jako příklad lze uvést definici z [9]: "Ontologie umožňují explicitní popis koncepcí pro reprezentaci znalostí ve znalostní bázi." Vytvoření kvalitní a komplexní znalostní báze je nutná podmínka pro implementaci umělé inteligence, neboť se jedná o obdobu lidských znalostí a vědění, na základě kterých jsme jako lidé schopni dále usuzovat, jednat a komunikovat a od umělé inteligence očekáváme podobné chování. Pro vzájemné porozumění dvou a více komunikujících prvků, ať už se jedná o lidi či počítače, je nezbytné, aby všichni účastníci komunikace uznávali tentýž popis světa, což umožňují právě ontologie. Ontologie představují jakýsi obecný slovník obsahující glosář definující jednotlivé pojmy a tezaurus, který popisuje vztahy mezi nimi.



Obrázek 1.2: Schéma vrstev sémantického webu. Vlastní zpracování podle [10]

Vzhledem k faktu, že je dnes zcela nepředstavitelné fungování světa bez Internetu a zároveň roste poptávka po automatizovaném strojovém zpracování dat, je vhodné vytvořit i takové webové prosředí, které bude narozdíl od aktuální podoby webu čitelné i pro počítače. Tím vznikla na počátku nového tisíciletí iniciativa *sémantického webu*, za níž stojí samotný tvůrce World Wide Webu Sir Timothy John Berners-Lee, jež je ředitelem konsorcia W3C dohlížejícím na další vývoj webu. Konsorcium definuje sémantický (neboli významový) web jako "... síť propojených dat takovým způsobem, aby byly snadno strojově zpracovatelné na celosvětové úrovni. Jedná se o nástavbu a vylepšení současného webu pomocí efektivní reprezentace dat za účelem zjednodušení vyhledávání, sdílení a kombinování informací." [10] Aby toho mohlo být dosaženo, konsorcium sestavilo schéma jednotlivých vrstev sémantického webu (viz obrázek 1.2) a vytvořilo nové standardy pro implementaci jednotlivých vrstev. Jednu z vrstev pak tvoří ontologie implementované v jazyce OWL, který je vhodný pro ex-

plicitní popis dat včetně jejich taxonomické struktury, definování pojmů a vztahů (relací) a spolu s RDFS tvoří jádro sémantického webu. Takto strukturovaný web pak umožňuje *softwarovým agentům* (někdy též nazývaných *boti*) automatizovaně vyhledávat data, porozumět jim, vytvářet nad nimi analýzy a dále je zpracovávat bez nutného zásahu člověka.
V posledních letech také vzrůstá význam tzv. *doménových ontologií*, které jsou zaměřené pouze na určitý obor - *doménu* a umožňují expertům snáze sdílet a vyměňovat si znalosti ve svém oboru. Díky doménovým ontologiím je jednodušší definovat a sjednotit používanou terminologii, jednoznačně vymezit vztahy mezi používanými pojmy a specifikovat jednotlivé koncepty v daném oboru. Tím lze při komunikaci předejít nejasnostem plynoucích z různého pojetí stejné problematiky či odlišného pojmenování a značení jednotlivých předmětů výzkumu. Podle [11] mezi hlavní přínosy doménových ontologií patří:

- provedení komplexní analýzy doménových znalostí z hlediska struktury informací a existujících vztahů mezi nimi
- sdílení znalostí a informací nejenom mezi lidmi, ale i tzv. *softwarovými agenty*
- možnost opakovaně používat doménové znalosti různými uživateli se specifickými požadavky
- vytvoření slovníku explicitně formulovaných definic používaných termínů
- oddělení doménových znalostí od konkrétních implementací

Cílem doménových ontologií je vytvořit sdílitelný a znovupoužitelný model konceptů z určitého oboru, který usnadní komunikaci a spolupráci mezi experty a zároveň umožní implementaci doménových znalostí v kontextu jednotlivých dílčích výzkumů a konkrétních řešených problémů [11].

1.2.3 Dělení ontologií

V různých pramenech lze najít několik přístupů k možnému členění ontologií. Například v [12] je uvedené rozdělení vývojových proudů ontologií podle jejich původu na:

1. **terminologické (lexikální) ontologie**, které se vyvinuly z knihovnických tezaurů a jsou orientované hlavně na formální definici termínů, vymezení taxonomických relací mezi nimi a stanovení obecných vztahů jako synonymie, meronymie a pod.
2. **informační ontologie**, které jsou nástavbou na relační databáze, kterým dodávají konceptuální schéma, zajišťují lepší integritu dat a umožňují díky větší abstrakci pojmové dotazování nad daty oproti běžným databázovým nástrojům

3. **znalostní ontologie** vychází z reprezentace znalostí pro umělou inteligenci pomocí logických a formálních modelů reálných objektů, které sestávají z přesně definovaných konceptů a relací

Další možné dělení uvádí [13], v němž se ontologie rozlišují podle toho, co je jejich předmětem, na:

1. **generické ontologie** usilující o zpracování všech obecných znalostí a zákonitostí, které platí napříč celým spektrem různých oborů
2. **doménové ontologie**, jejichž předmětem je určitá specifická oblast - doména
3. **úlohové ontologie** zabývající se různými procesy, například procesem diagnostiky, plánování, hodnocení a pod.
4. **aplikační ontologie** řešící konkrétní aplikace, tudíž se jedná o adaptované doménové nebo úlohové ontologie specifikované pro konkrétní úlohu

Dále lze ontologie podle stupně formalizace dělit na zcela neformální či semi-formální, které představují nejčastěji glosáře definující jednotlivé pojmy přirozeným jazykem a míra formálnosti u nich závisí na strukturalizaci dat, a na zcela formální. Otologie vyjádřené pomocí formálního jazyka využívají pak logická pravidla daného jazyka, např. různé logické kalkuly či deskripční logiku. Nicméně i formální ontologie obvykle obsahují část neformální, protože bývají doplněné o dokumentaci, která popisuje obsah ontologie přirozeným jazykem [5]. Ontologie mohou být formálně popsané celou řadou jazyků, mezi nejpoužívanější patří tradiční ontologické jazyky CycL, Ontolingua a OCML nebo značkovací (markup) jazyky SHOE, RDF, RDFS, DAML+OIL a OWL (více viz Kapitola 2.2).

1.3 Definice geo-ontologie

Geoprostorová ontologie, častěji zkráceně označovaná jako geo-ontologie, je specifický případ ontologie zobecněné na doménu geografie a geoprosotorvých dat. Pracuje tedy s takovými koncepty, které odpovídají prostorově lokalizovatelným objektům a jevům na zemském povrchu z oblasti fyzické i humánní geografie a definuje jak prostorové, tak i významové (sémantické) vztahy mezi nimi. Tyto objekty a jevy mohou být přírodního nebo antropogenního původu, vymezené přirozenými nebo uměle vytvořenými hranicemi a v prostoru a čase spojité či diskrétní [14]. Jak bylo zmíněno v předchozí sekci, jedním z cílů ontologie je vytvořit model konceptů, který je věrnou reprezentací reálného světa. Totéž platí i pro vytvoření geografického modelu reálného světa, kde navíc vystupují i prostorové informace a vztahy jako je například lokalizace prvku, sousednost nebo blízkost prvků, vzájemná propojenost prvků

nebo jejich geometrické vyjádření. Proto musí být u geo-ontologie zajištěna nejenom integrita logických vazeb, ale i celistvost vazeb topologických a mereologických, které vyjadřují prostorové vztahy mezi jednotlivými entitami [15].
V geo-ontologiích se objevují třídy popisující vlastnosti prostoru, jako například typ objektu (přírodní nebo vytvořený člověkem), typ krajiny (např. pobřeží, nížina, pohoří a pod.), povrch (pokrytí nebo porost), využití oblasti (těžba nerostných surovin, rybolov a pod.), typ regionu (formální nebo funkční) nebo použitý souřadnicový systém [14]. Implementované instance jednotlivých tříd pak představují skutečné objekty a jevy, které existují v reálném světě a mají určité vlastnosti.
Geo-ontologie představují moderní způsob pro popis geoprostorových informací. Oproti klasickému ukládání geodat do relačních databází, kde často chybí metadata či koncepční schéma dat, geo-ontologie poskytuje lepší abstrakci prostorových informací, jejich logickou kategorizaci a uspořádání a pro uživatele přináší variabilnější možnosti prostorových dotazů založených na sémantice dat [16]. Kategorizace všech prvků v ontologii a jejich hierarchická struktura spolu se sémantickým přístupem umožňují definovat celou řadu vzájemných vztahů mezi prvky a díky tomu je možné pokládat i složitější prostorové dotazy, které pracují s definovanými koncepty. Díky taxonomické reprezentaci geodat je i snadnější aplikace takové datové sady v GIS, jelikož popis struktury dat je komplexnější a pro uživatele srozumitelnější, zvláště u rozsáhlých a komplikovaných datasetů. Podle [16] by bylo optimální pro takové datasety získat vhodný nástroj pro automatickou extrakci geodat ze stávajících relačních databází, který by následně vytvořil geo-ontologii pro další, ruční i strojové, zpracování dat v GIS.

1.3.1 Přístupy k vytváření geografických konceptů

Vývoj geo-ontologií je významným krokem ke zlepšení možností sdílení geodat a geoinformací napříč rozmanitým spektrem koncových uživatelů. Toho je možné docílit jedině stanovením standardů a harmonizací dat. K tomu slouží vytváření koncepčních schémat, které mají sloužit jako abstrakce reálného světa či jako jeho schematická reprezentace [15]. Proces konceptualizace je zcela zásadní pro vytvoření kvalitní ontologie, která popisuje jak jednotlivé prvky, tak i jejich sémantické a prostorové vlastnosti. K vytváření jádrových geo-ontologických konceptů lze přistupovat podle [14] stejně jako ke geografickým tradicím, které jsou formulovány dle [17] následovně (pozn. překlad anglických termínů je převzatý ze slovenských skript [18]):

1. tradice prostorové vědy - k místu se přistupuje jako k objektu, který je možné vhodně geometricky reprezentovat (pomocí bodů, linií a polygonů) a lze přesně určit jeho polohu buď absolutně pomocí hodnot souřadnic v daném referenčním souřadnicovém systému, nebo pomocí vnějších vtahů k ostatním místům (geografickým objektům)

2. tradice areálové (regionální) vědy - rozděluje svět na územně malé jednotky, které se vyznačují určitou specifickou vlastností či přítomností nějakého jevu nebo prvku, např. funkční regiony typické svým využitím (zemědělství, těžba uhlí a pod.) nebo administrativní území daná hranicí působnosti místního samosprávního úřadu

3. tradice syntetické vědy - popisuje svět z pohledu vzájemných vztahů a interakcí mezi lidskou společností a přírodním prostředím, např. na základě faktorů jako půdní kryt (*land-cover*) či využití půdy (*land-use*)

4. tradice vědy o Zemi - představuje celostní pojetí fyzického světa, který je složený z různých geomorfologických prvků (vodstvo, vegetace, půda atd.)

Podle [15] lze rozlišit pouze dva základní druhy konceptů:

1. koncepty odpovídající fyzickým jevům a objektům (tzv. fyzické koncepty) - dále se dělící na:

   - koncepty jednotlivých geografických objektů, které jsou jednoznačně ohraničené buď prostorovou nespojitostí (např. hora, kaňon) či kvalitativní rozdílností (např. jezero, poušť) od okolních objektů
   - koncepty jevů spojitých v prostoru (např. teplota, sklon a svažitost, husotota osídlení)

2. koncepty představující jevy vytvořené společností reprezentující různé sociální a institucionální konstrukty (tzv. sociální koncepty) - dále se dělíci na:

   - koncepty jednotlivých objektů uměle vykonstruovaných buď institucionálními nebo právními konvencemi (např. parcela, samosprávní oblast, administrativní jednotka)
   - koncepty prostorově spojitých jevů, které jsou vymyšlené lidmi (např. míra kriminality, míra úmrtnosti novorozenců)

Toto rozdělení konceptů úzce souvisí s kategorizací geografických objektů podle způsobu jejich ohraničení - bona fide (přirozenými) či fiat (uměle vytvořenými) hranicemi [19].
Pro vytváření konceptů je nutné rozlišovat, zda se na daná místa nahlíží zvenčí či zevnitř. V [20] je jako konkrétní příklad uvedený objekt "dům", na který zevnitř může být nahlíženo jako na koncept "domov", ale při náhledu zvenčí se může jednat o koncept "budova". Je vhodné pro konceptualizaci a vytvoření kvalitní geo-ontologie nahlížet na místa objektivně zvenčí a vnímat je jako věci - objekty s určitými vlastnostmi - atributy, jako je třeba jméno, poloha, využití a pod. Každý koncept je pak definovaný svým názvem, jednoznačnou definicí a množinou atributů [15]. Problematikou konceptualizace pro potřeby geo-ontologií se podrobně zabývá také kniha [21].

1.3.2 Formální popis geo-ontologie

Formálně je obecná ontologie podle [16] definována jako pětice $O = \{C, R, HC, rel, A_0\}$, kde

- C je množina konceptů, které reprezentují jednotlivé prvky (entity) ontologie
- R je množina vztahů definovaných mezi koncepty
- HC je hierarchie konceptů (nebo-li taxonomické schéma ontologie)
- rel je funkce, která specifikuje relace v množině R ($rel : R \to C \times C$)
- A_0 je množina axiomů vyjádřených pomocí logického jazyka [2]

Doménová geo-ontologie je pak formálně definována jako obdobná pětice zohledňující navíc geografické koncepty a prostorové vztahy $O_s = \{C', R', HC, rel, A_0\}$, kde

- $C' = C \cup C_s$, kde množina C_s zahrnuje geografické koncepty
- $R' = R \cup R_s$, kde množina R_s zahrnuje prostorové vztahy

1.4 Vybrané příklady geo-ontologií

1.4.1 GeoNames Ontology

GeoNames není jen rozsáhlá geografická databáze geografických názvů a jmen, tzv. *toponym*, která je dostupná online, ale zahrnuje i vlastní ontologii a je plně integrovatelná do sémantického webu. Ontologie GeoNames umožňuje na webové stránky přidávat geoprostorové sémantické informace díky tomu, že každý z více než 11 milionů prvků databáze má svůj unikátní identifikátor URL. Pro každý prvek existují dva URL odkazy, jeden vede na HTML webovou stránku a druhý na RDF popis prvku. Ontologie je implementovaná v jazycích OWL a SKOS a pro definici všech geoprostorových prvků a jejich vlastností využívá termíny ze slovníku W3C *Basic Geo Vocabulary*. Jednotlivé prvky jsou mezi sebou propojené speciálními vazbami (tzv. *linked data*), např. "children" pro vztah kontinent-stát, "neighbours" pro sousední státy nebo "nearby" pro blízko sebe ležící objekty. První verze ontologie vyšla již v roce 2006, tedy necelý rok po vytvoření samotné databáze a v současné době se používá verze 3.2 vydaná 2019 [23].

[2] Logický jazyk je takový jazyk, který se od přirozeného jazyka liší absencí jakýchkoliv nejednoznačností, ať už slov s přeneseným významem či idiomů. V programování se pak jedná o jazyk vycházející z matematické logiky. V případě ontologií můžeme hovořit o kombinaci obou přístupů [22].

1.4.2 FAO Geopolitical Ontology

Geopolitická ontologie je velmi specifický příklad geografické ontologie, jejímž cílem je sémanticky popsat a zpracovávat data z oboru geopolitiky a umožňovat jejich snadnou výměnu. Nejrozsáhlejší geopolitická ontologie je vyvíjena světovou Organizací pro výživu a zemědělství Spojených národů (dále FAO) již od roku 2002 a v rámci veřejně přístupného informačního systému publikuje tematická data týkající se všech členských států OSN. Data jsou ve standardizovaném formátu, aby podporovala jejich interoperabilitu a možnosti sdílení, zároveň jsou neustále aktualizována a publikována v několika světových jazycích.
FAO Geopolitical Ontology konkrétně popisuje a spravuje data týkající se států, teritorií, regionů, sporných území a dalších oblastí, jejich společenství, organizací a ekonomických skupin. Definuje také vzájemné vztahy mezi jednotlivými entitami - například existence společné hranice, členství v organizaci, příslušnost ke skupině nebo správní nadřazenost/podřazenost území. Také řeší historický vývoj vzniku nebo zániku států a uspořádání jejich hranic, základní statistické údaje a faktické informace o jednotlivých oblastech o počtu obyvatel, měně, rozloze, HDP, atd. Je implementována v jazyce OWL a vymodelována podle zásad RDF, kdy každý prvek ontologie má pak unikátní identifikátor URI a plně odpovídá standardům W3C. FAO Geopolitical Ontology je díky užití standardů vhodnou předlohou pro další ontologie zaměřené na data popisující geopolitické entity [24].

1.4.3 Projekt TOWNTOLOGY

Projekt TOWNTOLOGY představuje možnou aplikaci (geo-)ontologií pro potřeby městského plánování a managementu. Prvně byl představen v roce 2005 a spadá pod agendu jedné z tzv. COST Actions mezinárodní organizace COST (European Co-operation in the field of Science and Technology). Ideou projektu je vytvořit několik doménových ontologií z oblasti městského inženýrství, které zcela dokážou obsáhnout veškeré aspekty spojené se správou měst. Měly by tedy pokrýt veškeré potřeby nejenom developerů a projektantů, ale i obyvatel měst, komunálních politiků, odborníků na městské plánování a dopravu, ekologů a dalších zainteresovaných osob.
V rámci projektu vznikly tři návrhy ontologií:

- **Road System Ontology** zastřešující veškerou dopravní infrastrukturu města včetně veřejné dopravy
- **Urban Mobility Ontology** zabývající se možnostmi pohybu po městě s důrazem na možnosti pěší přepravy, infrastruktury pro cyklodopravu, hendikepované osoby, možnosti parkování a částečně i turismus

- **Urban Renewal Ontology** řešící různé aspekty obnovy měst z mnoha pohledů (urbanistického, ekologického, právního, historického, sociálního, ekonomického, institucionálního a dalších).

Je nutné zdůraznit, že se jedná o případové studie, které vznikly pro akademické účely a nejsou využívané v praxi, pro kterou byly určeny [25].

1.4.4 Europian Road Ontology EUROTL

Do praxe ale byla uvedena ontologie, nebo spíše ontologický rámec (*framework*) EUROTL, jehož hlavním úkolem je spojovat specifičtější doménové ontologie týkající se dopravy, jako např. ontologie INSPIRE Road Network [3], ontologie infrastruktury podle datasetu AM4Infra [4] nebo PROV Ontology (PROV-O) [5] zabývající se proveniencí (místem původu) a je aplikovatelná pro různé domény.
Samotná EUROTL obsahuje pak takové definice, které jsou často opakované v sub-ontologiích, které spojuje. Jedná se například o koncepty provenience, používaných jednotek, časové a prostorové lokalizace nebo životního cyklu včetně k nim přidružených vlastností. Tyto koncepty pak již není nutné znovu definovat v jednotlivých doménových ontologiích a připojí se pomocí tzv. *linking rule sets*. Všechny ontologie sdružené v rámci EUROTL jsou vytvořeny podle standardů W3C [26].

1.5 Deskripční logika a mereotopologie

Aby bylo možné používat ontologie pro reprezentaci znalostí, je nutné k tomu používat formalizované a logické formulace. Použití přirozeného jazyka je naprosto nevhodné, neboť je velmi složitý, ovlivněný historickým vývojem, může být tedy nejednoznačný a obsahuje ve své syntaxi řadu výjimek. Proto se jako reprezentační jazyky zavádí jazyky umělé založené na matematické logice, s jednoznačným syntaxem i sémantikou. V ontologiích se jako reprezentační jazyk užívá deskripční logika, pomocí které lze vyjadřovat i mereotopologické vztahy, a proto se v následujícím textu zaměřím právě na tyto pro ontologie typické formalismy.
Jak mereotopologie, tak i deskripční logika mají svůj základ v matematické logice a predikátové logice prvního řádu, a proto zde nejprve uvedu jejich stručnou charakteristiku. Logika obecně je vědní disciplína zabývající se způsoby vyvozování závěrů a uplatňuje se v řadě dalších vědních oborů včetně matematiky, filozofie a počítačových věd. Matematická logika pak s využitím logiky a logických výrazů exaktně popisuje a definuje matematické vztahy,

[3] dostupná z: $https://www.roadotl.eu/static/eurotl-ontologies/inspire_doc/index-en.html\#overv$
[4] dostupná z: $https://www.roadotl.eu/static/eurotl-ontologies/AM4Infra_doc/index-en.html$
[5] dostupné z: $https://www.w3.org/TR/prov-o/$

provádí důkazy, tvoří axiomy a teorie. Klíčová je právě exaktnost a formálnost logických výrazů a tím i její univerzálnost a použití v řadě dalších oborů, protože umožňuje interpretaci výrazů bez konotací. Základní pojmy z oboru matematické logiky jsou uvedené v Příloze 1 a podrobněji se jí zabývá například kniha [27].
Predikátová logika prvního řádu (*First Order Logic*) je pak nástavbou na matematickou logiku, kterou obohacuje o zavedení kvantifikátorů a predikátů. Predikáty jsou taková tvrzení, o kterých má smysl rozhodovat, zda jsou pravdivá či nikoliv, a jedná se tedy o obdobu funkce nabývající booleovské hodnoty (*true* $\times$ *false, 1* $\times$ *0*). Místo konkrétních pojmenovaných entit se používají proměnné, za které pak lze dosazovat z daného universa či domény, a je možné je kvantifikovat ($\forall, \exists$). Tyto proměnné vstupují do predikátů a po vyhodnocení je výstupem buď pravda nebo nepravda. Kromě kvantifikátorů se používají také různé logické symboly ($\wedge, \vee, \neg, \rightarrow, \equiv, atd$) se vždy stejným významem, ale i nelogické symboly pro označení predikátů, funkcí a konstant, jejichž význam je různý podle interpretace. Predikáty logiky prvního řádu jsou úzce spjaté s teorií množin. Existují i logiky vyšších řádů, které navíc dovolují predikáty jako argumenty uvnitř predikátů nebo kvantifikaci predikátů a funkcí [52].

1.5.1 Deskripční logika

Pod pojmem deskripční logika je zastřešená celá skupina vzájemně příbuzných logických formalismů, tedy formálních jazyků pro reprezentaci znalostí, které jsou typicky používané v ontologiích. V současnosti nejpoužívanější ontologický jazyk OWL 2 je na deskripční logice založený. Stejně tak se hojně využívá ve znalostních systémech určených pro umělou inteligenci, neboť pomocí deskripční logiky lze formálně popsat příslušné koncepty a vztahy mezi nimi pomocí axiomů - logicky formulovaných platných tvrzení. Jedná se o specifickou logiku pojmů nebo také logiku konceptů [2], která se snaží vystihnout vztahy mezi jednotlivými pojmy, které jsou k sobě navzájem v různém postavení (nadřazený X podřazený pojem, obecné označení X konkrétní příklad). Základními modely používanými v deskripční logice jsou koncepty, role (tj. vztahy či relace) a individuály, které tvoří základní prvky formulovaných axiomů. Takto formulovaná tvrzení jsou vždy rozhodnutelná.
Poprvé se pojem deskripční logika objevil v 80. letech, ale v podstatě se jednalo pouze o nové pojmenování již existujících formalismů pro popis reprezentace znalostí. Dosavadní jazyky či rámce sloužící pro popis sémantiky dat nebyly formální, a proto bylo nutné zavést nový způsob jejich popisu pomocí formálního jazyka založeného na matematické logice. Během 80. a 90. let vznikala celá řada systémů založených na deskripční logice a postupem času se vylepšovaly i algoritmy používané pro tzv. *reasoning*, tedy pro odvozování nových explicitně neuvedených vztahů na základě existujících axiomů v daném systému. Reasonery používané v ontologiích v současné době (např. FaCT, FaCT++, Pellet) jsou již velmi rychlé a efektivní

a jsou založené na tzv. tablových algoritmech.
Existuje celá řada variant deskripční logiky, které se odlišují v tom, jaké logické vztahy je možné pomocí nich popsat, tedy v jejich expresivitě. Označují se typicky velkými písmeny, které v angličtině označují určitou množinu vztahů, která lze pomocí dané varianty popsat a kromě základních deskripčních logik existují i různé verze s rozšířeními. Konvence pojmenovávání různých variant deskripční logiky a přehled používané syntaxe je uvedený v příloze bakalářské práce. Základní příručkou pro deskripční logiku je pak rozsáhlá učebnice [34].

1.5.2 Mereotopologie

Mereotopologie představuje predikátovou logiku prvního řádu (*first-order logic*), která slučuje mereologické [6] a topologické [7] vztahy a její principy vychází jak z filozofických, tak matematických základů. První filozofické úvahy o vztahu mezi celkem a jeho částmi se datují k samým počátkům ontologie a metafyziky z dob Platona a Aristotela, ovšem největší rozvoj nastal až v 19. století, kdy se matematici snažili o formální vyjádření meronomických vztahů a kdy vznikla první teorie množin [29]. Za zakladatele mereotopologie jako takové je považován Alfred North Whitehead, který se jako první pokusil propojit vztah celek-část s topologickými pojmy jako sousednost, přilehlost či spojení. Na jeho práci navázal Bowman L. Clarke tím, že Whiteheadova tvrzení formalizoval a tím vznikla současná podoba mereotopologie [30].
Aplikace mereotopologie v ontologiích umožňuje formulovat zákonitosti týkající se hraničení, vzájemné propojenosti nebo sousednosti objektů a dalších vzájemných vazeb mezi jednotlivými objekty, jako je příslušnost části k celku, definuje vnitřní prostor celku a zavádí koncept povrchu, díry nebo bodu. Mereotopologie představuje v případě ontologií alternativu k formálnímu popisu struktury reálného světa pomocí teorie množin, které neumožňují úplné pokrytí všech mereologických a topologických vztahů. Význam mereotopologie lze vnímat i tak, že pomocí definování celku a jeho částí spolu s jednoznačným vymezením hranic, ať už *fiat* či *bona fide*, lze rozdělit libovolnou doménu na objekty nebo oblasti a ty dále dělit na další části a elementy, a tím získat jádrovou strukturu ontologie. Pomocí axiomů a definic lze formálně vyjádřit řadu skutečností a vztahů, které vymezují jevy jako spojitost, omezenost, propojenost, objem prostoru a pod. [30]. V Příloze 1 lze najít mereotopologické axiomy a definice vycházející z práce Barryho Smitha, který vydal ucelený přehled formulí a tvrzení sloužících k vyjádření popisu struktury reálného světa vycházející z myšlenek a práce celé řady autorů (A. C. Varzi [31], R. Casati [32], C. Eschenbach [33], ...).

[6] Mereologie je formální teorie vztahů mezi částí a celkem, zabývá se tedy vztahem část-celek, ale i vztahy mezi jednotlivými částmi celku [28].

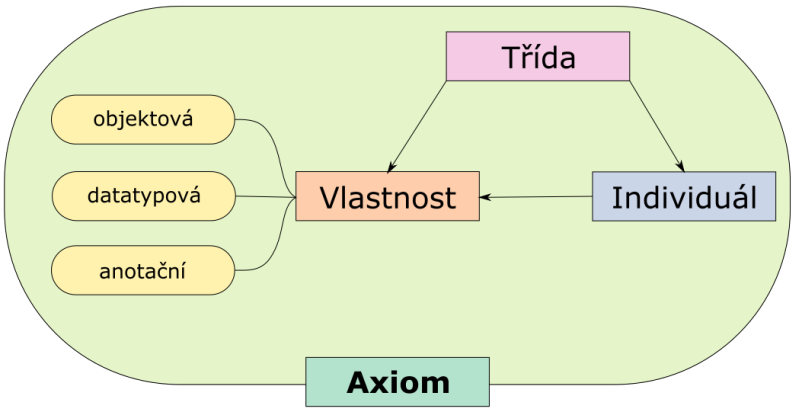
[7] Topologie je matematický obor, který se zabývá prostorem, ale na rozdíl od geometrie a metriky se nezabývá velikostí, tvarem nebo směrem, ale takovými vlastnostmi prostoru, které se při spojitých deformacích nemění [29].

2 Návrh ontologie

2.1 Struktura ontologie

Jako první blíže určil taxonomickou strukturu ontologie Thomas Gruber počátkem 90. let [6]. Základní struktura ontologií se ani v současné době a bez ohledu na použitý nástroj a jazyk pro jejich implementaci zpravidla zásadně neliší od původně navrženého řešení. Ovšem pro označení jednotlivých součástí ontologie se používají různé termíny, které navíc nemusí mít ve všech případech české ekvivalenty. Zjednodušeně bez vymezení jednotlivých stavebních prvků lze strukturu ontologie rozdělit na 3 části (*boxy*), které dohromady uchovávají jak samotná data, tak i souvislosti mezi nimi [35]:

- T-Box (*terminology/taxonomy box*) = terminologická složka, která zahrnuje všechny třídy včetně jejich hierarchického uspořádání a datotypových vlastností
- A-Box (*assertional box*) = složka individuálů, která definuje konkrétní instance tříd a jejich popis
- R-Box (*role box*) = složka rolí a vztahů, která definuje objektové vlastnosti tříd a individuálů



Obrázek 2.1: Prvky ontologie. Převzato z [36]

Schéma 2.1 pak znázorňuje obecnou strukturu ontologie. V následujícím textu se zaměřím hlavně na pojmy užívané v editoru ontologií *Protégé*, který využívá jazyk OWL 2, protože

pomocí nich je vytvořena geo-ontologie veřejné správy a samosprávy a obecně se jedná o široce používaný nástroj. V závorkách jsou pak uvedené další používaná označení stejného prvku ontologie.

Třídy (kategorie, koncepty, rámce)

Třídy jsou hlavní složky ontologie, na kterých je vybudovaná celá taxonomie (hierarchie) nadtříd a podtříd. Představují soubor individuálů se stejnými vlastnostmi a obsahují formální pravidla, co musí individuál splnit, aby náležel dané třídě. Popisují tedy vlastnosti jednotlivých položek ontologie. Třídy jsou propojené nejenom mezi sebou, ale i s jednotlivými individuály. Na rozdíl od objektově orientovaného programování třídy neobsahují procedurální metody, společné ale je pro objektově orientované programování i ontologie schopnost dědění. Základním prvkem hierarchie musí být kořenový prvek (*root element*) reprezentovaný svojí třídou, např. v jazyce OWL 2 je kořenovým prvkem *owl: Thing*, který tvoří nadtřídu pro všechny ostatní třídy. Všechny třídy a individuály jsou potomky kořenového prvku. Rozlišují se dva typy tříd [37]:

- **primitivní třídy**, které tvoří primitivní kostru ontologie, jsou založené na logických operacích a mají definované jen nutné podmínky
- **definované třídy**, které mají definované jak podmínky nutnosti, tak i postačitelnosti pro příslušnost individuálu do třídy

Individuály (individua, instance, jedinci, entity)

Individuály reprezentují konkrétní objekty z reálného světa či jeho části (domény), kterou daná ontologie zpracovává. Představují specifické jedince či fakta, které již není možné dále dělit a jsou tedy koncovými prvky hierarchie. Individuály by měly příslušet nějaké třídě a vyhovovat definici konceptu, který tato třída reprezentuje. V ontologii se ale mohou vyskytovat i jedinci bez příslušnosti ke konkrétní ontologické třídě, ačkoli vždy jsou alespoň potomky kořenové třídy *owl: Thing*, nebo naopak mohou spadat do několika tříd zároveň. Určení, zda daný prvek ontologie bude definovaný jako třída či jako individuál, je často subjektivní a závisí na úhlu pohledu tvůrce ontologie, kontextu situace nebo podrobnosti tvořeného modelu.

Axiomy (tvrzení, postuláty, pravidla)

Axiomy představují všechna explicitní tvrzení a pravidla, která jsou v dané ontologii pravdivá a vymezují vztahy mezi jednotlivými prvky ontologie. Axiomy vymezují jak datatypové a anotační, tak i objektové vlastnosti tříd (např. ekvivalenci či disjunkci tříd), vztah subsumpce nebo rozklad na podtřídy. Axiomy v sobě zahrnují nejenom třídy a individuály,

ale i soubor vlastností. Lze říci, že ontologie je množinou platných axiomů. Axiomy jsou formulované pomocí logického jazyka, aby bylo možné je jednoznačně interpretovat.

Vlastnosti (role, funkce, atributy, sloty, relace)

Vlastnosti navzájem propojují jednotlivé prvky ontologie (tj. třídy a individuály) a zároveň je i blíže popisují. Vlastnost nemůže existovat samostatně bez vazby na nějakou třídu či individuál, který specifikuje. I vlastnosti mohou mít vlastní hierarchii a lze definovat specializovanější vlastnosti již dříve deklarovaných funkcí a vztahů. Vlastnosti lze rozdělit na 3 skupiny:

- **objektové vlastnosti** - specifikují vztah mezi dvěma entitami (tj. individuály nebo třídami) a určují strukturu ontologie
- **datatypové vlastnosti** - omezují datový typ hodnoty konkrétního jedince (tj. zda se jedná o textový řetězec, celé nebo desetinné číslo, případně výčet hodnot, kterých může objekt nabývat a pod.)
- **anotační vlastnosti** - představují dodatečná metadata obsahující popis dat, vysvětlení pojmů, alternativní terminologii nebo ekvivalenty v cizím jazyce, odkazy na zdroje atd.

Nejrozsáhlejší skupinu obvykle tvoří objektové vlastnosti, které definují sémantické vztahy a integritní omezení. Je pro ně možné definovat definiční obor a obor hodnot, tedy například se specifikuje, z jakých tříd mohou pocházet propojovaní jedinci. Objektové vlastnosti pak lze dále charakterizovat jako:

- **inverzní** - jestliže prvek A je vlastností x spojený s prvkem B a prvek B je spojený s prvkem A vlastností y, pak x a y jsou inverzní vlastnosti
- **funkcionální** - pro daný prvek vlastnost nabývá jen jedné hodnoty, tj. pro každý prvek z definičního oboru existuje jen jedna hodnota z příslušného oboru hodnot
- **tranzitivní** - je-li prvek A ve vztahu s prvkem B a prvek B je ve vztahu s prvkem C, pak s prvkem C je ve vztahu i prvek A
- **symetrické** - jestliže vlastnost propojuje prvky A a B, pak ta samá vlastnost musí propojovat i prvek B s prvkem A (opakem je antisymetrická vlastnost)
- **reflexivní** - pomocí reflexivní vlastnosti může být prvek propojený sám se sebou (opakem je antireflexivní vlastnost)

Vlastnosti jako takové pak samy mohou být dále omezené oborem hodnot, datovým typem nebo kardinalitou [37].

Sémantické vztahy v ontologiích

Sémantické typy vazeb mezi prvky ontologie vznikají na základě vlastností jednotlivých prvků a taxonomické struktury ontologie. Podle [37] se rozlišují následující sémantické vztahy:

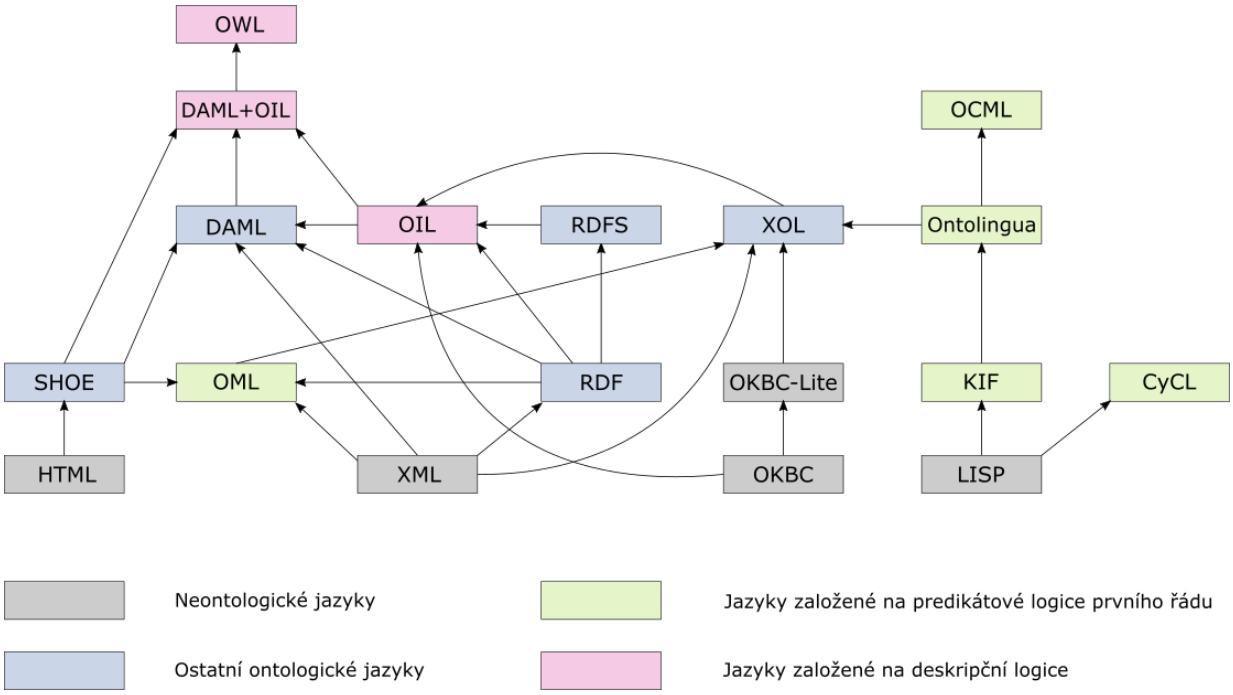
- **hyponymie** - vyjadřuje vztah podřazenosti prvku jinému prvku, například příslušnost specifické třídy třídě obecnější (nadtřída město - podtřída statutární město) a tedy vyjadřuje vztah subsumpce, bývá nejčastěji vyjádřená vztahem IS-A (tedy *statutární město* IS-A *město*)
- **synonymie** - značí podobnost významů 2 nebo více prvků (např. *primátor* má podobný význam jako starosta statutárního města)
- **homonymie** - znamená stejný význam 2 nebo více prvků (např. prvek *městský obvod* je totožný s prvkem *městská část*)
- **meronymie/holonymie** - vyjadřuje vazbu část-celek, tedy stav, kdy je jeden prvek částí jiného prvku (např. *obec s pověřeným obecním úřadem* je částí prvku *obec s rozšířenou působností*)
- **hypernymie** - je opakem hyponymie a vyjadřuje nadřazenost jednoho prvku druhému (např. třída *okres* je nadřazený individuálu *Plzeň-sever*)

Klasifikátory (reasonery)

Další součástí ontologií jsou tzv. klasifikátory, které ovšem nepatří mezi základní prvky ontologie, nýbrž představují nástroj pro kontrolu konzistence ontologie po jejím vytvoření. Kromě prověření splnění všech vlastností a vztahů deklarovaných axiomy umožňuje následně i odvodit další informace, které nejsou v ontologii výslovně uvedené. Klasifikátory dokážou odhalit nesplnění disjunkce tříd, špatné určení domény či oboru hodnot, a prověřují správné provedení vztahu subsumpce nebo zda daná třída může mít nějakou instanci.
Proces odhalování nových vztahů mezi prvky na základě deklarovaných vlastností se nazývá *reasoning*, a proto se pro klasifikátory používá označení *reasoner*. Existuje celá řada klasifikačních nástrojů, například editor Protégé využívá nástroje *FaCT++ (Fast Classification of Terminologies)*, *Hermi* či *Pellet* [36]. Srovnání různých klasifikačních nástrojů včetně zhodnocení jejich kladů a záporů, funkčnosti a možností aplikace lze najít v [38]. Na některé nedostatky klasifikátorů jsem narazila také během realizace Geo-ontologie veřejné správy a samosprávy - viz kapitola 4.4.

2.2 Jazyky pro návrh ontologie

Je nutné zapisovat ontologie pomocí nějakého formalizovaného jazyka, který umožní naplnit samotnou podstatu myšlenky ontologie, tedy hlavně její sdílitelnost a znovupoužitelnost. První jazyky pro tvorbu ontologií se objevily již počátkem 90. let a vycházely hlavně z predikátové logiky prvního řádu. Po vydání značkovacích jazyků HTML, XML a RDF, které jsou zároveň webovými standardy vydané konsorciem W3C, se začaly rozvíjet ontologické jazyky založené na těchto třech standardech. Na následujícím schématu 2.2 je pak znázorněná evoluce ontologických jazyků a jejich kategorizace podle logického přístupu k formulaci axiomů.



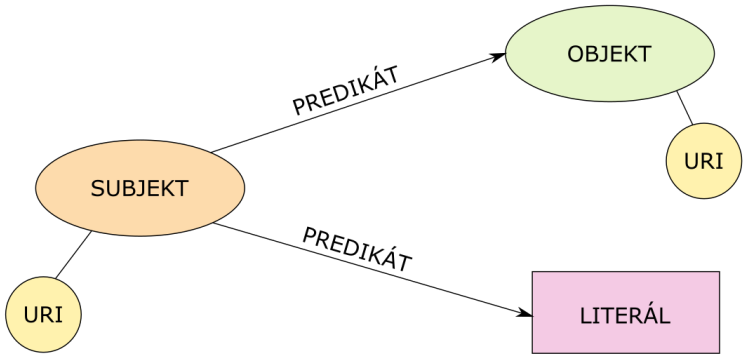
Obrázek 2.2: Vývoj a kategorie ontologických jazyků. Převzato z [37]

V současné době je nejrozšířenějším jazykem *Web Ontology Language OWL*, který je spolu s RDF podporovaný nástrojem Protégé. V dalším textu se tedy budu věnovat hlavně těmto jazykům. Více informací o jednotlivých ontologických jazycích a standardech lze najít například v [5], [37] nebo [39].

2.2.1 Resource Description Format RDF

Resource Description Format, do češtiny překládaný jako rámec pro popis zdrojů, je standardizovaný jazyk a zároveň i datový model určený pro popis dat a datových zdrojů v prostředí webu. Datové zdroje pocházející z webových stránek reprezentuje ve formě metadat.

RDF dokáže popisovat nejenom webové stránky a dokumenty, ale i webové služby a další libovolné zdroje dat. Pomocí RDF je možné definovat sémantické vlastnosti dat, což se využívá právě v ontologiích, kde tvoří jakousi základní úroveň, na které se dále staví pomocí jazyka OWL (viz dále). Obecně lze říci, že každá informace na webu, kterou je možné jednoznačně identifikovat, je pak možné pomocí RDF namodelovat včetně všech jejích vlastností a vztahů k dalším informacím.
Základem RDF je definování trojic **Subjekt - Predikát - Objekt**, které vyjadřují jednoduchá tvrzení a jejich skládáním je možné sdělovat tvrzení složitější. Subjektem je vždy daný popisovaný zdroj, který musí mít přiřazený svůj URI identifikátor. Predikát pak určuje vztah mezi subjektem a objektem (např. přiřazení určité vlastnosti). Objekt představuje hodnotu, kterou subjekt nabývá, je-li predikát splněný. Objektem může být jiný datový zdroj identifikovaný svým URI či jakýkoliv literál (tj. řetězec, celé nebo desetinné číslo, datum, a pod.). RDF trojice je znázorněna na schématu 2.3. Samotný RDF nedefinuje svoji syntaxi, a proto se zapisuje několika různými způsoby. Tím nejrozšířenějším je pak *RDF/XML*, protože se jedná o standardizovaný značkovací jazyk, který je strojově čitelný a přenositelný bez omezení mezi platformami. Další možnosti zápisu RDF jsou například použití ohodnoceného orientovaného grafu, jehož prvky tvoří definované trojice, což je dobře čitelné pro člověka, ale nečitelné pro stroj, nebo speciální syntaxe N-Triples či Notation 3 [37].



Obrázek 2.3: RDF trojice. Převzato z [41]

RDF samo o sobě neumožňuje popsat vazby mezi zdroji a jejich vlastnostmi, a proto na něj byla vytvořena nástavba v podobě *RDF Schema*. RDFS navíc přidává možnost definovat vztah podřazenosti mezi třídami zdrojů i mezi predikáty (tj. vztah subsumpce) a také pro vlastnosti stanovit definiční obor a obor hodnot, a tím pokrýt základní sémantický popis

struktury ontologie. Další omezení vlastností (např. kardinalitu, disjunktnost tříd, podmínky příslušnosti k třídě a pod.) ale už definovat neumí a je třeba využít prostředků dalšího ontologického jazyka, typicky pak OWL [5].

2.2.2 Web Ontology Language OWL

Web Ontology Language označovaný přesmyčkou jeho akronymu jako OWL je dalším standardem z dílny WC3. Prvně byl publikován v roce 2004 a stal se nástupcem jazyka DAML+OIL. V roce 2008 pak byla na základě podnětů uživatelů vytvořena nová obohacená verze jazyka označovaná jako OWL 2. Jak je zřejmé z předchozího textu, OWL vychází z jazyků RDF a RDFS, které začleňuje do vlastních konstruktů a dále je rozšiřuje, neboť OWL je jazyk bohatší na sémantiku. Z RDF(S) přebírá jeho syntaxi.
Původní jazyk OWL se dělí na tři druhy (dialekty), které se liší mírou expresivity a tím i náročností procesu odvozování nad ontologií implementovanou v daném dialektu.

- **OWL Lite** je nejjednoduší a nejméně expresivní varianta jazyka OWL, která je určená pouze pro vytvoření základní taxonomie. Z RDFS přebírá strukturalizaci na třídy a vlastnosti, vymezení oboru hodnot a definičního oboru a možnost použít jen jednoduchou kardinalitu. Takto zapsaná ontologie je pak velmi jednoduchá na zpracování pomocí klasifikátorů.

- **OWL DL** (*Description Logic*) představuje kompromis mezi plnou verzí jazyka, která je velmi expresivní a náročná na výpočetní zpracování, a takovou implementací, která zaručuje odvozování nových vlastností pomocí klasifikátorů. OWL DL zahrnuje kompletně OWL Lite a k tomu přidává možnost disjunkce, výčtových typů, další možnosti kardinality vazeb a negaci výroků.

- **OWL Full** je kompletním sjednocením jazyka OWL s RDF a RDFS, které obsahuje všechny elementy a konstruktory těchto jazyků. Pro svoji velkou vyjadřovací sílu je ovšem velmi náročný na zpracování a nezaručuje možnost dalšího odvozování nad ontologií.

V nové verzi OWL 2 je zachovaný pouze dialekt OWL Full a OWL DL a k nim navíc přidává další prvky a nové možnosti. Hlavním zjednodušením je odlehčená syntaxe, která umožňuje snadnější zápis některých stávajících konstruktů. Tato verze umožňuje efektivnější užití klasifikátorů k odvozování nových vztahů a vlastností, aniž by byla snížena vyjadřovací síla jazyka. Přibyly zde také nové konstruktory pro popis vlastností, pro datové typy je nyní možné omezit jejich rozsah a u vazeb lze stanovit jejich kardinalitu bez omezení či definovat vazbu prvku na sebe sama (unární vazba). OWL 2 nabízí oproti první verzi rozšíření anotací

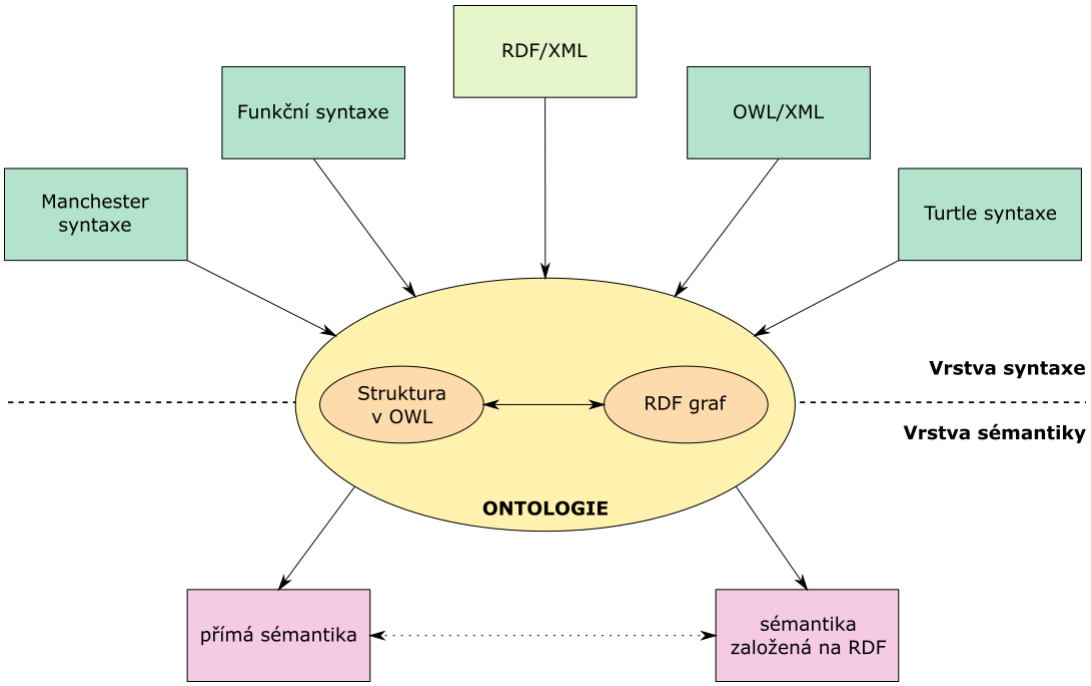
i na axiomy. Dalším rozdílem je, že povoluje pojmenovávat třídy a jedince stejným názvem. Jazyk OWL 2 zavádí 3 jazykové profily, které jsou vhodné pro různé situace:

- **EL** je vhodný pro práci s třídami podle pravidel deskripční logiky EL++ a dokáže zpracovat i rozsáhlejší ontologie, ovšem za cenu nižší expresivity
- **QL** (*querry language*) se používá pro vyhodnocení dotazů nad daty, jejichž struktura vychází z modelu relační databáze, kde se pro menší počet tříd ukládá velké množství specifických individuálů
- **RL** se využije pro práci s daty uloženými v podobě RDF trojic a umožňuje jejich uložení v podobě grafu

Strukturu jazyka lze rozdělit na sémantickou a syntaktickou vrstvu, kdy každou z vrstev je možné získat několika možnými způsoby. I samotná ontologie může být vytvořena buď pomocí specifikace její struktury jazykem OWL 2 a nebo může být uložená v RDF grafu. Obě tyto různé struktury je možné mezi sebou vzájemně převádět pomocí speciálního nástroje, např. v dokumentu [43] je sepsaný přehled konstruktorů v jazyce OWL 2 a jejich ekvivalent v syntaxi RDF. Syntaktická vrstva má za úkol umožnit popsat a uložit strukturu ontologie v takové podobě, která bude přenositelná mezi uživateli a aplikacemi. Primárním výměnným syntaktickým formátem je RDF/XML, který je povinně podporovaný všemi nástroji pro návrh ontologií. Existují další 4 možné syntaxe:

- **OWL/XML** pro zpracování XML nástroji
- **funknčí syntax** zachycující formální strukturu pro aplikaci klasifikátorů
- **Manchester syntax**, který je snadno čitelný a srozumitelný pro člověka
- **Turtle syntax** užívaný pro zápis RDF trojic

Co se týká sémantiky, tak je možné ji do struktury ontologie zahrnout dvěma možnými způsoby. První možností je zahrnout významovou složku dat přímo do struktury ontologie, ale je nutné dodržet všechna pravidla použité deskripční logiky EL++. Druhou možností je použít sémantiku založenou na RDF a přidat význam dat do RDF grafu reprezentujícího strukturu ontologie. Pro lepší představu, jak vypadá struktura jazyka OWL 2, slouží schéma 2.4, které je přejaté z oficiální dokumentace jazyka od konsorcia W3C [42].



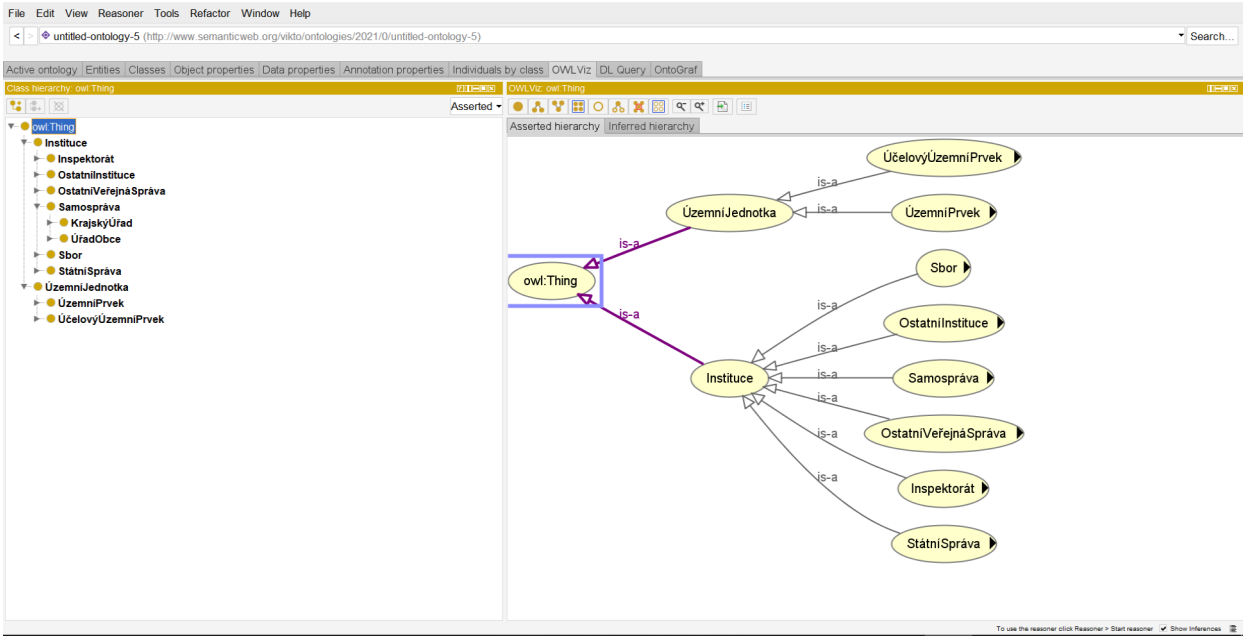
Obrázek 2.4: Struktura jazyka OWL 2. Zjednodušené schéma převzaté z [42]

2.3 Editor Protégé

Pro vytvoření geo-ontologie v rámci této bakalářské práce jsem využila nástroj Protégé (verze 5.5.0). Software je vyvíjený univerzitním výzkumným centrem *Stanford Center for Biomedical Informatics Research at the Stanford University School of Medicine* již od roku 1987 a je jedním z nejvyužívanějších nástrojů pro editaci ontologií vůbec. Jedná se o volně dostupný software, který je implementovaný v jazyce Java a tudíž je jeho použití nezávislé na operačním systému. Mezi jeho další přednosti patří také to, že plně podporuje standardy konsorcia W3C (RDF, RDFS, OWL 2) a je možné základní verzi programu doplnit o zásuvné moduly (např. plugin pro vizualizaci ontologie OWLViz 5.0.3 nebo plugin Cellfie pro import dat z Excelu). V rámci nástroje Protégé je možné využít několik podporovaných klasifikátorů v rámci procesu odvozování (FaCT++, HermiT, Pellet) a pro otestování správné definice tříd a jejich taxonomie podle pravidel deskripční logiky lze použít plugin DL Querry. Další výhodou Protégé je velmi rozsáhlá a podrobná dokumentace [1], včetně návodů na použití jednotlivých pluginů a rozšíření. Stejně tak díky existenci široké uživatelské základny je k dispozici podpora v podobě tutoriálů a individuálního řešení uživatelských dotazů. V současné době je k dispozici kromě desktopové aplikace i webová verze *WebProtégé*, které jsou vzájemně plně kompatibilní. Webová aplikace navíc obsahuje rozšíření o nástroje umožňující spolupráci více autorů na jedné ontologii a možnost notifikací v případě provedených změn.

[1] dostupné z: *https* : *//protegewiki.stanford.edu/wiki*

Pro potřeby této bakalářské práce nejsou tyto nástroje třeba, a proto byla využita pouze desktopová verze Protégé. Na následujícím obrázku 2.5 je ukázka z uživatelského prostředí editoru.



Obrázek 2.5: Ukázka uživatelského prostředí editoru Protégé rozšířeného o plugin OWLViz

2.4 Metodika návrhu ontologie

Spolu s rozvojem ontologií počátkem 90. let bylo nutné stanovit i standardizované postupy pro jejich navrhování a tvorbu. Sestavování metodik, popis dílčích úkolů a jejich praktické ověření je úkolem ontologického inženýrství. Do současné doby vzniklo velké množství metodik, mezi nejvýznačnější se řadí metodika METHONTOLOGY, On-To-Knowledge, metodika od autorů Uschold a King nebo metodika Cyc. Podstatné ale je, že více než polovina ontologií vzniká živelně bez použití nějaké již existující metodiky, jak vyplývá ze studie [40]. Je ale nutné podotknout, že návrh ontologie se odvíjí od konkrétního záměru autorů, jejich přístupu a potřebám, vymezené doméně a podrobnosti zpracování, předchozích zkušeností autorů s tvornou ontologií a dalších faktorů a je tedy obtížné formulovat jednotný standardizovaný postup, který by byl univerzálně aplikovatelný.
V [2] je shrnutý doporučený postup a základní pravidla, která by měla být dodržena při návrhu ontologie:

1. stanovení účelu a rozsahu ontologie:
    - určení, jaké úlohy se budou pomocí ontologie řešit

- vymezení domény, tj. části reálného světa, kterou má ontologie modelovat
- vyřešení problému rozsahu a podrobnosti, aby nedošlo k bezbřehému nabalování okrajových pojmů a vztahů
- uvedení zdrojů, ze kterých se čerpají data pro naplnění ontologie
- vymezení potenciálních uživatelů ontologie a jejich požadavků
- stanovení tzv. kompetenčních otázek (formulace konkrétních otázek, které by měla ontologie umět zodpovědět)

2. specifikace terminologie (vytvoření glosáře)

   - určení základních termínů, které se vztahují k dané doméně
   - doplnění popisu, vysvětlení a definic základních pojmů
   - pomocí takto vytvořeného glosáře se identifikují ontologické prvky (třídy, individuály a vlastnosti)

3. odlišení ontologických typů

   - vychází z vytvořeného glosáře, který definuje jednotlivé prvky ontologie
   - stanovení, zda bude určitý prvek ontologie deklarovaný jako třída či individuál a v některých případech i jako vlastnost
   - zohlednění přístupu k modelované doméně a účelu hotové ontologie

4. vytvoření taxonomie

   - vymodelování kostry ontologie z jednotlivých prvků na základě vlastností a sémantických vztahů
   - výběr vhodného přístupu k tvorbě taxonomie:
     - postup zdola nahoru (zoběcňování), kdy se konkrétní prvky řadí do obecnějších tříd
     - postup shora dolů, kdy se postupuje od nejobecnějších konceptů až ke konkrétním prvkům
     - postup ze středu, kdy se nejprve zařadí středně obecné koncepty a k nim se přidávají konkrétnější prvky a obecnější nadřazené koncepty

5. podrobná deklarace tříd a vlastností

   - detailní specifikace tříd a vlastností, které se k dané třídě vztahují
   - rozlišení primitivních a definovaných tříd

- stanovení globálních a lokálních omezení
- vytvoření hierarchické struktury vlastností
- využití principu dědičnosti, kdy se vlastnosti nadřazených tříd přenáší na jejich potomky

6. kontrola ontologie a vytvoření dokumentace

   - nasazení klasifikátorů k otestování konzistence ontologie
   - evaluace naplnění stanovených cílů
   - sepsání podrobné dokumentace k ontologii
   - možné automatické vygenerování části dokumentace za použití anotačních vlastností prvků ontologie
   - doplnění dalšího slovního popisu, schémat a pod.

7. nasazení a údržba ontologie

   - možné vytvoření další nástavby na původní ontologii či její verzování

3 Geo-ontologie veřejné správy a samosprávy

3.1 Specifikace geo-ontologie veřejné správy a samosprávy

Tato kapitola přináší základní informace o vytvořené geo-ontologii a blíže popisuje postup při jejím navrhování a následném naplnění daty pro Plzeňský kraj. Při návrhu geo-ontologie jsem postupovala především podle obecné metodiky od Svátka a Vacury [2]. Jako další zdroje informací, jak postupovat při návrhu ontologie, jsem použila odborné články [44] a [45], které vychází hlavně ze zkušeností autorů a ukazují konkrétní zvolené postupy a metody pro tvorbu ontologie. Vzhledem k tomu, že mnou vytvářená ontologie pracuje s administrativním členěním státu, byla pro mne do jisté míry vzorem i geo-politická ontologie FAO [24].
Dále jsem provedla šetření, zda již existuje podobná ontologie nebo alespoň ontologie zahrnující stejné koncepty jako mnou řešená ontologie a zda je možné je mezi sebou propojit či na shodné prvky v obou ontologiích odkazovat. Vzhledem k tomu, že mnou vytvořená ontologie obsahuje třídy specifické pro české prostředí, jejichž definice často vychází ze zákonů, byla nalezena shoda pouze u dvou nejobecnějších tříd „Instituce“ a „Územní jednotka“. Třída „Instituce“ je shodná s třídou „Organization“ v *The Organization Ontology* [1] a třída „Územní jednotka“ pak s třídou „Place“ v ontologii *Schema.org* [2].

3.1.1 Účel a vymezení rozsahu

Doménová geo-ontologie veřejné správy a samosprávy je vytvořena za účelem utřídění působnosti orgánů státní správy a samosprávy a dalších institucí veřejné správy v kontextu územního členění. V rámci rozsahu ontologie se řeší působnost jednotlivých institucí ve smyslu jejich lokace a prostorové působnosti, nikoli ve smyslu jejich aktivit a vykonávaných činností. Agenda řešených institucí a popis výkonu činností státní správy tedy není předmětem této práce a pokud budou zmíněny v rámci popisu dílčích prvků ontologie, pak jen okrajově. Cílem ontologie je určit prostorové vazby mezi úrovněmi administrativního členění kraje a působením jednotlivých institucí, jejichž hierarchie často vychází právě z územního členění regionu. Doménou je tedy územní působnost orgánů veřejné správy a samosprávy. Ontologie

[1] Dostupné z: $https://www.w3.org/TR/vocab-org/\#ontology_reference$
[2] Dostupné z: $https://schema.org/Place$

by měla obsáhnout také administrativní dělení kraje na menší územní jednotky a jednoznačně identifikovat a popsat vztahy mezi těmito územními prvky. Dalším úkolem ontologie je přiřadit takto vymezené územní prvky různým úrovním orgánů státní správy. Výsledná ontologie by měla být použita jako základ pro mapovou aplikaci, která pomůže občanům lépe se zorientovat v územní působnosti jednotlivých úřadů a institucí.

3.1.2 Použitá data a jejich zdroje

Ontologie je naplněna daty týkajících se Plzeňského kraje, ale její struktura je vytvořena tak, aby byla použitelná i pro ostatní kraje České republiky. Tudíž obsahuje i takové prvky, které se v rámci Plzeňského kraje nevyskytují (např. vojenský újezd). Použitá data pochází z veřejně přístupných registrů dat (Registr územní identifikace, adres a nemovitostí RÚIAN), z registrů a databází Českého statistického úřadu a Portálu veřejné správy. Dále jsem informace čerpala z různých zákonů a vyhlášek, které se týkají výkonu státní správy. Konkrétně se pak jedná o tyto zákonné a podzákonné právní normy:

- Zákon č. 51/2020 Sb., o územně správním členění státu
- Vymezení krajů: Ústavní zákon č. 347/1997 Sb., o vytvoření vyšších územněsprávních celků
- Vymezení okresů: Vyhláška č. 564/2002 Sb., o stanovení území okresů České republiky a území obvodů hlavního města Prahy
- Vymezení ORP a POÚ: Zákon č. 314/2002 Sb., o stanovení obcí s pověřeným obecním úřadem a stanovení obcí s rozšířenou působností + Vyhláška 388/2002 Sb., o stanovení správních obvodů obcí s pověřeným obecním úřadem a správních obvodů obcí s rozšířenou působností
- Vymezení obcí: Zákon č. 128/2000 Sb., Zákon o obcích (obecní zřízení)
- Vymezení orgánů státní správy: Zákon č. 2/1969 Sb., o zřízení ministerstev a jiných ústředních orgánů státní správy České republiky

Při návrhu podtříd třídy „Instituce“ mi byly velmi nápomocné přednášky prof. JUDr. Petra Průchy, CSc. z Právnické fakulty Masarykovy univerzity v Brně [46] a dokument „Veřejná správa v České republice“ vydaný Ministerstvem vnitra [47].

3.2 Vytvoření jádrové struktury

Vytvoření jádrové struktury ontologie spočívá ve vymezení a definování hlavních konceptů, jejich vlastností a vazeb mezi nimi. Poté, co jsem měla rozvrženou základní kostru

ontologie, začala jsem vytvářet jednotlivé prvky jádrové struktury přímo v editoru Protégé. Vytváření nových objektů v editoru Protégé je velmi jednoduché a intuitivní, případně je k dispozici kromě dokumentace i celá řada návodů, jak deklarovat nové třídy, vlastnosti a individuály.
Následovalo pak otestování správnosti navržených vazeb a obecně konzistence ontologie včetně odvozování dalších vztahů pomocí reasoningu. V rámci prvotního návrhu ontologie nebyly jako individuály vkládány konkrétní územní jednotky či instituce, ale pouze obecné testovací objekty a to v malém počtu, aby bylo možné odhalit chyby v navržené struktuře a odstranit je v krátkém čase. Nevýhodou všech reasonerů je totiž poměrně dlouhý průběh odvozování, pokud je v ontologii velké množství axiomů, které je třeba vyšetřit. Postupně byly tímto způsobem eliminovány chyby a nedostatky ve struktuře ontologie a bylo pak možno přistoupit k plnění ontologie konkrétními daty. Jednotlivé kroky jsou podrobněji popsané v následujícím textu.

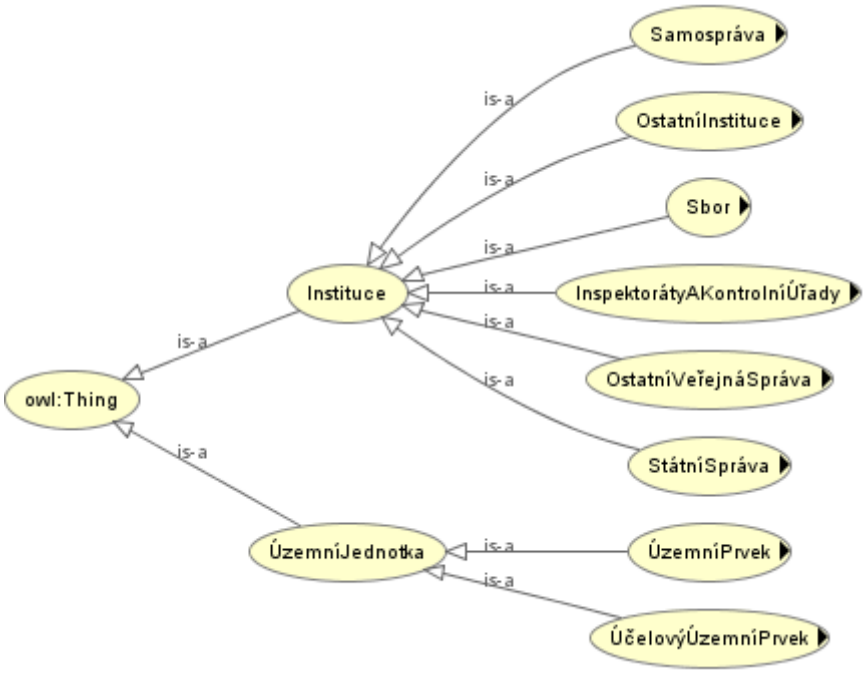
3.2.1 Identifikace jednotlivých prvků ontologie

Jak již bylo zmíněno ve 2. kapitole, základními prvky ontologie jsou třídy, vlastnosti a individuály. Začala jsem s identifikací základních konceptů, které jsou předmětem ontologie a tvoří nejobecnější třídy. Jsou to jednoznačně třídy *Územní jednotka* a *Instituce*, které jsou striktně disjunktní a dále jsem pak pracovala při odvozování jejich podtříd s každým konceptem zvlášť. Vznikly tak dvě oddělené „větve“ hierarchie podtříd, které jsou vzájemně propojené pomocí vztahů definovaných objektovými vlastnostmi.

Konkrétně se jedná o dvojici vzájemně inverzních vlastností *působí v* a *spadá do působnosti*, které určují, která instituce působí v jaké územní jednotce a naopak, která územní jednotka spadá do působnosti jaké instituce. Ještě byla vytvořena jedna dvojice inverzních vlastností *je částí* a *skládá se z* popisující vzájemnou skladebnost územních jednotek. Dále bylo třeba stanovit datatypové vlastnosti, které mají konkrétní entity v ontologii a průběžně vytvářet anotační vlastnosti k jednotlivým prvkům ve chvíli, kdy byly do ontologie vkládány. Tím vznikala okamžitě i metadata popisující jednotlivé objekty a základ pro vytvoření dokumentace ontologie. V poslední řadě, když byla vytvořená struktura ontologie otestována a byla ověřena její konzistence pomocí uměle vytvořených obecných individuálů, byla ontologie naplněna konkrétními individuály pro Plzeňský kraj.

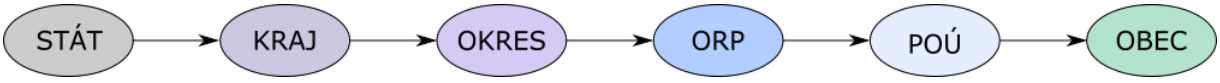
Identifikace podtříd třídy *Územní jednotka*

Při identifikaci podtříd třídy „Územní jednotka“ jsem vycházela pouze ze skutečností definovaných v současně platné legislativě ČR. K 1. lednu 2021 vešel v platnost nový Zákon č.



Obrázek 3.1: Struktura hlavních konceptů ontologie. Vlastní zpracování.

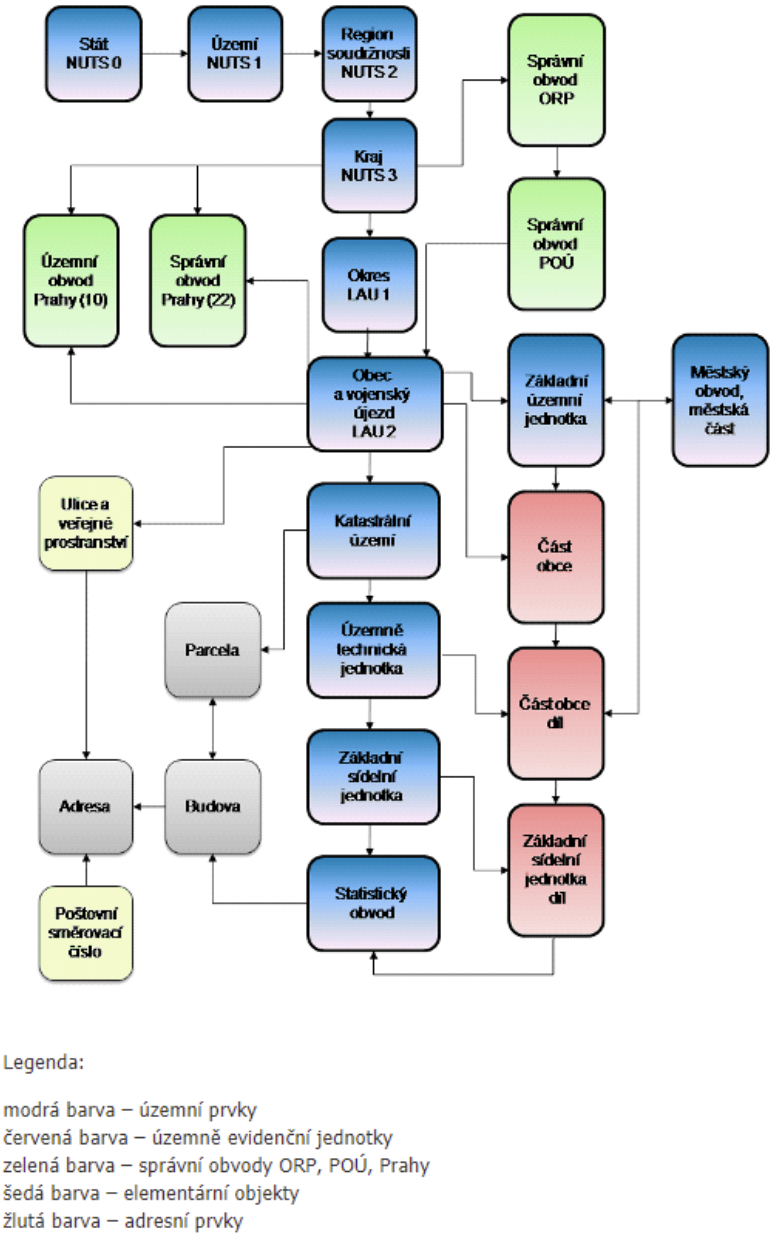
51/2020 Sb., o územně správním členění státu, který nahradil velmi zastaralý Zákon o územním členění státu (č. 36/1960 Sb.). Tento nový zákon by měl zaručovat skladebnost územního členění státu tak, aby každý celek vyšší úrovně byl tvořen pouze celými celky nižší úrovně a nedocházelo k rozdělení nižších územních celků mezi více celků vyšší úrovně. Přesto se stále na území republiky objevují výjimky, kdy není skladebnost znázorněná na schématu 3.2 splněna.



Obrázek 3.2: Skladebnost územního členění státu. Vlastní zpracování dle Zákona č. 51/2020 Sb., o územním členění státu

Jednu z výjimek ze skladebnosti územních jednotek, kterou tento nový zákon odstranil, bychom mohli najít i v Plzeňském kraji, kde se týkala správního obvodu POÚ Holýšov na Domažlicku. Správní obvod POÚ Holýšov ležel v okrese Domažlice, zároveň ovšem spadá pod správní obvod ORP Stod, jehož ostatní správní obvody POÚ se nachází v okresu Plzeň-jih. Neskladebnost územně správního členění státu je v tomto případě tedy způsobena tím, že správní obvod ORP Stod zasahuje do území více okresů. Od nabytí účinnosti zákona došlo k přesunu dotčených devíti obcí Bukovec, Čečovice, Černovice, Holýšov, Horní Kamenice, Kvíčovice, Neuměř, Štichov a Všekary do okresu Plzeň-jih a skladebnost je tedy na celém území Plzeňského kraje již kompletně zaručena. Podobných situací, kdy byl správní obvod ORP rozdělen mezi více okresů, bylo před nabytím účinnosti nového zákona šest a v současné době

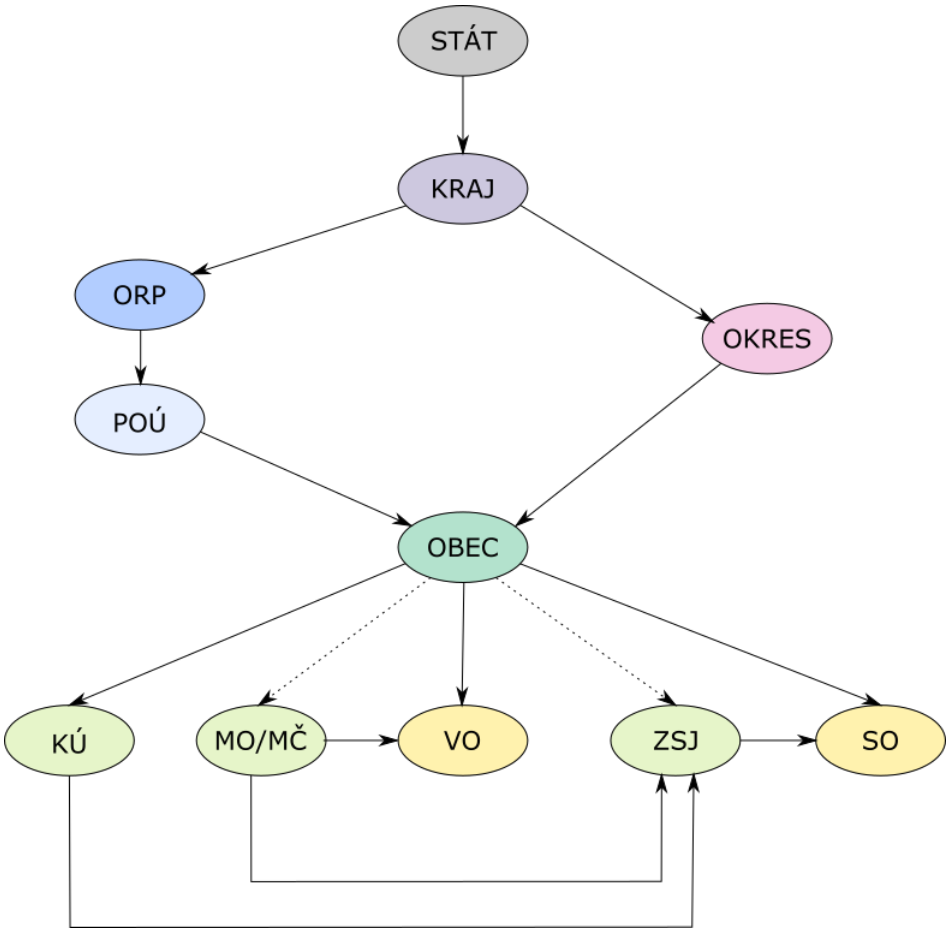
stále existuje jedna výjimka pro ORP Turnov, který zasahuje do území 3 okresů, a náprava by nebyla proto tak snadná jako v ostatních pěti případech. Pokud by v budoucnu došlo k rozšíření ontologie i na ostatní kraje České republiky, bylo by nutné s tímto faktem počítat hlavně při propojování územních jednotek s působností příslušných orgánů státní správy.



Obrázek 3.3: Schéma soustavy územních prvků a územně evidenčních jednotek aplikované v RSO. Převzato z [48].

Na základě výše uvedených faktů jsem odvodila vlastní zjednodušené schéma vazeb mezi jednotlivými územními jednotkami, které je obecně použitelné pro celé území České republiky včetně Libereckého kraje, kde se nachází „neskladebný“ správní obvod ORP Turnov.

Inspirací pro mne také bylo schéma soustavy územních prvků a územně evidenčních jednotek, aplikované v Registru sčítacích obvodů Českého statistického úřadu [48], které je ovšem pro potřeby vytvoření geo-ontologie veřejné správy a samosprávy až moc podrobné. Mnou navržené schéma představuje jakýsi kompromis mezi velmi jednoduchým schématem vyplývajícím ze Zákona o územně-správním členění státu (schéma 3.2) a příliš podrobným schématem Českého statistického úřadu (schéma 3.3).



Obrázek 3.4: Skladebnost územního členění státu s ohledem na výjimky ze skladebnosti. Vlastní zpracování.

Důležitá je hlavně spodní část schématu, která řeší nejnižší územními prvky (katastrální území, městský obvod/městská část, základní sídelní jednotka) a účelové územní prvky (volební okrsek, sčítací obvod). Znázorňuje i skladebné vazby mezi nimi navzájem a obcí, ve které se nachází.

Identifikace podtříd třídy *Instituce*

Identifikace tříd spadajících do konceptu *Instituce* už nebyla tak snadná a jednoznačná jako v případě územních jednotek. Z legislativy a dostupných materiálů Ministerstva vnitra

[47] lze vyvodit jen část existujících orgánů státní a veřejné správy a samosprávy a hierarchických vztahů mezi nimi. Bylo proto třeba provést důkladné vlastní šetření z celé řady písemných pramenů, hlavně z přednášek pana doc. Průchy [46], které mi velmi pomohly pochopit základní principy fungování veřejné správy a její organizace. Velmi nápomocné mi také byly dvě webové stránky, a to *Portál veřejné správy* [3] a *statnisprava.cz*, kde jsem nalezla velké množství užitečných informací o veřejné správě a kompletní přehled všech existujících orgánů státní správy a samospráv. Dále jsem čerpala informace z webových stránek jednotlivých institucí, z nichž jsem přebírala stručný popis jejich agendy pro vyplnění anotační vlastnosti u každé třídy ontologie a také vymezení prostorové působnosti úřadů a jejich podřízených poboček či pracovišť.
Vzhledem k tomu, že základním principem ontologií je vyhledávání co nejobecnějších konceptů a vytváření taxonomické struktury mezi nimi, bylo třeba identifikovat několik obecných kategorií, které vytvoří nadtřídy pro konkrétní instituce. Vzniklo proto 6 obecných tříd:

- Státní správa
- Samospráva
- Ostatní veřejná správa
- Sbory
- Inspektoráty a kontrolní úřady
- Ostatní instituce

Do kategorie *Státní správa* spadají např. Finanční úřady, Matriky, Stavební úřady nebo Úřad práce. Mezi orgány *samosprávy* patří krajské úřady a úřady obcí různých typů (městský úřad, úřad obce s rozšířenou působností apod.) a to včetně všech svých odborů a oddělení. V kategorii *Ostatní veřejná správa* nalezneme další instituce státní správy jako Drážní správní úřad, Katastrální a Pozemkový úřad, Ředitelství silnic a dálnic nebo Úřad pro zastupování státu ve věcech majetkových. *Sbory* pak tvoří Armáda ČR, Celní správa ČR, Hasičský záchranný sbor, Policie ČR a Vězeňská služba ČR. Mezi *Inspektoráty a kontrolní úřady* se řadí na příklad Česká inspekce životního prostředí, Inspektoráty práce, Nejvyšší kontrolní úřad, Úřad pro ochranu osobních údajů nebo Ústav pro kontrolu léčiv. Dále pak zbývá celá řada institucí, které se nepodařilo přiřadit ani k jedné z těchto pěti skupin a spadají pak pod obecný koncept *Ostatní instituce*.
Tato poslední kategorie je dále členěná na další méně obecné třídy, a to podle zaměření daných institucí (v závorkách jsou uvedené vybrané příklady konkrétních institucí spadajících do dané nadtřídy):

[3] dostupné z: *https* : *//portal.gov.cz/*

- *Archiv* (např. Centrální depozitář cenných papírů, Národní archiv, Státní oblastní archivy)
- *Bezpečnost* (Národní úřad pro kybernetickou a informační bezpečnost, Bezpečnostní informační služba)
- *Cestovní ruch a místní rozvoj* (Centrum pro regionální rozvoj České republiky, Česká centrála cestovního ruchu)
- *Doprava* (Řízení letového provozu, Státní fond dopravní infrastruktury)
- *Finance a obchod* (Česká národní banka, Správa státních hmotných rezerv)
- *Kultura a památková péče* (Divadelní ústav, Národní památkový ústav)
- *Soudnictví a právo* (Exekutorská komora, Státní zastupitelství, Nejvyšší správní soud)
- *Školství* (Český institut pro akreditaci, Národní pedagogický institut)
- *Výzkumné a zkušební ústavy* (Akademie věd ČR, Česká rozvojová agentura, Ústav územního rozvoje)
- *Zdravotnictví* (Hygienická správa, Veterinární správa, Zdravotní pojišťovny)
- *Životní prostředí* (Agentura ochrany přírody a krajiny, Lesy ČR, Správa jeskyní ČR)

Kompletní seznam všech institucí veřejné správy, které jsou zahrnuté v geo-ontologii veřejné správy a samosprávy, je uvedený v Příloze 2 ve formě tabulek pro jednotlivé obecné kategorie. Uvědomuji si, že mnou navržené a použité rozčlenění orgánů veřejné správy do výše uvedených kategorií není jediným možným řešením a je věcí diskuze, zda by jiné řešení nebylo vhodnější. Mým cílem bylo vytvořit takové schéma, které je logicky uspořádané a srozumitelné i pro čtenáře, který není expertem na veřejnou správu, ale zároveň aby bylo dostatečně vypovídající o struktuře veřejné správy a plně odpovídalo skutečnosti. Nabízí se zde možnost v rámci další práce provést dotazníkové šetření, zda jsou takto navržené kategorie skutečně srozumitelné pro běžného uživatele či by mohlo být obdobným způsobem provedeno otestování více možných variant řešení.

3.2.2 Vytvoření taxonomie

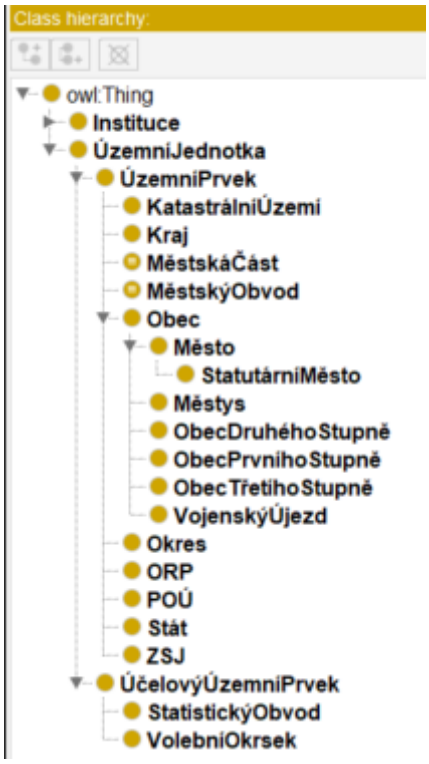
Vytvoření taxonomie ontologie spočívá v nalezení hierarchických vztahů mezi jednotlivými třídami, tedy vztahů nadřazenosti a podřazenosti, disjunkce či ekvivalence tříd. Tyto vztahy z velké míry vyplývají již z definování jednotlivých prvků ontologie, protože při vymezování nejobecnějších konceptů a méně obecných tříd se začíná formovat i základní struktura

vazeb mezi třídami. Základní nadtřídou pro všechny další třídy ontologie je vždy superclass *owl: Thing*. Jak již bylo zmíněno dříve, předmětem geo-ontologie jsou dva obecné disjunktní koncepty *Územní jednotka* a *Instituce*, které jsou přímými potomky třídy *owl: Thing*. Další struktura obou větví geo-ontologie je pak podrobně znázorněná na následujících obrázcích. Jedná se o ukázky přímo z editoru Protégé, který díky pluginu OWLViz [4] umožňuje vizualizovat taxonomii vytvořených tříd pomocí přehledných diagramů a následně je možné tato schémata přímo vyexportovat např. ve formátu .png. Nevýhodou pluginem vytvořených schémat ovšem je, že se s vykreslenými objekty nedá manipulovat a přesouvat je tak, aby byly lépe uspořádané podle logického významu a nikoliv tak, jak jsou automaticky řazeny. Další nevýhodou pluginu je jeho problematické užívání ve chvíli, kdy mají objekty v ontologii vyplněnou anotační vlastnost *rdfs: label*, proto jsou názvy tříd v diagramech uvedené bez mezer, protože takto jsou pojmenované v samotné ontologii. Výsledná geo-ontologie ale obsahuje všechny prvky řádně označené pomocí popisků *(rdfs: label)* a výše popsané vytvoření diagramů je z ní téměř nemožné.

Podtřídy třídy *Územní jednotka*

Třída *Územní jednotka* má dva potomky *Územní prvek* zahrnující různé samosprávní celky a *Účelový územní prvek* zastřešující uměle vytvořené územní celky za nějakým účelem. Jednotlivé třídy podřazené těmto dvěma kategoriím jsou uvedené na následujícím obrázku. Kromě vztahu subsumpce se zde objevuje i ekvivalence, a to mezi třídami *Městský obvod* a *Městská část*. Obě tato označení pro samosprávné části územně členěných statutárních měst jsou z hlediska Zákona o obcích rovnocenná a je na rozhodnutí zastupitelstva statutárního města, které použije. Samosprávné městské části nalezneme v Brně a v Opavě, městské obvody pak v Plzni, Ostravě, Pardubicích či Ústí nad Labem. Některá statutární města v České republice (České Budějovice, Děčín, Mladá Boleslav nebo Olomouc) se sice dělí na městské části, ty ale nejsou samosprávné. Další statutární města členění na městské části nemají. Specifická situace je pak v Praze, kde je územní členění na samosprávné jednotky prováděno dle zákona č. 131/2000 Sb., o hlavním městě Praze, a obecně závazné vyhlášky č. 55/2000 Sb., Statut hl. m. Prahy, a je velmi složité.

[4] Plugin OWLViz není ve výchozí verzi desktopové aplikace Protégé a je třeba ho dodatečně doinstalovat. Pro jeho správnou funkci vykreslování schémat je nutné mít také nainstalovaný grafický opensource softwarový balík *Graphviz* v příslušné verzi.

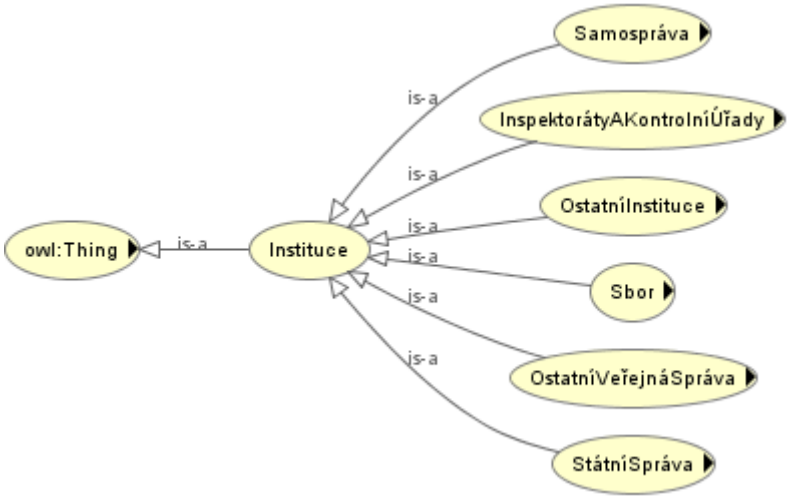


Obrázek 3.5: Přehled podtříd třídy *Územní jednotka*. Vlastní zpracování.

Obrázek 3.6: Taxonomie podtříd třídy *Územní jednotka*. Vlastní zpracování.

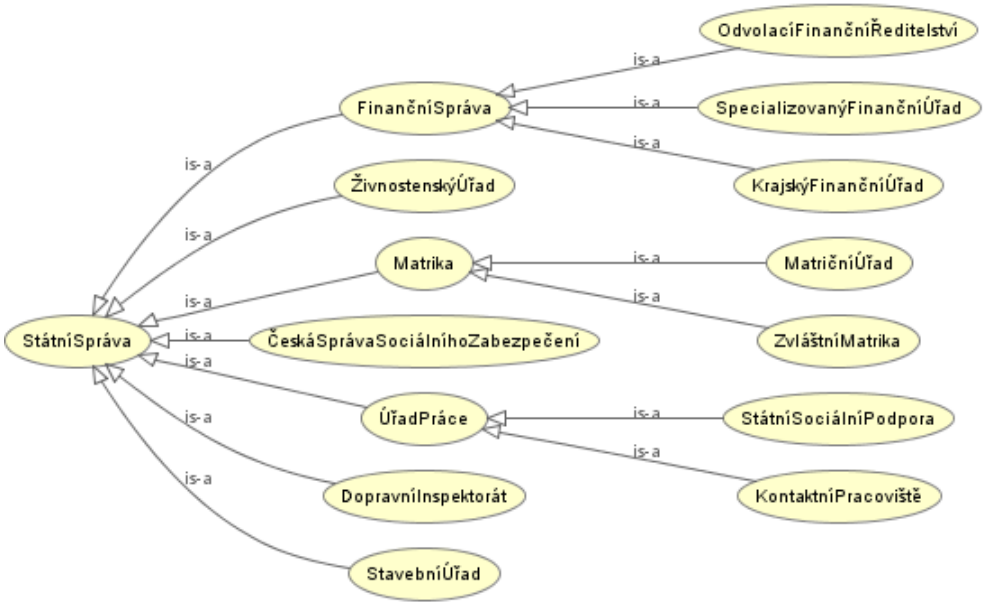
Podtřídy třídy Instituce

Mezi podtřídami třídy *Instituce* se vyskytují hierarchické vztahy podřazenosti a nadřazenosti a zároveň o všech třídách platí, že jsou disjunktní, tedy že nemají žádný průnik, resp. by jejich průnikem byla prázdná množina. Výsledná taxonomie tříd vznikala již při identifikaci jednotlivých tříd.



Obrázek 3.7: Základní rozdělení konceptu *Instituce* na méně obecné kategorie. Vlastní zpracování.

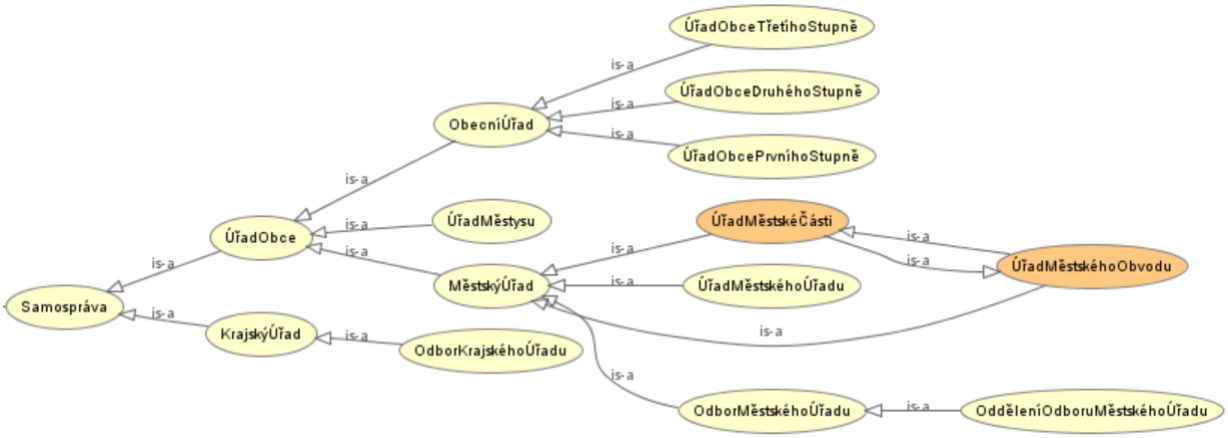
Na následujících obrázcích je pak ukázána detailní taxonomie v těchto šesti kategoriích.



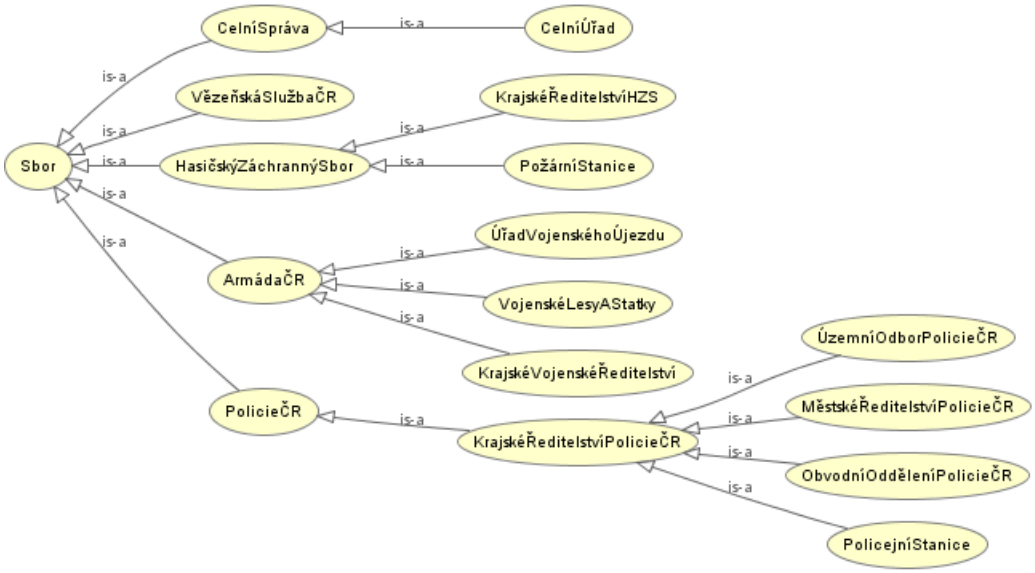
Obrázek 3.8: Taxonomie v kategorii *Státní správa*. Vlastní zpracování.

Najdeme i zde jeden vztah ekvivalence mezi třídou *Úřad městské části* a *Úřad městského*

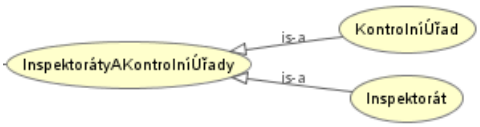
obvodu, které působí v uvedených samosprávných částech statutárních měst.



Obrázek 3.9: Taxonomie v kategorii *Samospráva*. Vlastní zpracování.



Obrázek 3.10: Taxonomie v kategorii *Sbory*. Vlastní zpracování.



Obrázek 3.11: Taxonomie v kategorii *Inspektoráty a kontrolní úřady*. Vlastní zpracování.

Obrázek 3.12: Taxonomie v kategorii *Ostatní veřejná správa*. Vlastní zpracování.

Obrázek 3.13: Taxonomie v kategorii *Ostatní instituce*. Vlastní zpracování.

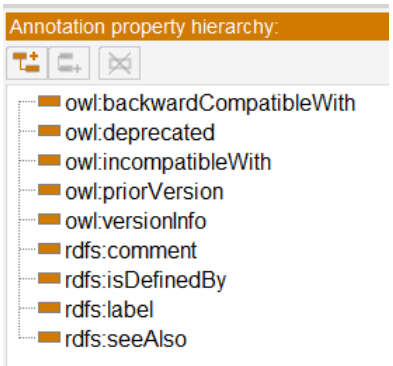
3.2.3 Podrobná deklarace tříd a vlastností

Každý objekt v ontologii je jednoznačně pojmenovaný a identifikovatelný pomocí svého IRI (=Internationalized Resource Identifier) a vazbami na další prvky ontologie - v editoru Protégé se obecně označují jako *entity*. Každá třída musí být minimálně potomkem (*Subclass of)* třídy *owl: Thing* a může být předkem libovolného množství dalších tříd. Třídy mohou mít své instance v podobě konkrétních individuálů, tedy prvků s datatypovými vlastnostmi nabývajících určitých hodnot. Další vazby mezi třídami jsou definované pomocí objektových vlastností, které kromě názvu a IRI mohou mít také svoji taxonomii. Všechny objektové vlastnosti jsou potomky *owl: topObjectProperty* a můžou mít svoje podřazené *sub-properties*. Podobně i datatypové vlastnosti jsou potomky *owl: topDataProperty* a také mohou mít *sub-properties*. Anotační vlastnosti jsou v tomto ohledu speciální, protože některé jsou již předem definované a nelze je měnit či jim přidávat další potomky. Deklarace jednotlivých typů *entit* je popsána v následujícím textu tak, jak byla prováděná během tvorby Geo-ontologie veřejné správy a samosprávy.

Deklarace anotačních vlastností

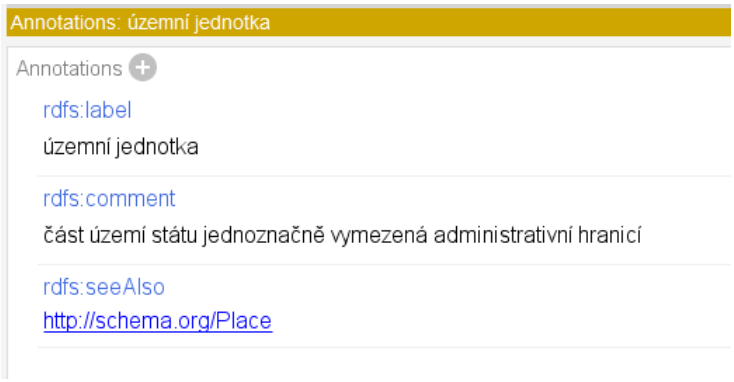
Deklarace anotačních vlastností je oproti ostatním objektům ontologie specifická díky tomu, že jsou již v ontologii vytvořené a to buď pomocí jazyka OWL 2 nebo RDFS, ale je možné přidávat i další vlastní anotační vlastnosti, což jsem ale pro potřeby vytvářené ontologie nevyužila. Anotační vlastnosti jsou v editoru označené oranžovým obdélníčkem. Již existující anotační vlastnosti jsou uvedené níže na obrázku a z nich byly v Geo-ontologii veřejné správy a samosprávy využity následující:

- *rdfs: label* - popisek, který se v ontologii používá pro označení objektu místo jeho IRI
- *rdfs: comment* - komentář k objektu, pro potřeby geo-ontologie slouží pro popis objektu
- *rdfs: seeAlso* - odkaz na obdobný objekt v jiné existující ontologii



Obrázek 3.14: Předem definované anotační vlastnosti. Vlastní zpracování.

U anotačních vlastností je možné určit jejich definiční obor a obor hodnot, ale tuto možnost jsem nevyužila. Anotační vlastnosti je možné ale i žádoucí přidávat ke každému objektu v ontologii, protože vytváří metadata k ontologii a zároveň zvyšují její srozumitelnost pro široké spektrum uživatelů. Stejně tak je možné přidávat anotace i anotačním vlastnostnem.



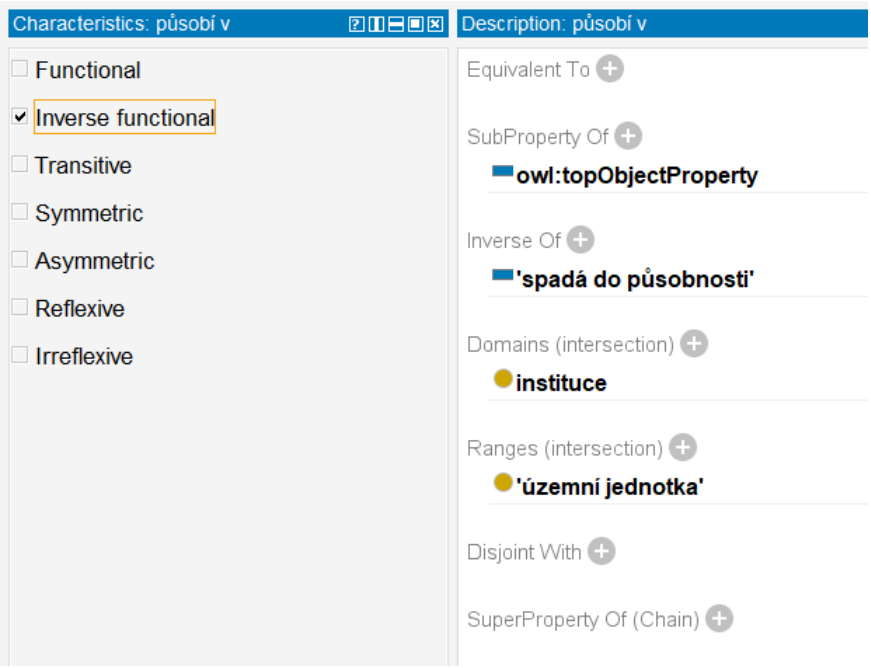
Obrázek 3.15: Ukázka anotačních vlastností třídy *Územní jednotka*. Vlastní zpracování.

Deklarace objektových vlastností

Nové objektové třídy je možné v editoru vytvářet v záložce *Object Properties* v pohledu *Object Property Hierarchy*, kde se nová třída přidá buď pomocí tlačítka v horní části pohledu, přes vyskakovací nabídku po kliknutí pravou myší do okna nebo pomocí klávesové zkratky. Tyto možnosti jsou obecně platné i pro další typy objektů ontologie. Nejprve je nutné zadat IRI vytvářené vlastnosti a následně ji blíže specifikovat v pohledu *Object Property Description*. V tomto pohledu jsou k dispozici tyto omezení vybrané vlastnosti:

- *Equivalent To* - název jiné objektové vlastnosti, se kterou je ekvivalentní
- *SubProperty Of* - název vlastnosti, která je rodičem vybrané vlastnosti
- *Inverse Of* - výčet vlastností, které jsou k dané vlastnosti inverzní
- *Domains (intersection)* - výčet tříd, které tvoří doménu dané vlastnosti
- *Ranges (intersection)* - výčet tříd, které jsou oborem hodnot dané vlastnosti
- *Disjoint With* - výčet vlastností se kterými je vybraná vlastnost disjunktní
- *SuperProperty Of (Chain)* - výčet vlastností, pro které je vybraná vlastnost nadtřídou

Objektové vlastnosti jsou v editoru Protégé vyznačené modrým obdélníčkem. Kromě výše uvedených omezení objektových vlastností je zde možnost zadat charakteristiku (*Characteristics*) vlastnosti z nabízených možností: *Functional, Inverse functional, Transitive, Symmetric, Asymmetric, Reflexive, Irreflexive.*



Obrázek 3.16: Ukázka deklarace objektové vlastnosti *působí v*. Vlastní zpracování.

V Geo-ontologii veřejné správy jsou pouze čtyři objektové vlastnosti vymezující vztahy mezi třídami. Jde o inverzní vlastnosti 'působí v' a 'spadá do působnosti', které určují vzájemný vztah mezi nějakou institucí a územní jednotkou. Vztahy mezi územními jednotkami navzájem pak určuje druhá dvojice objektových vlastností 'je částí' a 'skládá se z' popisující jejich vzájemnou skladebnost.

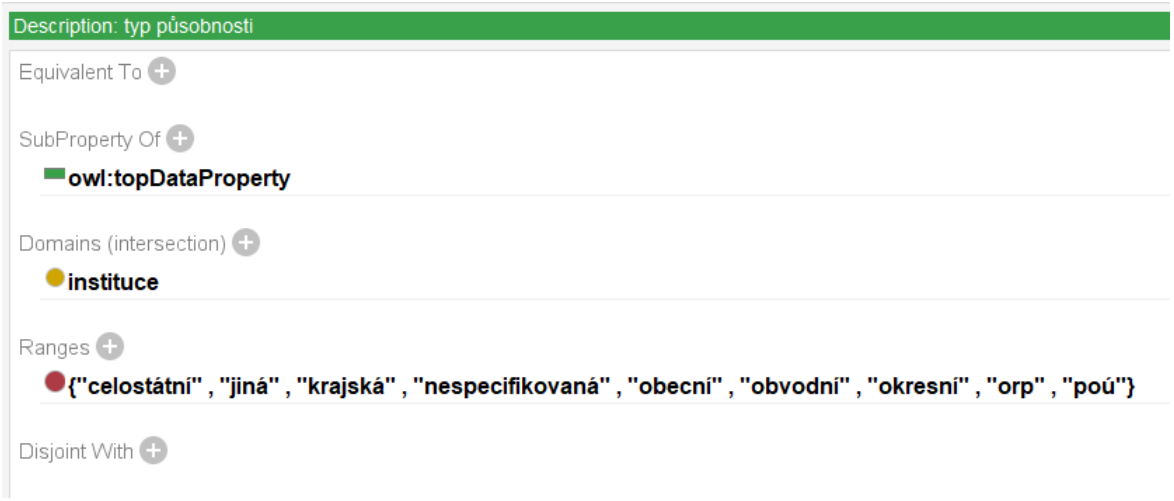
Deklarace datatypových vlastností

Datatypové vlastnosti se vytváří a upravují v záložce *Data Properties* obdobně jako objektové vlastnosti. V editoru jsou odlišené pomocí zeleného obdélníčku. Pro každou datatypovou vlastnost lze určit:

- *Equivalent To* - název jiné datatypové vlastnosti, se kterou je ekvivalentní
- *SubProperty Of* - název vlastnosti, která je rodičem vybrané vlastnosti
- *Domains (intersection)* - výčet tříd, jejichž instance mohou mít tuto datatypovou vlastnost
- *Ranges* - obor hodnot, kterých datatypová vlastnost může nabývat
- *Disjoint With* - výčet vlastností se kterými je vybraná vlastnost disjunktní

Pro určení oboru hodnot datatypové vlastnosti se použije buď datový typ hodnot (např. *xsd: string, xsd: integer, xsd: boolean, owl: rational*) nebo výčet možných hodnot. V ontologii

je k dispozici celá řada definovaných datových typů pomocí jazyků OWL, RDF, RDFS a XSD a v editoru jsou odlišené červeným puntíkem. Je možné přidávat a definovat i vlastní datové typy dle potřeby. V Geo-ontologii veřejné správy a samosprávy jsem použila pro většinu datatypových vlastností již předdefinované standardní datové typy a pro jednu vlastnost jsem určila obor hodnot pomocí výčtu konkrétních hodnot. V ontologii se objevuje celkem pět datatypových vlastností: *adresa, název instituce, typ působnosti, ID a číslo jednotky.*



Obrázek 3.17: Ukázka deklarace datatypové vlastnosti *typ působnosti*. Vlastní zpracování.

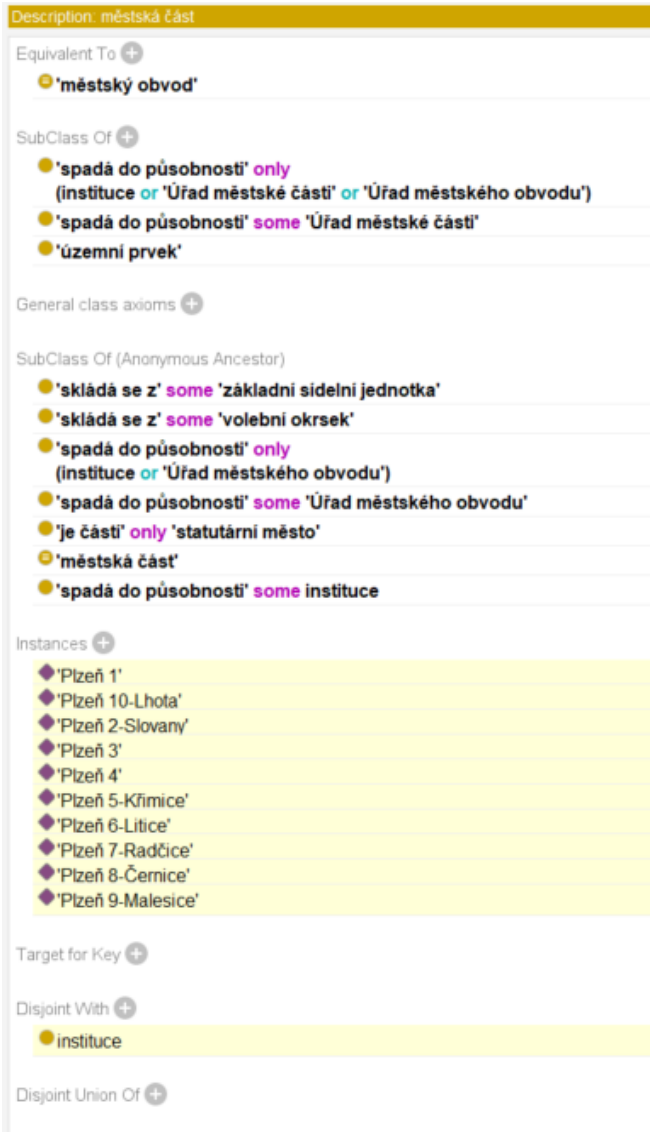
Deklarace tříd

Deklarace nové třídy začíná výběrem jejího předka a následně zadáním jejího IRI. Žádnou třídu není možné deklarovat bez uvedení jejího předka, ale je možné později vytvořené vztahy mezi třídami měnit. Je možné zadat i více nových tříd najednou pomocí okna *Enter hierarchy*, do kterého se vypisují IRI nově vkládaných tříd a pomocí naformátování textu v okně se vytvoří i jejich hierarchie. Takto vytvořené nové třídy jsou primitivními třídami a v editoru jsou označené žlutým puntíkem. V okně *Description* je možné specifikovat následující taxonomické a objektové vlastnosti dané třídy:

- *Equivalent To* - název jiné třídy, se kterou je vybraná třída ekvivalentní (třídy v tomto vztahu již nejsou primitivními třídami, ale třídami definovanými a jejich označení se v editoru změní na žlutý puntík se třemi vodorovnými čárkami)
- *SubClass Of* - výčet tříd, kterých je vybraná třída potomkem
- *General class axioms* - výčet obecných axiomů, které zmiňují vybranou třídu (v geo-ontologii se u žádné třídy nevyskytují)
- *SubClass Of (Anonymous Ancestor)* - výčet zděděných vlastností *SubClass Of* po předcích (tyto vlastnosti jsou typicky odvozeny při reasoningu)

- *Instances* - výčet instancí vybrané třídy (tj. soupis všech individuálů, které jsou instancemi dané třídy)
- *Target for Key* - výpis objektových a datatypových vlastností, které se týkají instancí dané třídy
- *Disjoint With* - výčet tříd, se kterými je vybraná třída disjunktní
- *Disjoint Union Of* - specifikuje že vybraná třída je hlavní třídou v *DisjointUnion* axiomu

Při specifikaci objektových vlastností, kterých je daná třída potomkem, je možné tyto vlastnosti blíže specifikovat použitím klíčových slov omezujících počet (*some, only, value, min, max, exactly*), kolika instancí těchto tříd se vlastnost týká, a dále je spojovat pomocí logických spojek (*and, or, not*). Například axiom „kraj *je částí* **some** stát **and only** stát“ říká, že nějaká instance třídy kraj je částí nějaké a pouze jedné instance třídy stát a nazývá se axiom uzávěru (*Closure Axiom*). Všechny axiomy, které se týkají specifikace tříd, se nazývají *class experessions* a zapisují se pomocí Manchaster syntaxe jazyka OWL 2.
Při deklaraci inverzních vztahů mezi třídami není nutné vytvářet obě varianty daného vztahu, ale je dostačující zadat vztah pouze „jedním směrem“ a opačný vztah se odvodí díky reasoningu. Reasonery dále odvozují pro třídy i vlastnosti ekvivalence a disjunkce. V editoru Protégé se odvozené vlastnosti vyznačují světle žlutým podbarvením a jsou viditelné jen při aktivování nějakého reasoneru. Stejně tak není nutné pro každou třídu definovat vlastnosti zděděné po předcích, protože jsou tyto vlastnosti samy automaticky doplňované na základě již existujících axiomů.
V Geo-ontologii veřejné správy je celkem 167 tříd a jejich identifikací jsem se zabývala již v předchozích částech této kapitoly. Všechny třídy ontologie jsou také uvedené na schématech v sekci **3.2.2 Vytvoření taxonomie** a v dokumentaci, která je přílohou bakalářské práce. Z tohoto důvodu zde neuvádím kompletní výčet všech tříd.



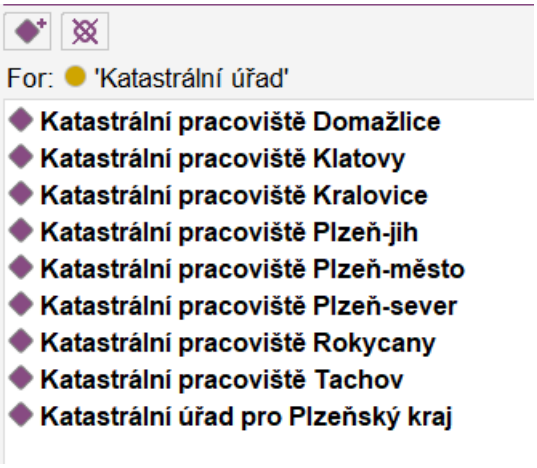
Obrázek 3.18: Ukázka deklarace třídy *Městská část* po reasoningu. Vlastní zpracování.

Deklarace individuálů

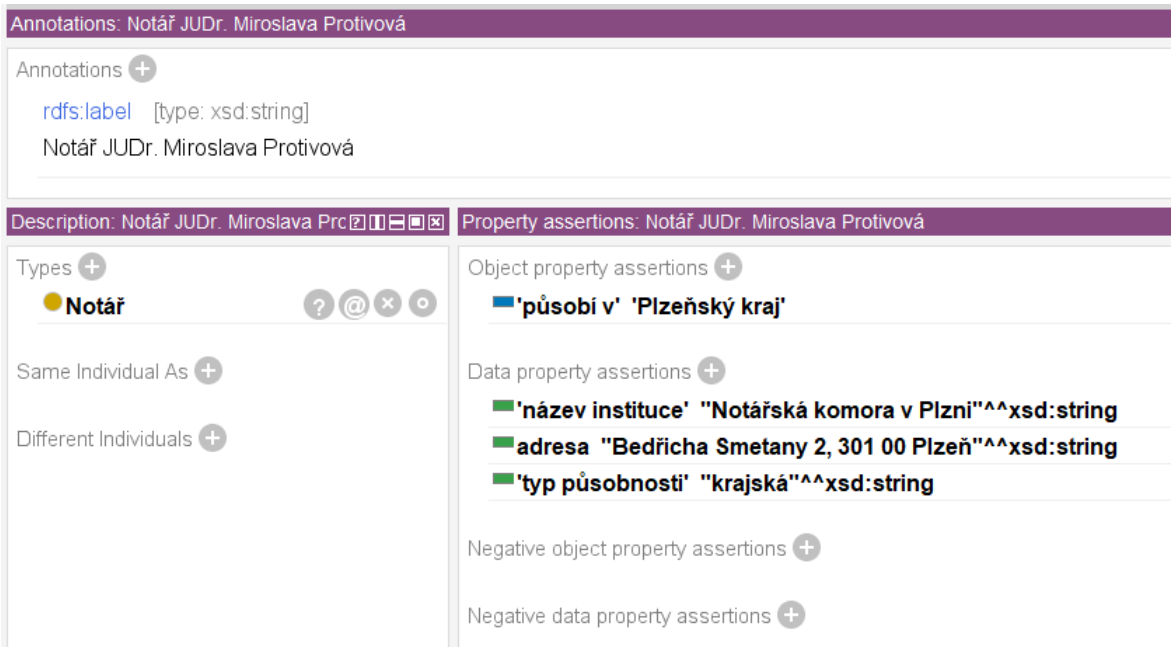
Každý individuál musí být instancí nějaké třídy ontologie a od ní pak přebírá veškeré její vlastnosti. Pro individuály je typické, že mají udané konkrétní hodnoty pro příslušné datatypové vlastnosti. V editoru je pro práci s individuály určená záložka *Individuals by class*, kde jsou roztříděné podle příslušnosti k nějaké třídě, a nebo záložka *Individuals*, kde jsou všechny individuály pohromadě řazené podle abecedy. Od ostatních entit jsou odlišené pomocí fialového kosočtverce. Pro každý individuál se může určit:

- *Types* - název tříd, jejichž je individuál instancí
- *Same Individual As* - výčet shodných individuálů

- *Different Individuals* - výčet odlišných individuálů (nevyužívá se)
- *Object Property Assertions* - výčet objektových vlastností, které se týkají daného individuálu
- *Data Property Assertions* - výčet datatypových vlastností, které se týkají daného individuálu
- *Negative Object Property Assertions*
- *Negative Data Property Assertions*



Obrázek 3.19: Ukázka všech instancí třídy *Katastrální úřad*. Vlastní zpracování.



Obrázek 3.20: Ukázka deklarace instance třídy *Notář*. Vlastní zpracování.

Poznámka k podrobné deklaraci tříd a vlastností

Výše popisovaný postup deklarace tříd a vlastností ontologie je pouze jednou z možností, jak ontologii vytvořit. Obecně lze říci, že existují dva přístupy k vytvoření ontologie:

1. plně využít grafické uživatelské prostředí editoru Protégé a nové axiomy do ontologie přidávat pomocí dostupných tlačítek či rozbalovacích nabídek a využívat již definované objekty

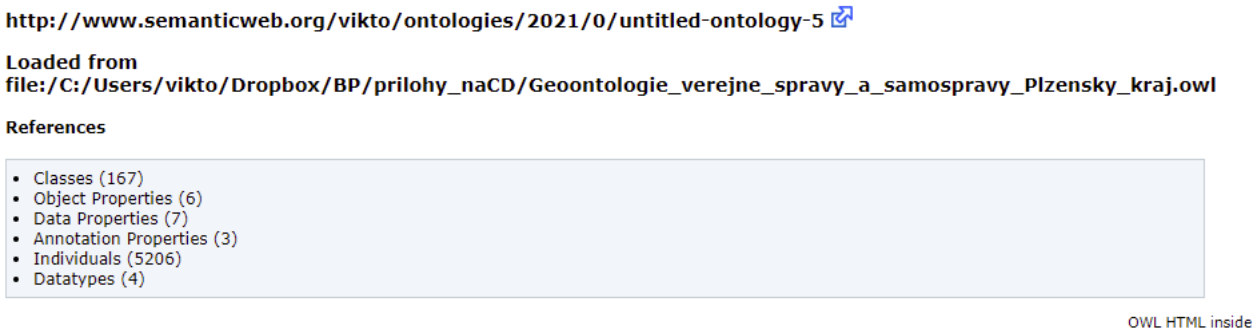
2. kombinovat možnosti uživatelského prostředí editoru s vlastním psaním axiomů v jazyce OWL 2, převážně s využitím Manchaster syntaxe, a vytvářet tak nové objekty ontologie či je následně opravovat rychleji a neomezovat se pouze na předdefinované objekty.

První zmiňovaný postup je poměrně snadný a díky intuitivnímu prostředí editoru a možnosti nápovědy a oficiální dokumentace [49] je vhodný pro začátečníka či nenáročného uživatele. Při prvotním seznamování s editorem a vytváření prvních testovacích verzí ontologie jsem pracovala hlavně tímto způsobem a potupně jsem si osvojila řadu klávesových zkratek, které práci v editoru značně urychlují.
Druhý přístup se mi osvědčil hlavně pro vytvoření finální velmi rozsáhlé verze ontologie, kdy bylo nutné do ontologie přidat obrovské množství axiomů (v řádu desítek tisíc) a to by nebylo možné provést prvním způsobem. Podrobněji se hromadnému plnění ontologie daty věnuje další kapitola.

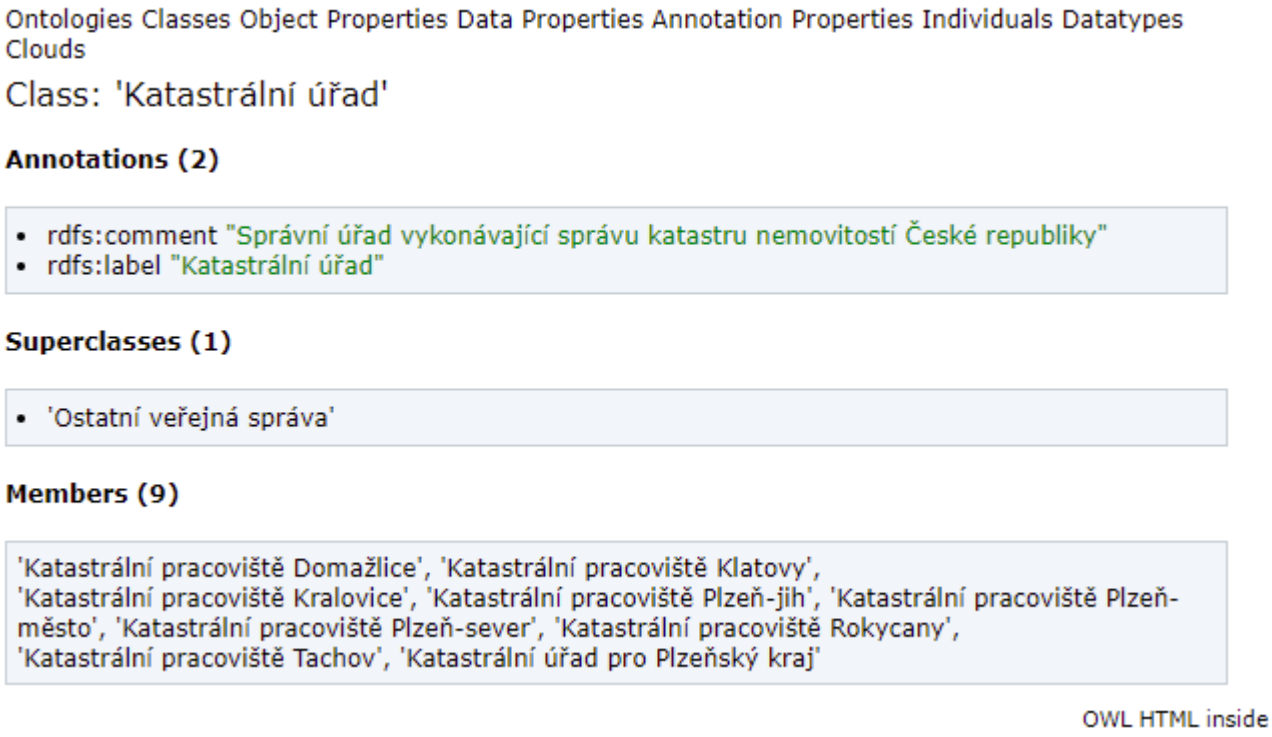
3.3 Dokumentace ontologie

Dokumentaci ontologie lze získat přímo z editoru Protégé pomocí pluginu OWLDoc. Tento plugin na základě všech axiomů vytvoří balík statických HTML stránek, které je možné exportovat a následně procházet pomocí webového prohlížeče. Pro interaktivní procházení HTML stránek dokumentace je součástí balíku i javascriptový kód a soubor s kaskádovými styly. Dále jsou v balíku jednotlivé HTML soubory pro všechny entity ontologie roztříděné do složek podle jejich typu (*classes, annotation properties, data properties, object properties, individuals*).
Procházení dokumentace v prohlížeči je velice snadné a intuitivní, neboť všechny objekty jsou vzájemně propojené pomocí odkazů a celou dokumentací lze procházet proklikáváním na požadované prvky, případně se vrátit na výchozí stránku dokumentace či na přehled všech entit podle jejich typu. Jediným nedostatkem je, že takto vyexportovaný balík HTML stránek může nabývat v případě rozsáhlé ontologie s velkým množstvím prvků obrovských velikostí v řádu desítek Gigabytů. Konkrétně vyexportovaná dokumentace kompletní geo-ontologie pro Plzeňský kraj, která obsahuje 5393 prvků, má 54 GB a takto velký soubor s sebou přináší

řadu problémů s jeho sdílením a přenášením dalším uživatelům. Pak je nutné využít nějaký komprimační nástroj a dokumentaci přenášet jako již zkomprimovaný archiv. Na následujících obrázcích jsou ukázky z takto získané dokumentace, která je jako celek součástí přílohy.

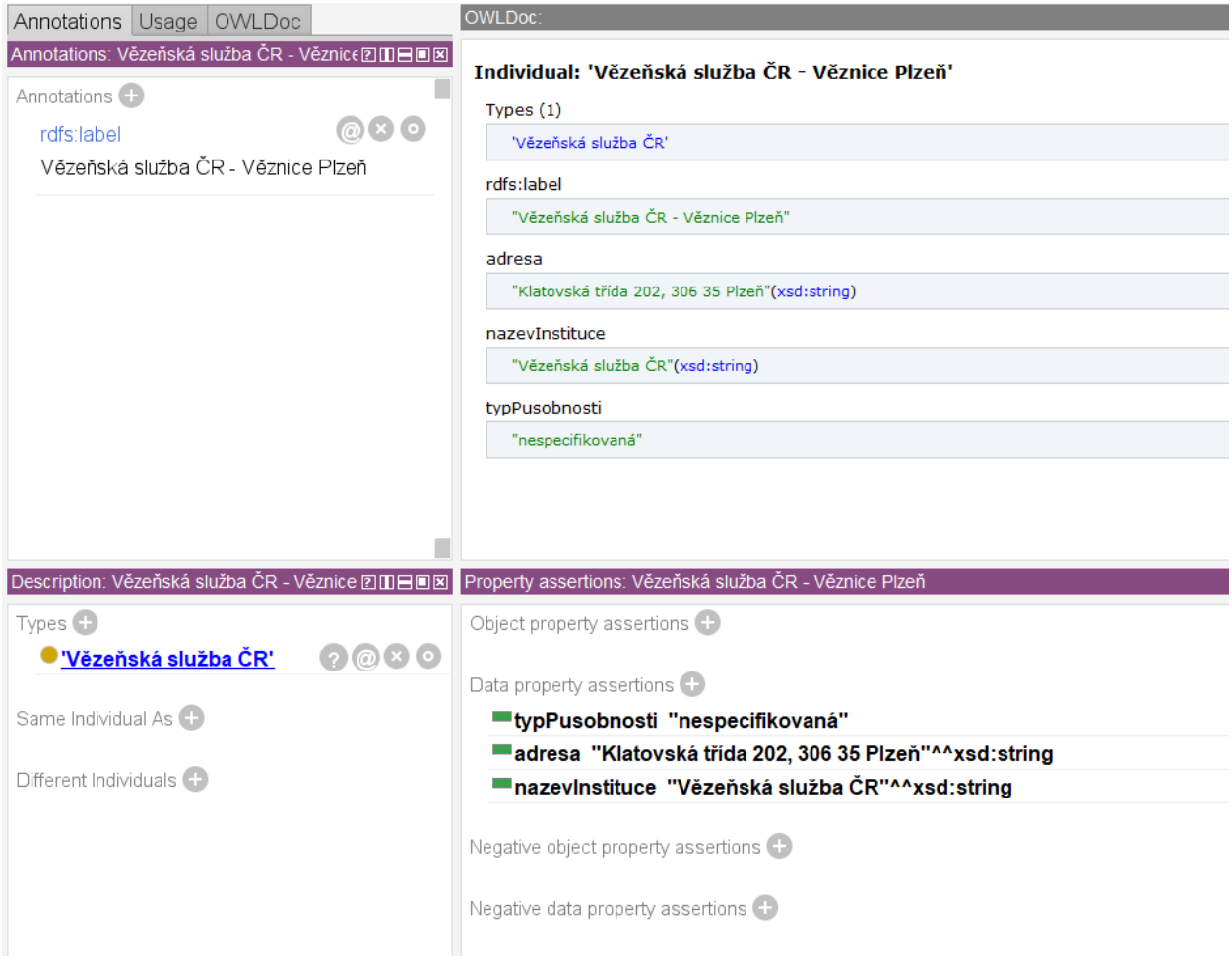


Obrázek 3.21: Úvodní webová stránka dokumentace ontologie. Vlastní zpracování.



Obrázek 3.22: Ukázka dokumentace třídy Katastrální úřad. Vlastní zpracování.

Kromě exportu kompletní dokumentace plugin OWLDoc umožňuje také přidat do editoru nové pohledové okno, ve kterém se po vybrání libovolného objektu ontologie zobrazí jeho dokumentace.



Obrázek 3.23: Pohledové okno pluginu OWLDoc s dokumentací vybraného objektu. Vlastní zpracování.

4 Naplnění ontologie daty pro Plzeňský kraj

4.1 Obecné informace o Plzeňském kraji

Plzeňský kraj se nachází v jihozápadní části České republiky a stejně jako ostatní kraje byl zřízen k 1.1.2000 na základě Ústavního zákona o vytvoření vyšších územních samosprávných celků a o změně ústavního zákona České národní rady č. 1/1993 Sb., Ústava České republiky. Dále je pro existenci kraje závazný Zákon o krajích č. 129/2000 Sb.(krajské zřízení), který upravuje postavení kraje a jeho samosprávných orgánů.

S necelými 600 000 obyvateli je Plzeňský kraj až osmým v pořadí co do počtu obyvatel, ovšem rozlohou je třetím největším krajem. Hustota obyvatelstva (76 $obyv./km^2$) je proto třetí nejnižší v České republice. Co se týká počtu obcí, tak se svými 501 obcí je na pátém místě a s počtem měst (z 501 obcí jich má statut města 57) je na třetí příčce. Do roku 2016 se na území Plzeňského a Středočeského kraje nacházel vojenský újezd Brdy, který se skládal z 27 katastrálních území na ploše okresů Rokycany (7 k.ú.), Plzeň-jih (3 k.ú.), Beroun (6 k.ú.) a Příbram (11 k.ú.). Tato katastrální území se připojila ke katastrům „civilních“ obcí a v současné době se tedy v Plzeňském kraji žádný vojenský výcvikový prostor nenachází.

V Plzeňském kraji je zřetelné dominantní postavení krajského města Plzně, které je centrem celého regionu. Kromě institucí s krajskou či okresní působností a samosprávy města zde sídlí i řada orgánů státní a veřejné správy s působností na okresy či ORP, které se nachází v bezprostředním sousedství města. Za všechny můžu jmenovat například pobočky Úřadu práce a Správy sociálního zabezpečení pro okresy Plzeň-jih a Plzeň-sever, které mají sídlo v Plzni spolu s pobočkou pro okres Plzeň-město. Stejně tak Městský úřad Nýřany, který je úřadem obce s rozšířenou působností pro 54 obcí, má tři pobočky, z nichž dvě sídlí mimo území SO ORP v Plzni. Na tomto příkladu je ukázáno, že je Geo-ontologie veřejné správy a samosprávy navržená tak, aby dokázala pokrýt i tyto neobvyklé či výjimečné situace. Také fakt, že je Plzeň statutárním městem, jehož samospráva je prováděna jinak než v „obyčejných“ městech, ukazuje univerzálnost použití navržené struktury.

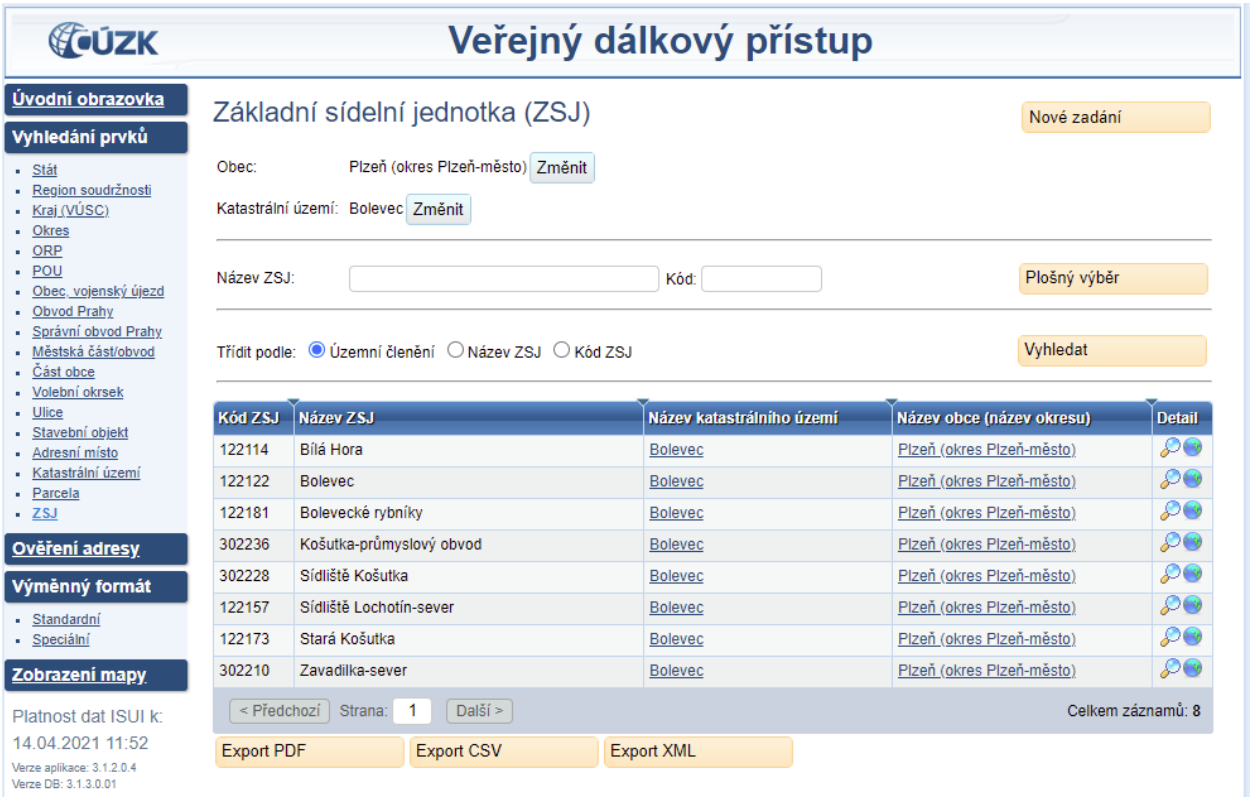
Jediný koncept, který není možné naplnit konkrétními daty z Plzeňského kraje, je právě vojenský újezd a k němu příslušné vojenské správní úřady. Nicméně v geo-ontologii je počítáno i s těmito třídami reprezentující vojenské újezdy a jejich správu, aby mohla být použitelná i pro jiné kraje. Všechny ostatní koncepty se v kraji vyskytují a je možné pomocí konkrétních dat ukázat funkčnost celého návrhu geo-ontologie.

OKRES	ORP	POÚ	Obce	ZSJ	KÚ	VO	SO
Domažlice	Domažlice	Domažlice	30	84	53	51	149
		Kdyně	15	44	34	24	60
		Poběžovice	13	56	52	16	67
	Horšovský Týn	Horšovský Týn	10	59	48	21	79
		Staňkov	8	17	15	12	26
Klatovy	Horažďovice	Horažďovice	20	74	55	50	92
	Klatovy	Klatovy	31	198	146	90	304
		Nýrsko	5	36	33	15	57
		Plánice	7	24	21	15	28
	Sušice	Kašperské Hory	5	31	26	5	39
		Sušice	26	181	134	56	240
Plzeň-jih	Blovice	Blovice	13	42	28	24	57
		Spálené Poříčí	6	25	21	13	31
	Nepomuk	Nepomuk	26	67	63	56	87
	Přeštice	Přeštice	30	73	50	56	104
	Stod	Dobřany	5	10	8	10	33
		Holýšov	9	13	11	14	24
		Stod	10	23	19	18	45
Plzeň-město	Plzeň	Plzeň	4	122	26	189	753
		Starý Plzenec	11	22	17	19	49
Plzeň-sever	Kralovice	Kralovice	25	57	49	31	77
		Manětín	6	43	40	8	49
		Plasy	13	27	23	26	46
	Nýřany	Město Touškov	15	40	34	16	53
		Nýřany	13	36	20	29	87
		Třemošná	15	28	22	26	70
		Všeruby	11	36	34	11	40
Rokycany	Rokycany	Radnice	17	37	29	25	46
		Rokycany	34	79	47	54	159
		Zbiroh	17	33	23	22	42
Tachov	Stříbro	Bezdružice	5	25	22	8	34
		Stříbro	19	72	59	26	113
	Tachov	Bor	7	63	54	12	86
		Planá	8	46	42	13	76
		Tachov	12	58	38	26	116
Celkem	**15**	**35**	**501**	**1881**	**1396**	**1087**	**3418**

Obrázek 4.1: Přehled administrativního členění Plzeňského kraje. Vlastní zpracování.

4.2 Použitá data a jejich zdroje

Data pro naplnění geo-ontologie byla sbírána již v rámci semestrálních projektů, neboť se jedná o velmi rozsáhlý dataset z několika různých zdrojů a velkou část dat bylo nutné zpracovat manuálně po jednotlivých položkách. S tím také souvisí harmonizace dat do jednotné podoby ve formě tabulky ve formátu .xlsx. Pro použití získaných dat v geo-ontologii bylo nutné je aktualizovat, aby byla v souladu se skutečný stavem a legislativou platnou v době finalizace bakalářské práce, tedy v dubnu 2021. Stejně tak bylo třeba tabulky upravit do takové podoby, která je vhodná pro použití pluginu *Cellfie* (více viz další sekce).
Data týkající se územního členění jsem čerpala z Registru územní identifikace, adres a nemovitostí (dále RÚIAN) spravovaného Českým úřadem zeměměřickým a katastrálním [1]. Do registru je možné nahlížet pomocí Veřejného dálkového přístupu a stahovat z databáze data ve výměnném formátu (.pdf, .csv, .xml) podle zadaných kritérií. Pro účely této práce bylo nejvhodnější použít formát .csv, neboť je možné tato data zobrazit v MS Excel v tabulce, uložit je jako .xlsx soubor a dále s nimi pracovat jako s běžnou „excelovskou“ tabulkou.



Obrázek 4.2: Ukázka Veřejného dálkového přístupu do RÚIANu. Vlastní zpracování.

V RÚIANu jsou k dispozici údaje o základních územních prvcích, územně evidenčních

[1] Dostupné z: *https* : *//vdp.cuzk.cz/vdp/ruian/vymennyformat/vyhledej*

jednotkách a některých účelových územních prvcích (v současné době jsou zde k dispozici údaje o volebních okrscích). Stahování dat je velmi snadné, bohužel pro menší územní jednotky (katastrální území, základní sídelní jednotky, volební okrsky, a pod.) je možné údaje stahovat pouze pro jednotlivé obce, což je ve chvíli, kdy uživatel potřebuje údaje pro celý kraj, poměrně zdlouhavé.

Údaje pro statistické obvody byly získány z Registrů sčítacích obvodů a budov (dále RSO) spravovaného Českým statistickým úřadem, kde je možné data z databáze zobrazovat a stahovat po vybraných územních celcích ve formátu .dbf [2]. MS Excel umí zpracovat i tento datový formát a následně jej uložit jako .xslx soubor. Před samotným stažením dat z RSO je možné si v parametrech výběru nastavit připojení údajů z dalšího číselníku registru. Vzhledem k tomu, že jsou statistické obvody vázané na základní sídelní jednotky, připojila jsem si jako další sloupec kódy příslušných ZSJ a díky tomu pak bylo možné takto získaná data provázat přes kód ZSJ s daty, které jsem již měla k dispozici z RÚIANu. Statistické obvody mají kromě unikátního kódu i pořadové číslo v rámci okresu, proto jsem data stahovala po jednotlivých okresech, aby jejich úprava a propojení s daty z RÚIANu bylo přehlednější a až na závěr jsem je uložila do jedné tabulky. Výsledný sešit pro územní členění obsahuje tři listy: ZSJ, VO a SO s příslušnými územními jednotkami.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Kód ZSJ	ZSJ	Kód k.ú.	Katastrální území	Kód Obce	Název Obce	Typ obce	Kód POÚ	Název POÚ	Kód ORP	Název ORP	Kód Okresu	Název Okresu	Kód Kraje	Název Kraje
2	710	zsj Babylon	600717	kú Babylon	553433	Babylon	obec prvního stupně	1015	poú Domažlice	485	orp Domažlice	3401	okres Domažlice	43	Plzeňský kraj
3	21369	zsj Česká Kubice	621366	kú Česká Kubice	553549	Česká Kubice	obec prvního stupně	1015	poú Domažlice	485	orp Domažlice	3401	okres Domažlice	43	Plzeňský kraj
4	21377	zsj Nová Kubice	752746	kú Starý Spálenec	553549	Česká Kubice	obec prvního stupně	1015	poú Domažlice	485	orp Domažlice	3401	okres Domažlice	43	Plzeňský kraj
5	34550	zsj Dolní Folmava	634557	kú Dolní Folmava	553549	Česká Kubice	obec prvního stupně	1015	poú Domažlice	485	orp Domažlice	3401	okres Domažlice	43	Plzeňský kraj
6	34568	zsj Horní Folmava	634565	kú Horní Folmava	553549	Česká Kubice	obec prvního stupně	1015	poú Domažlice	485	orp Domažlice	3401	okres Domažlice	43	Plzeňský kraj
7	152722	zsj Nový Spálenec	752720	kú Nový Spálenec	553549	Česká Kubice	obec prvního stupně	1015	poú Domažlice	485	orp Domažlice	3401	okres Domažlice	43	Plzeňský kraj
8	152731	zsj Spáleneček	752746	kú Starý Spálenec	553549	Česká Kubice	obec prvního stupně	1015	poú Domažlice	485	orp Domažlice	3401	okres Domažlice	43	Plzeňský kraj
9	152749	zsj Starý Spálenec	752746	kú Starý Spálenec	553549	Česká Kubice	obec prvního stupně	1015	poú Domažlice	485	orp Domažlice	3401	okres Domažlice	43	Plzeňský kraj
10	26085	zsj Díly	626082	kú Díly	553557	Díly	obec prvního stupně	1015	poú Domažlice	485	orp Domažlice	3401	okres Domažlice	43	Plzeňský kraj

Obrázek 4.3: Ukázka listu ZSJ z výsledné tabulky vhodné pro import dat do ontologie. Vlastní zpracování.

Pro získání údajů o jednotlivých institucích veřejné správy a samosprávy bohužel není k dispozici žádný registr či databáze, ze kterého by bylo možné tyto údaje snadno získat, jako tomu bylo u územních jednotek. Údaje jsou k dispozici na webu *statnisprava.cz*, ale pouze k prohlížení a hromadné stahování údajů není možné. Tato část získávání údajů byla jedna z časově nejnáročnějších, protože jsem nenašla jinou možnost, než údaje pro jednotlivé instituce manuálně kopírovat z webových stránek do MS Excel. Výsledný sešit v MS Excel obsahuje několik listů podle kategorie institucí a v každé tabulce jsou následující údaje pro každý záznam: identifikátor, název úřadu, název pobočky, typ působnosti, působnost a adresa. Identifikátor představuje uměle vytvořený unikátní kód pro každý záznam ve formátu PK-zkratka instituce-pořadové číslo (např. PKHZS12 znamená dvanáctý záznam pro kategorii

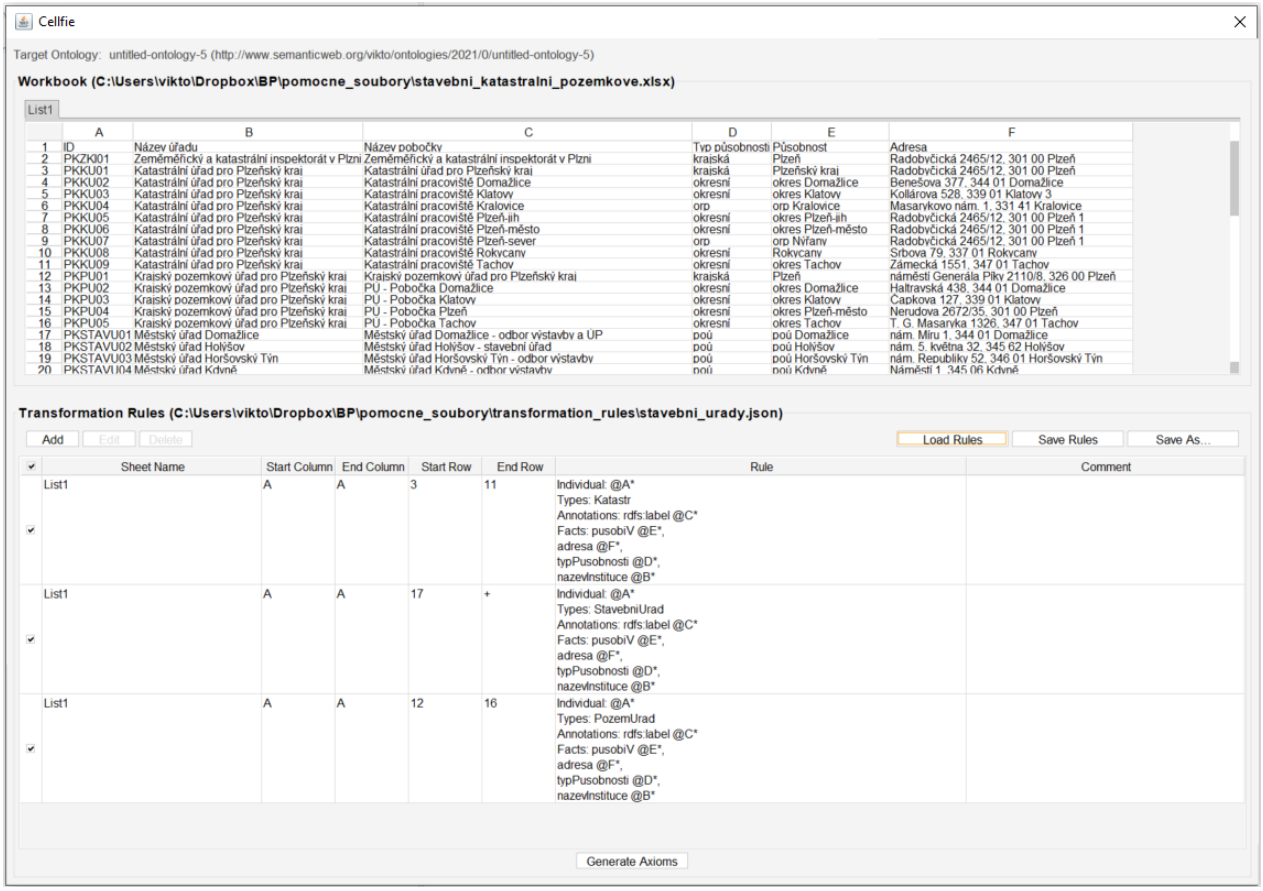
[2] Dostupné z: $http://apl.czso.cz/irso4/export1.jsp?kodcis=55$

Hasičský záchranný sbor a PK je označení pro Plzeňský kraj). Tyto identifikátory bylo nutné vytvořit proto, aby každý vkládaný individuál do ontologie měl unikátní IRI. U územních jednotek byl tento požadavek vyřešen tím, že každá územní jednotka má již svůj unikátní kód ve zdrojovém registru.
Všechny tabulky vytvořené pro import dat jsou součástí odevzdaných souborů. Jedná se o následující soubory:

- PK-uzemni-cleneni.xlsx
- PK-instituce.xlsx
- ORP-Klatovy-uzemni-cleneni.xlsx
- ORP-Klatovy-instituce.xlsx

4.3 Plnění ontologie daty

Jak již bylo zmíněno v minulé kapitole, jsou dvě možnosti pro vytváření nových axiomů ontologie - „manuálně“ v GUI editoru nebo psaním kódu v jazyce OWL 2. Pro plnění ontologie konkrétními daty je možné použít oba způsoby, ale každý se hodí v jiné situaci. První způsob jsem použila jen v několika jednotkách případů, kdy pro nějakou třídu existuje pouze jedna instance a bylo snadnější zadat hodnoty pomocí editoru. Ve většině případů ale bylo třeba k třídám ontologie vytvořit více (pro některé třídy až tisíce) individuálů a k tomu je nutné využít sofistikovanější nástroj pro hromadný import dat. Pro nahrávání dat z excelovských tabulek je k dispozici speciální plugin *Cellfie*, který umožňuje transformaci tabulkových dat do formátu OWL ontologie. Plugin je od verze 5.5.0 součástí samotného editoru Protégé a není nutné ho doinstalovávat a jeho použití je podrobně popsáno v [50] a [51].
Vstupem pro vytvoření nových axiomů z tabulky je soubor ve formátu .xlsx, obsahující jednu tabulku nebo případně celý sešit s několika listy, který se zobrazí v horní části okna pluginu. Je vhodné mít tabulku připravenou v takovém formátu, aby každý řádek představoval jeden nový vkládaný objekt a sloupce pak jeho IRI (tj. unikátní název), skutečný název vkládaný jako anotační vlastnost *rdfs: label* a hodnoty dalších vlastností objektu. Následně je nutné sestavit transformační či převodní pravidla (*transformation rules*) pro převedení dat z tabulky do podoby axiomů v jazyce OWL 2. Pro jednu tabulku či sešit lze vytvořit celou řadu transformačních pravidel a ty lze následně ukládat jako .JSON soubor a opětovně je používat či upravovat.
Transformační vztahy pro vytvoření nových axiomů se definují pomocí Manchaster syntaxe jazyka OWL 2. Nejprve je nutné zadat název listu a definovat oblast v tabulce, ze které data převádíme (tj. počáteční a koncový sloupec a řádek nebo počáteční řádek a symbol +,



Obrázek 4.4: Ukázka použití pluginu *Cellfie* pro převod dat z excelovské tabulky do ontologie. Vlastní zpracování.

chceme-li vybrat vše až do konce tabulky). V prvním řádku samotného pravidla se definuje, o jaký objekt se jedná (tj. *Class, Individual*) a ve kterém sloupci tabulky se nachází IRI nově vytvářeného objektu nebo se přímo zadá jeho název. Pro objekty *Class* se dále uvádí jeho vlastnosti jako *EquivalentTo* nebo *SubClassOf* a opět se doplní příslušná hodnota či její umístění v tabulce. Pro *Individual* je možné zadat název třídy, které je instancí, uvozen jako *Types:* opět s názvem již existující třídy či odkazem do tabulky. Po uvedení *Facts:* následuje výčet dalších vlastností individuálu a jejich hodnot, případně i jejich datového typu. Jak *Class*, tak *Individual* může mít definované i anotační vlastnosti jako *Annotations:*. Všechny hodnoty přebírané z tabulky je možné také formátovat pomocí Manchaster syntaxe (např. formát čísel, vynechání bílých znaků, velikost písmen, nahrazení symbolů a pod.). Způsob odkazování na hodnoty v tabulce je vždy jednotný a to pomocí symbolu @ s označením buď konkrétní buňky v tabulce (např. @A2) nebo celého sloupce (např. @A*). Užitečný návod, jak formulovat transformační pravidla lze najít na příklad v [51].

Jsou-li definovaná všechna potřebná transformační pravidla pro převod dat, vygenerují se nové axiomy, které je pak možné přidat do existující ontologie. Pomocí pluginu *Cellfie* je možné vytvářet všechny typy objektů v ontologii, nejenom individuály, ale i třídy a nové

vlastnosti. Z takto vygenerovaných axiomů je možné vytvořit i úplně novou ontologii, což bylo vyzkoušeno při vytváření ukázkové „malé“ ontologie pro ORP Klatovy.

Ukázka transformačního pravidla pro import dat:

```
{
  "Collections": [
    {
      "sheetName": "List1",
      "startColumn": "A",
      "endColumn": "A",
      "startRow": "26",
      "endRow": "+",
      "comment": "",
      "rule":
       "Individual: @G*
       \n   Types: @I*
       \n   Annotations: rdfs:label @H*
       \n   Facts: id @G*(xsd:integer),
       \n          jeCasti @J*,
       \n          jeCasti @N*,
       \n          skladaSeZ @A*,
       \n          skladaSeZ @D*",
      "active": true
    }
  ]
}
```

Všechny .JSON soubory s transformačními pravidly, které byly použité pro naplnění ontologie daty, jsou součástí přílohy bakalářské práce.

4.4 Problémy při plnění ontologie konkrétními daty

Při hromadném plnění ontologie daty jsem narazila na problém s nedostatkem alokované paměti pro běh JVM [3], který se objevoval ve dvou případech. Zaprvé při vkládání velkého množství nových objektů do ontologie najednou a zadruhé při spuštění reasoningu, bylo-li v ontologii uloženo velké množství individuálů. Podle informací v souboru Protege.l4j.ini je pro

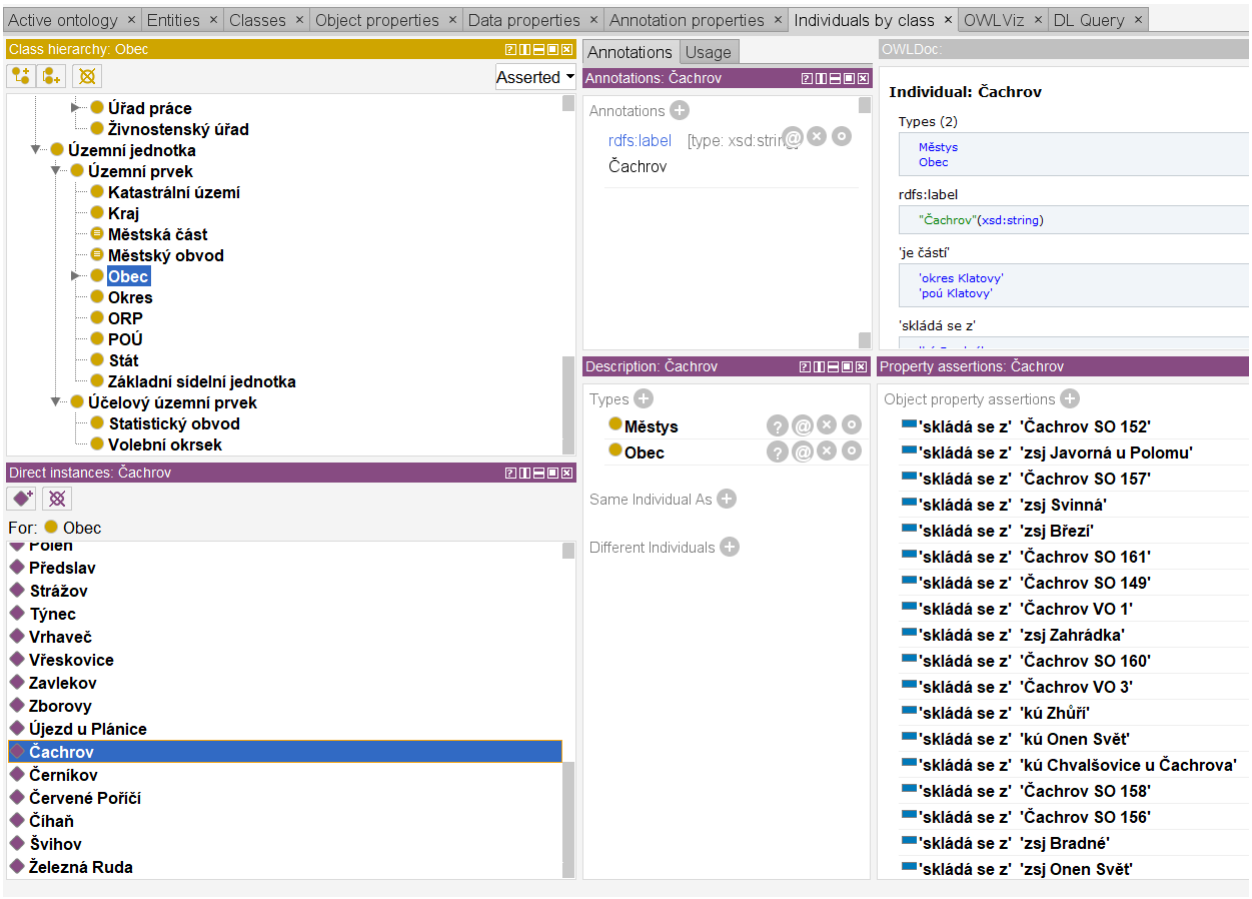
[3]JVM, Java Virtual Machine, je nástroj pro spouštění programů a skriptů v jazyce Java, případně i v jiných programovacích jazycích.

běh JVM alokovaná paměť 500 MB, což je zřejmě nedostačující. Vyzkoušela jsem proto v inicializačním souboru nastavit uvolněnou paměť nejprve na 2000 MB a následně na 4000 MB, což už je v případě mého počítače s 8GB RAM v podstatě mezní hodnota pro uvolnění paměti pro činnost nesystémového softwaru. Nicméně ani toto navýšení alokované paměti zmíněné problémy nevyřešilo. Z dostupných zdrojů jsem nedohledala, jakému alespoň přibližnému množství najednou zpracovávaných axiomů tato velikost uvolněné paměti odpovídá, a proto zde uvedu alespoň závěry, k jakým jsem při práci na geo-ontologii dospěla.
Při plnění ontologie daty pro Plzeňský kraj první situace nastala ve chvíli, kdy jsem pomocí pluginu Cellfie spustila vygenerování nových axiomů pro téměř 3 500 individuálů představující jednotlivé statistické obvody, kdy na každý individuál připadá šest nových axiomů. Nicméně pro import téměř 1 100 nových volebních okrsků a všech příslušných axiomů (v tomto případě připadala na jeden individuál pětice axiomů) byla alokovaná operační paměť dostačující. Osvědčilo se mi při hromadném plnění nahrávat do ontologie cca 5 000 nových axiomů najednou, neboť při větším objemu byl průběh velice pomalý a nebo úplně nemožný z důvodu nedostatku paměti.
S druhým případem, tedy s tím, že přestávaly fungovat reasonery, jsem se setkávala, když byla ontologie naplněna více než 50 000 axiomy. Pro takto velké množství axiomů, které je v rámci reasoningu nutné prošetřit, ověřit jejich konzistenci a navíc z nich odvodit další nové vztahy, je pak tento proces výpočetně velmi náročný a zdlouhavý a s postupným navyšováním počtu nových axiomů již nemožný. Objem zpracovávaných dat už je pak příliš velký na to, aby ho dokázala JVM zpracovat a vždy spuštěná akce skončila chybou `java.lang.OutOfMemoryError: GC overhead limit exceeded` a reasoning neproběhl.

4.4.1 Řešení problémů

Pro vyřešení výše uvedených potíží jsem nenašla lepší řešení, než snížit počet axiomů v jedné ontologii, a proto jsem se rozhodla do Geo-ontologie veřejné správy a samosprávy v Plzeňském kraji nepřidávat individuály pro účelové územní prvky, které s problematikou státní správy a územní působností jednotlivých institucí přímo nesouvisí. Navíc jejich vypuštěním z ontologie ubyde téměř 30 000 axiomů, což znatelně urychlí jak reasoning, tak i načítání objektů v ontologii při jejich prohlížení a editaci. Odevzdávaná geo-ontologie tedy obsahuje všechny třídy včetně tříd pro volební okrsky a statistické obvody, ovšem jejich konkrétní instance v ontologii uložené nejsou. Nicméně součástí digitální přílohy bakalářské práce jsou všechny soubory potřebné pro jejich případné nahrání do ontologie, tedy jak vstupní .xlsx soubor se všemi volebními okrsky a statistickými obvody, tak i příslušná transformační pravidla pro hromadný import pomocí pluginu Cellfie.
Aby bylo možné demonstrovat funkčnost navrženého modelu v jeho kompletní podobě včetně reasoningu, vytvořila jsem i „menší“ ukázkovou verzi geo-ontologie obsahující veškerá data

pro území jednoho ORP včetně všech individuálů i pro účelové územní prvky. Konkrétně se jedná o území správního obvodu ORP Klatovy a struktura ontologie je naprosto totožná s plnou verzí pro celý Plzeňský kraj. Liší se tedy pouze v počtu konkrétních instancí jednotlivých tříd, protože obsahuje pouze ty individuály, které se přímo nachází na území SO ORP Klatovy. K této ontologii nebylo třeba vytvářet žádné nové pomocné soubory, pouze stačilo upravit rozsah transformačních pravidel. Součástí digitální přílohy je tedy pouze zdrojový kód této ontologie.



Obrázek 4.5: Ukázka z geo-ontologie pro území ORP Klatovy. Vlastní zpracování.

Závěr

Cílem práce bylo nejprve se seznámit s problematikou ontologií a následně navrhnout geo-ontologii veřejné správy a samosprávy a vytvořený model otestovat na konkrétních datech pro Plzeňský kraj. V rámci rešerše odborné literatury jsem se kromě obecné problematiky ontologií a ontologického inženýrství zaměřila také na predikátovou logiku prvního řádu a deskripční logiku, které tvoří základ pro formulování axiomů ontologie. Dále jsem se věnovala vyhledávání již existujících ontologií, které obsahují stejné koncepty jako mnou zpracovávaná ontologie, aby bylo možné je navzájem provázat. Před zahájením vytváření samotné struktury ontologie jsem se podrobně seznámila s teoretickými přístupy a různými metodikami návrhu ontologií, s jednotlivými jazyky používanými pro popis ontologie a nakonec i s editorem Protégé (verze 5.5.0) určeným pro vytváření ontologií v jazyce OWL 2.

Při navrhování struktury ontologie jsem vycházela hlavně z platné legislativy České republiky a ontologie by měla být v souladu s právním stavem k 1. dubnu 2021. V procesu vymezování základních konceptů týkajících se veřejné správy jsem se zaměřila hlavně na definování obecných kategorií v takové podobě, aby výsledná struktura byla co nejsrozumitelnější pro uživatele a zároveň co nejvěrněji vystihovala skutečný stav. Jsem si plně vědoma toho, že mnou navržené členění není jediným správným řešením a je otázkou například i navazující diplomové práce pokusit se navrhnout jiná řešení a následně otestovat jejich použitelnost v dalších aplikacích.

Po dokončení obecného modelu jsem vytvořenou strukturu postupně plnila konkrétními daty pro Plzeňský kraj, aby byla otestována použitelnost modelu v praxi. Při naplňování prázdné ontologie velkým množstvím dat jsem se setkala s několika problémy a nedostatky, které jsou spíše technického rázu (nedostatečná alokovaná paměť', přílišná výpočetní náročnost) a jejich možné řešení je popsáno v závěrečné kapitole bakalářské práce. Na základě ověření funkčnosti modelu pro Plzeňský kraj lze říci, že navržená geo-ontologie veřejné správy a samosprávy by měla být použitelná i pro další kraje České republiky s výjimkou hlavního města Prahy, jehož administrativní a správní členění je velmi specifické a značně odlišné od ostatní třinácti krajů.

Zdroje

[1] Vedral, J. (2020) Kontrolní závěr z kontrolní akce 19/14 - Zavedení elektronické identifikace a zajištění elektronického přístupu ke službám veřejné správy *Nejvyšší kontrolní úřad, Praha* Červenec 2020

[2] Svátek, V., Vacura, M. (2007). Ontologické inženýrství. *DATAKON 2007* Říjen 2007. 1-32.

[3] Anzenbacher, A. (1981) Úvod do filosofie. Vyd. 3., V Portálu 2. Praha: Portál, 2010. 377 s. ISBN 978-80-7367-727-5.

[4] Aristotelés. Metafyzika A 1-2. Přel. Antonín Kříž. *3. vyd., v nakl. Petr Rezek 2. Praha: Rezek, 2008*

[5] Svátek, V. (2002). Ontologie a WWW. *DATAKON 2002* Říjen 2002. 1-35.

[6] Gruber, T. R. (1993) A translation approach to portable ontology specifications. *Knowledge Acquisition* May 1993. 199–220.

[7] Borst, W. N. (1997). Construction of Engineering Ontologies. *PhD thesis, University of Twente, Enschede* 1997.

[8] Studer, R., Benjamins, V. R., & Fensel, D. (1998) Knowledge engineering, principles and methods. *Data and Knowledge Engineering, 25(1-2)* 1998. 161–197

[9] Bernaras, A., Laresgoiti, I., & Corera, J. (1996) Building and reusing ontologies for electrical network applications. *In Proceedings of the 12th ECAI* 1996. 298–302

[10] Berners-Lee, T. J., Hendler, J., & Lassila, O. (2001) The Semantic Web. *Scientific American Magazine* May 2001.

[11] Noy, N. F., & McGuiness, D. L. (2001) A Guide to Creating Your First Ontology *Ontology Development 101* January 2001. 1-24

[12] Uschold, M., Gruninger, M. (1996) Ontologies: principles, methods and applications. *The Knowledge Engineering Review, Vol.11:2* February 1996. 93-136

[13] van Heijst, G., Schreiber, G., & Wielinga, B.J. (1997) Using Explicit Ontologies in KBS Development. *Int. J. Human-Computer Studies - special issue* 1996. 183 - 292

[14] Henriksson, R., Kauppinen, T., & Hyvönen, E. (2008). Core Geographical Concepts: Case Finnish Geo-ontology. *In Proceedings of the First International Workshop on Location and the Web.* January 2008. 57-60.

[15] Fonseca, F., Camara, G. (2019). Geo-Ontologies. *The GIS Manual - Chapter 12* September 2013. 157-164.

[16] Baglioni, M., Masserotti, M.V., Renso, CH., & Spinsanti, L. (2007). Building Geospatial Ontologies from Geographical Databases. *Conference Paper: GeoSpatial Semantics, Second International Conference, GeoS 2007, Mexico City* November 2007.

[17] Pattison, W.D. (1964). Four Traditions of Geography. *Journal of Geography 63* May 1964. 211-216

[18] Matlovič, R., Matlovičová, K. (2015) Geograficke myslenie *FHPV, Prešovská univerzita v Prešově, GRAFOTLAČ Prešov, s.r.o.* 2015. 37-38

[19] Smith, B. & D. Mark. (1998). Ontology and geographic kinds. *Proceedings of the Eighth International Symposium on Spatial Data Handling, held in Vancouver, British Columbia, Canada.* 1998. 308–320

[20] Salter, C.L., & Hobbs, J.J. (1998). Essentials of World Regional Geography. *Pacific Grove, CA: Brooks/Cole-Thomson* January 1998. ISBN 978-0030232817

[21] Kavouras, M. & Kokla, M. (2008). Theories of Geographic Concepts: Ontological Approaches to Semantic Integration *CRC Press* September 2019 ISBN 978-0367387914

[22] Sowa, J. (2010). The Role of Logic and Ontology in Language and Reasoning. *Theory and Applications of Ontology: Philosophical Perspectives* January 2010. 231-263

[23] GeoNames Ontology: Documentation (2019)
dostupné z: http : *//www.geonames.org/ontology/documentation.html*

[24] Kim, S., Iglesias-Sucasa, M., & Viollier, V. (2013). The FAO Geopolitical Ontology: A Reference for Country-Based Information *Journal of Agriculture & Food Information, 14* 2013. 50-65

[25] Berdier, Ch., & Roussey, C. (2007) Urban Ontologies: the Towntology Prototype towards Case Studies *Computational Intelligence (62)* 2007. 143-155

[26] EUROTL Framework Overview
dostupné z: https : *//www.roadotl.eu/static/eurotl − ontologies/index.html*

[27] Walicki, M. (2011) Introduction to Mathematical Logic *World Scientific Publishing Company* December 2011. ISBN: 978-9814343879

[28] Simons, P. M. (1991) Free Part-Whole Theory, in K. Lambert (ed.) *Philosophical Applications of Free Logic, New York and Oxford: Oxford University Press* 1991. 285-306.

[29] O'Connor, J. J., & Robertson, E. F. (1996) A history of Topology *Math History. School of Mathematics and Statistics University of St Andrews, Scotland* May 1996.

[30] Smith, B. (1996) Mereotopology: A Theory of Parts and Boundaries *Data and Knowledge Engeneering, 20* 1996. 287-303.

[31] Varzi, A. C., (2003) Mereology *The Stanford Encyclopedia of Philosophy (Spring 2019 Edition), Edward N. Zalta (ed.)* January 2019

[32] Casati, R., & Varzi, A. C., (1994) Holes and Other Superficialities *Cambridge, Mass. and London: MIT Press.* 1994

[33] Eschenbach, C., & Heydrich, W. (1995) Classical mereology and restricted domains *International Journal of Human-Computer Studies, Volume 43, Issues 5–6* November 1995, 723-740

[34] Baader, F., McGuiness, D. L., Nardi, D., Patel-Schneider, P. F. (2010) The Description Logic Handbook: Theory, Implementation and Applications *Cambridge University Press, 2nd Edition*

[35] Borgewardt, S. (2018) Logic-Based Ontology Engineering Part 1: Introduction to Ontology Engineering, OWL 2, and Description Logics *Institute of Theoretical Computer Science, Chair of Automata Theory, Technische Universität Dresden* 2018

[36] Čerba, O. (2013) Ontologie. *Přednáška z předmětu Socioekonomická geografie pro geomatiku (KMA/SGG), Západočeská univerzita Plzeň* 2010

[37] Čerba, O.(2011) Ontologie jako nástroj pro návrhy datových modelů vybraných témat příloh směrnice INSPIRE. *Praha: Univerzita Karlova v Praze. Disertační práce.* 2011

[38] Bock, J., Haase, P., Ji, Q., Volz, R. (2007) Benchmarking OWL Reasoners. *VLBD '07, Very Large Data Base Endowment, Vienna* September 2007. 23-28

[39] Corcho, O., Fernández-Lopez, M., Goméz-Peréz, A (2003) Methodologies, tools and languages for building ontologies: Where is their meeting point?. *Data & knowledge engineering, 2003, Vol. 46, No. 1)* January 2003. 41-64.

[40] Cardoso, J. (2007) The Semantic Web Vision: Where are We? *IEEE Intelligent Systems, September/October 2007* September 2007 pp. 22-26

[41] RDF 1.1 Concepts and Abstract Syntax *W3C Reccomendation 25 February 2014* Dostupné z: *https* : *//www.w3.org/TR/*2014*/REC* − *rdf*11 − *concepts* − 20140225*/resources* − *and* − *statements*

[42] OWL 2 Web Ontology Language Document Overview (Second Edition). *W3C Documentation* Dostupné z: *http* : *//www.w3.org/TR/owl*2 − *overview/*

[43] OWL 2 Web Ontology Language Quick Reference Guide. *W3C Documentation* Dostupné z: *http* : *//www.w*3*.org/*2007*/OWL/refcardA*4

[44] Pundt, H. (2007) From Idea toward Ontology *10th AGILE International Conference on Geographic Information Science* 2007

[45] Noy, F. N., McGuinness, D. L. (2001) Ontology Development 101: A Guide to Creating Your First Ontology *Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880* March 2001

[46] Průcha, P. (2012) Správní právo : obecná část. *8., dopl. a aktualiz. vyd.,. Brno: Doplněk, 2012.* 2012

[47] Marek, J., Pánková, M., Šímová, P. (2005) Veřejná správa v České republice *Ministerstvo vnitra ČR, úsek veřejné správy, 2. rozšířené vydání, Praha* 2005

[48] Soustava územních prvků a územně evidenčních jednotek | ČSÚ. Český statistický úřad | ČSÚ [online].
Dostupné z: *https* : *//www.czso.cz/csu/rso/soustava* − *uzemnich* − *prvku* − *a* − *uzemne* − *evidencnich* − *jednotek*

[49] Protégé 5 Documentation. GitHub Pages [online].
Dostupné z: *http* : *//protegeproject.github.io/protege/*

[50] GitHub - protegeproject/cellfie-plugin: Protégé plugin for creating OWL ontologies from spreadsheets. GitHub: Where the world builds software · GitHub [online].
Dostupné z: *https* : *//github.com/protegeproject/cellfie* − *plugin*

[51] MappingMasterDSL · protegeproject/mapping-master Wiki · GitHub. GitHub: Where the world builds software · GitHub [online].
Dostupné z: *https* : *//github.com/protegeproject/mapping* − *master/wiki/MappingMasterDSL*

[52] Vaníček, J. et al. (2008) Predikátová logika *Teoretické základy informatiky. Kernberg Publishing s.r.o., Praha* 2008

[53] Varzi, A.C. (1996). Parts, Wholes, and Part-Whole Relations: The Prospects of Mereotopology. *Data Knowl. Eng., 20* November 1996. 259-286.

Seznam digitálních příloh

- geo-ontologie veřejné správy a samosprávy (pro Plzeňský kraj)
- geo-ontologie veřejné správy a samosprávy (pro SO ORP Klatovy)
- sešit s daty pro územní členění
- sešit s daty pro instituce
- transformační pravidla
- balík s vyexportovanou dokumentací

Digitální přílohy jsou uložené na cloudovém úložišti Google Drive a jsou dostupné na tomto odkazu:
https://drive.google.com/drive/folders/1hAFDV0kq6c9pQGrrDTjrnKY93ysy4EUv?usp=sharing.

Seznam vložených příloh

- Příloha 1
  - Úvod do matematické logiky a predikátové logiky prvního řádu
  - Mereotopologické axiomy a definice
  - Deskripční logika
- Příloha 2
  - Podtřídy třídy Územní jednotka
  - Podtřídy třídy Instituce

A Přílohy

A.1 Příloha 1

A.1.1 Úvod do matematické logiky a predikátové logiky prvního řádu

Vzhledem k faktu, že mereotopologie i deskripční logika vychází z predikátové logiky prvního řádu, je vhodné si nejprve definovat základní termíny používané v logice [52]:

- **universum** je zde synonymem pro definiční obor
- **proměnná** je zástupcem libovolného objektu z universa, rozlišují se volné a vázané proměnné: ve formuli $\forall y P(x, y)$ je y proměnná vázaná kvantifikátorem $\forall$ a x je proměnná volná
- **term** označuje jakoukoli individuální proměnnou nebo funkci (tj. funkční vztah, který přiřazuje prvkům definičního oboru jiné prvky definičního oboru).
- **výrok** je obecně v logice jakékoli tvrzení, u kterého má smysl řešit, zda je pravdivé nebo nepravdivé. V predikátové logice je jako výrok označované takové tvrzení, které zmiňuje nějaký konkrétní objekt z dané množiny D a jeho vlastnost
  Příklad: Výrok "Plzeň má více obyvatel než Klatovy" zmiňuje z množiny obcí konkrétní objekty Plzeň a Klatovy a lze ihned vyhodnotit jako pravdivý).
- **predikát (výroková formule)** je zobecněním výroku a na rozdíl od výroku nezmiňuje elementární prvek množiny, ale proměnnou zastupující prvky definičního oboru. Predikát je funkcí vracející booleovskou hodnotu, ale výrokové formě jako takové nelze přisuzovat pevnou pravdivostní hodnotu - ta závisí až na dosazení za proměnné z definičního oboru
  Příklad: Pro zobecnění výroku "Plzeň má více obyvatel než Klatovy" použijeme následující postup: predikát "má více obyvatel než" označíme P a proměnné z definičního oboru obcí označíme x, y. V obecné podobě pak výrokovou formuli zapíšeme jako $P(x, y)$ a pro konkrétní prvky z definičního oboru $p =$ Plzeň a $k =$ Klatovy formuli zapíšeme a hned ji můžeme i vyhodnotit jako $P(p, k) = 1$).
- **kvantifikátor** je symbol vyjadřující míru přítomnosti dané vlastnosti (predikátu) v definičním oboru - $\forall$ znamená, že predikát platí pro všechny prvky definičního oboru

a ∃ vyjadřuje, že existuje aspoň 1 prvek z definičního oboru, pro který predikát platí, aby bylo dané tvrzení pravdivé.
Příklad: Tvrzení "existuje aspoň jedno město, které má více obyvatel než Klatovy" můžeme zapsat jako $\exists x P(x, k)$ a víme, že pro hodnotu $x = p$ je tvrzení pravdivé.

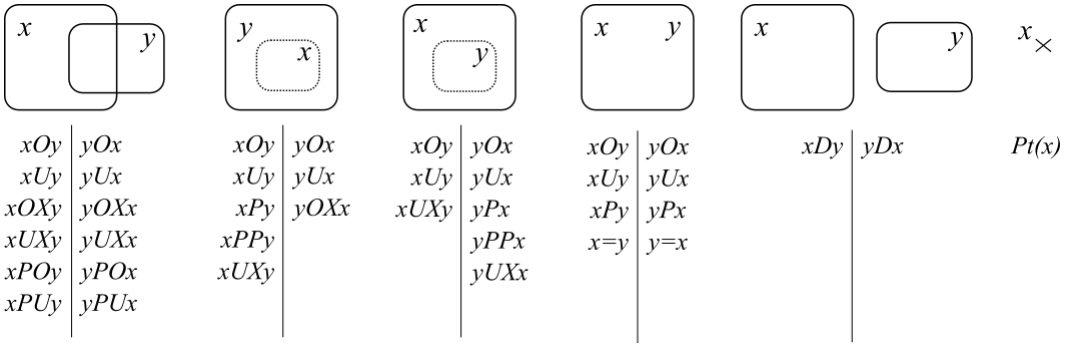
- **další používané symboly**:

$\neg$	negace
$\wedge$	a (zároveň)
$\vee$	nebo
$\rightarrow$ nebo $\Rightarrow$	implikace
$\leftrightarrow$ nebo $\Longleftrightarrow$	právě tehdy, když (ekvivalence)
$\equiv$	je ekvivalentní, odpovídá
$:=$	je definováno jako
x'	doplněk x

A.1.2 Mereotopologické axiomy a definice

Úvodem je třeba poznamenat, že neexistuje shoda v zápisu logických formulí mezi jednotlivými autory a v dalším textu je použitý formát zápisu vycházející z práce Barryho Smitha. Na příklad predikát P, že x je částí y lze zapsat následujícími způsoby: xPy, Pxy a $P(x, y)$, kde právě první uvedená možnost zápisu je podle Barryho Smitha, druhý zápis používá Achille C. Varzi a třetí varianta je převzatá z klasické predikátové logiky.
Pro axiomatické vyjádření mereotopologických vztahů je nutné vedle základních operátorů predikátové logiky definovat 2 nelogická primitiva - *část P (part)* a *vnitřní část IP (interior part)*. Mezi primitivy P a IP pak definujeme *spojení C (connection)*. Primitivum *část P* pak častěji uvádíme jako tvrzení "x je součástí y" nebo jinými slovy "x je část y" a zapíšeme ho jako xPy. Tento vztah vyjadřuje, že x je jakoukoliv částí y a může i znamenat, že je v určitém případě x identické s y. Obdobně lze zapsat, že "x je vnitřní částí y" jako $xIPy$.
Dále je třeba uvést mereologické primitivum *je řádnou částí PP (is a proper part of)*, které lze vyjádřit také jako $xPPy$ a platí pro něj následující definice: $xPPy := xPy \wedge \neg yPx$, tedy že x je řádnou částí y, když y není částí x. Pokud všechny řádné části objektu X jsou také řádnými částmi objektu Y, pak je X částí Y. Také platí, že mají-li dva objekty totožné své řádné části, pak se jedná o shodné objekty.
Další mereologické vztahy, které podmiňují definici mereotopologických axiomů, jsou uvedené v následujícím přehledu a na obrázku A.1

x **překrývá** y:	$xOy := \exists z(zPx \wedge zPy)$	*overlap*
x **je překryto** y:	$xUy := \exists z(zPx \wedge zPy)$	*underlap*
x **kříží** y:	$xOXy := xOy \wedge \neg xPy$	*over-crossing*
x **je překříženo** y:	$xUXy := xOy \wedge \neg yPx$	*under-crossing*
x **řádně překrývá** y:	$xPOy := xOXy \wedge yOXx$	*proper overlap*
x **je řádně překryto** y:	$xPOy := xUXy \wedge yUXx$	*proper underlap*
x **je oddělené** od y:	$xDy := \neg xOy$	*is discrete from*
x **je bod**:	$Pt(x) := \forall y(yPx \to y = x)$	*is a point*



Obrázek A.1: Znázornění základních mereologických vztahů Vlastní zpracování podle [53]

Nyní můžeme vyvodit první dva mereotopologické axiomy definující základní vlastnosti vztahu *x je část y* xPy:

AP1: $xPy := \forall z(zOx \to zOy)$
AP2: $xPy \wedge yPx \to x = y$ P je antisymetrická relace

AP1 říká, že x je částí y, když pro všechna z platí, že překrývá-li z prvek x, pak z překrývá také y. Antisymetrie znamená, že x a y jsou částí jeden druhého jen tehdy, jsou-li x a y shodné. Z axiomů **AP1** a **AP2** lze vyvodit následující teorémy:

TP1: xPx P je reflexivní relace (tj. x je částí samo sebe)
TP2: $xPy \wedge yPz \to xPz$ P je tranzitivní relace

Tranzitivita relace znamená, že je-li x část y a zároveň je y část z, pak x je částí z.
Pro další definice je nutné zavést tzv. *termo-tvorný (term-forming) operátor* σ *a* ι, který slouží k rozlišení termů od výroků a pomocí něj lze rozeznat singulární term odkazující na unikátní prvek z universa, pro který je formule pravdivá. Dále se zavádí formule φ vyjadřující podmínku s jednou volnou proměnnou x, o které můžeme říci, že je splněna, pokud je věta φx pravdivá aspoň pro jednu hodnotu x. Můžeme pak definovat *sumu (sum, fusion)* všech entit, které splňují podmínky φ. Tato *suma* se zapisuje jako $\sigma x(\varphi x)$ a platí pro ni následující definice:

$$\sigma x(\varphi x) := \iota y(\forall w(wOy \iff \exists v(\varphi v \wedge wOv))),$$

což znamená, že suma $\sigma x(\varphi x)$ odpovídá určité entitě y, pro kterou platí, že ji překrývá nějaká entita w, která zároveň překrývá jinou entitu v splňující podmínku φ. Z této definice pak plyne další teorém:

$$\textbf{TP3}: \quad y = \sigma x(\varphi x) \rightarrow \forall x(\varphi x \rightarrow xPy)$$

Ten říká, že je-li y sumou entit splňujících podmínku φ, pak pro všechna x, která splňují stejnou podmínku φ, vyplývá, že jsou částí y. Není definována prázdná suma, a tedy ani $\sigma x(\varphi x)$, pokud podmínka φ není splněna.

Z axiomu identity $\exists x(x = x)$ se vyvozuje teorém dokazující existenci *universa* (tj. definičního oboru):

$$\textbf{TP4}: \quad \exists x \forall y(yPx)$$

Dále definujeme následující vztahy mezi entitami, které mají svůj základ v teorii množin:

sjednocení	$x \cup y := \sigma z(zPx \vee zPy)$
průnik	$x \cap y := \sigma z(zPx \wedge zPy)$
doplněk	$x' := \sigma z(zDx)$
rozdíl	$x - y := \sigma z(zPx \wedge xDy)$

Na tyto vztahy pak navazuje další teorém týkající se principu doplňku:

$$\textbf{TP5}: \quad xPy \wedge x \neq y \rightarrow \exists z(z = y - x)$$

Kromě topologického primitiva *je část P* xPy dále definujeme primitivum *je vnitřní částí IP* $yIPy$, což je speciální případ xPy, kdy x se nedotýká hranice y. Případ, kdy x je řádnou částí y $xPPy$ a dotýká se zevnitř hranice, nazýváme x jako tečnou řádnou část y a vztah označíme jako $xTPPy$. Obecně pro jakkoli překrývající se, protínající se, splývající, sjednocené či jinak se dotýkající entity x a y platí, že *jsou propojené xCy*. Pokud se entity dotýkají pouze z vnějšku svými hranicemi, pak říkáme, že *jsou propojené zvenčí xECy*.
Pro *vnitřní část IP* můžeme odvodit následující vlastnosti:

AIP1:	$xIPy \rightarrow xPy$	IP je speciální případ P
AIP2a:	$xIPy \wedge yPz \rightarrow xIPz$	
AIP2b:	$xPy \wedge yIPz \rightarrow xIPz$	
AIP3:	$xIPy \wedge xIPz \rightarrow xIP(y \cap z)$	podmínka pro průnik
AIP4:	$\exists x(\varphi x) \wedge \forall x(\varphi x \rightarrow xIPy) \rightarrow \sigma x(\varphi x)IPy$	podmínka pro sumu
AIP5:	$\exists y(xIPy)$	podmínka pro existenci vnitřní části

Z axiomu **AIP5** lze vyvodit teorémy, že celé universum je vnitřní částí samo sebe a každá entita x je vnitřní částí universa:

$$\textbf{TIP1}: \quad \mathbf{I}IP\mathbf{I}$$
$$\textbf{TIP2}: \quad \forall x(xIP\mathbf{I})$$

Aby bylo možné definovat, že x *je hranicí* y xBy, je nutné definovat vztah x *protíná* y xXy:

$$xXy := \neg xPy \wedge \neg xDy$$

nebo obdobně pro $y \neq \mathbf{I}$:

$$xXy := xOy \wedge xO(\mathbf{I} - y).$$

Z toho vyplývá, že žádná entita nemůže protínat samu sebe a universum protíná všechny entity, kromě entit identických s universem.

Dále je nutné definovat vztah x *překlenuje* y $xSty$ jako

$$xSty := \forall z(xIPz \rightarrow zXy)$$

a z toho plynoucí $xSty \rightarrow \neg xIPy$. Tyto definice říkají, že entita x překlenuje entitu y v takovém případě, kdy vše, čeho je x vnitřní částí, protíná y a aby mohlo x překlenovat y, pak nesmí být jeho vnitřní částí. Z této definice se odvozuje teorém **TIP3**:

$$\textbf{TIP3}: \quad xPy \rightarrow xIPy \vee xSty$$

Teorém **TIP3** říká vlastně totéž, co tvrzení $\neg xIPx \rightarrow xStx$, že každá část nějakého objektu je buď jeho vnitřní částí a nebo ho překlenuje.

Entity, které překlenují nějakou jinou entitu, lze rozdělit do dvou skupin na hranice a tečny. První možnost, kdy x *je hranicí* y xBy, definujeme takto:

$$xBy := \forall z(zPx \rightarrow zSty),$$

a druhý případ, kdy x je *tečnou* y xTy pak definuje:

$$xTy := \exists z(zPy \wedge zBy),$$

Tečna entity y je tedy taková entita, která má jako svoji část hranici entity y. Pro všechny tečny platí, že danou tečnou entitu i překlenují. Pro všechny hranice entity y platí, že jsou zároveň i jeho tečnami a tudíž nemohou být jeho vnitřní částí. Z toho pak plynou následující teorémy:

TB1:	$xBy \equiv \forall z(zPx \rightarrow zTy)$	všechny části hranice jsou tečnami
TB2:	$xBy \wedge yBz \rightarrow xBz$	vztah je tranzitivní
TB3:	$xPy \wedge yBz \rightarrow xBz$	
TB4:	$xT(y \cup z) \rightarrow xTy \vee xTz$	hranice sjednocení entit
TB5:	$\forall y(\varphi x \rightarrow xBy) \rightarrow \sigma x(\varphi x)By$	hranice sumy

Dále je možné definovat pro entitu x její *uzavření cl(x)*:

$$cl(x) := x \cup \sigma y(yBx) \text{ pro } x \neq \mathbf{I},$$

což znamená, že uzavření entity x $cl(x)$ odpovídá sjednocení této entity se sumou všech y, které jsou hranicí x. Pro *uzavření x cl(x)* platí Kuratowského axiomy:

AK1:	$xPcl(x)$
AK2:	$cl(cl(x)) = cl(x)$
AK3:	$cl(x \cup y) = cl(x) \cup cl(y)$

Následující definice *maximální hranice bdy(x)* odpovídá standardní topologické definici hranice entity jako průsečnice uzavření $cl(x)$ této entity s uzavřením jejího doplňku $cl'(x)$. Definici *maximální hranice bdy(x)* lze také zapsat jako:

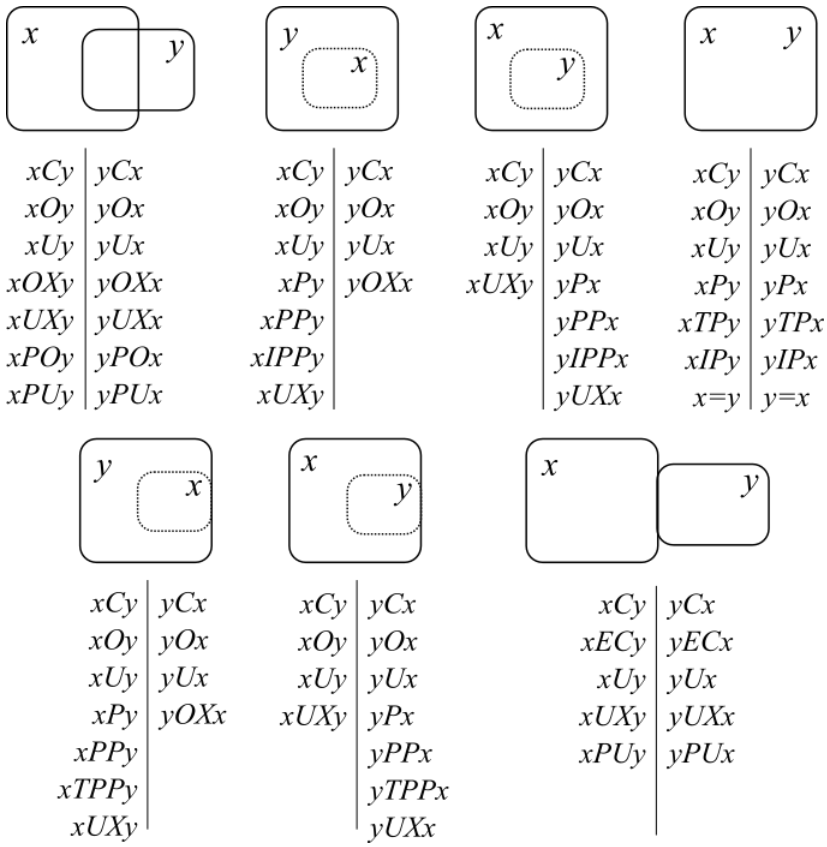
$$bdy(x) := \sigma y(yBx)$$

Stejně tak definice *vnitřního prostoru int(x)* koresponduje s topologickou definicí vnitřního prostoru, který je určený jako rozdíl entity a její hranice. *Vnitřní prostor int(x)* definujeme také jako:

$$int(x) := \sigma y(yIPx).$$

O entitě můžeme říci, že je otevřená, pokud je identická se svým vnitřním prostorem $int(x)$ a je uzavřený její doplněk.

Na následujícím obrázku A.2 jsou znázorněny vybrané zásadní mereotopologické vztahy, které byly definovány v předchozím textu. Z ilustrace je názorně vidět, že mereotopologie v sobě skutečně propojuje základní mereologické vztahy s topologickými prvky.



Obrázek A.2: Znázornění základních mereotopologických vztahů Vlastní zpracování podle [53]

A.1.3 Deskripční logika

Deskripční logika vychází z predikátové logiky, ale liší se používaná terminologie, a proto zde na úvod představím rozdíly v používaných pojmech [34]:

- **Univerzální koncept** je zde synonymem pro definiční obor nebo universum - značí se Δ nebo $\top$ a také se nazývá jako nejobecnější třída (*top*)
- **Prázdný koncept** je ekvivalentem prázdné množiny - značí se $\emptyset$ nebo $\bot$ a také se nazývá jako nejspecifičtější třída (*bottom*)
- **Koncept** je označení pro obecnou množinu, která je podmnožinou univerzálního konceptu - značí se $\mathbf{C}, \mathbf{D}$, případně dalšími velkými písmeny
- **Individuály** jsou konkrétní prvky spadající do nějakého konceptu a značí se x, y, případně dalšími malými písmeny. Příslušnost individuálu x konceptu $\mathbf{C}$ se zapíše jako $C(x)$
- **Role** jsou množiny uspořádaných dvojic (tj. binární relace), které udávají vztahy mezi koncepty a individuály - značí se $\mathbf{R}$

- **Axiomy** jsou logická tvrzení v deskripční logice pracující s koncepty, individuály a rolemi
- **Konstruktory ∀ a ∃** se v deskripční logice používají pro vyjádření toho, kolik objektů splňuje daný axiom. Nazývají se jako univerzální (∀) a existenční (∃) omezení a lze jejich význam vyjádřit pomocí predikátové logiky následovně:

$$\forall R.C: \quad x; \forall y(R(x,y) \rightarrow C(y)),$$

kde R je role, C je koncept a x,y jsou individuály a platí pro ně, že se jedná o takovou množinu prvků x, kde všechna y, pro která platí R(x,y), spadají pod koncept C(y). Obdobně pro

$$\exists R.C: \quad x; \exists y(R(x,y) \wedge C(y)),$$

platí, že se jedná o takovou množinu prvků x, pro které existuje y, kde platí zároveň R(x,y) a C(y).

- **další používané symboly**:

¬	negace	doplněk množiny
⊓	a (zároveň)	průnik množin (konjunkce)
⊔	nebo	sjednocení množin (disjunkce)

Pro roli R je možné omezit její kardinalitu, resp. omezit počet objektů, pro které daný axiom platí:

$$\leq nR \text{ resp. } \geq nR: \quad |R(x)| \leq n \text{ resp. } |R(x)| \geq n$$

Lze vyjádřit i tzv. kvalifikované omezení kardinality nejenom pro roli R, ale i koncept C:

$$\leq nR.C \text{ resp. } \geq nR.C: \quad |(R(x) \wedge C(y))| \leq n \text{ resp. } |(R(x) \wedge C(y))| \geq n$$

Výše uvedené logické vztahy lze libovolně kombinovat a vytvářet kromě jednoduchých tvrzení i složité axiomy.

V následujících příkladech uvedu v ontologiích často používané vztahy vyjádřené pomocí deskripční logiky.

$$\textbf{Subsumpce:} \quad C \sqsubseteq D,$$

kde koncept C je nadřazený konceptu D - např. *Obec* $\sqsubseteq$ *Městys*.

Instance: $C(x)$,

kde individuál x je instancí konceptu C - např. *Město(Klatovy)*.

Objektová vlastnost: $R(x, y)$ nebo $R(C, D)$,

kde individuály x,y nebo koncepty C,D jsou ve vztahu R - např. *je částí(Kraj, Stát)*.

Inverzní objektová vlastnost: $R^{-}(C, D)$,

kde role R a R^{-} jsou vzájemně inverzní - např. R = *působí v* a R^{-} = *spadá do působnosti*.

Ekvivalence: $C \equiv D$,

kde koncept C je ekvivalentní s konceptem D - např. *Městský obvod* $\equiv$ *Městská část*.

Disjunkce: $|C \sqcup D| = \emptyset$,

kde koncepty C a D jsou disjunktní, tedy mají prázdný průnik - např. $|$ *Instituce* $\sqcup$ *Územní jednotka*$| = \emptyset$.

Přehled různých deskripčních logik

Různé formy deskripční logiky se od sebe odlišují svojí expresivností, tedy schopností vyjádřit určité vztahy. Rozlišují se tři základní typy deskripční logiky:

1. $\mathcal{AL}$ = *Attributive language* - nejzákladnější jazyk umožňující vyjádřit:
   - atomickou negaci (tj. negaci konceptů, které se nevyskytují na levé straně axiomů)
   - průnik konceptů
   - univerzální omezení
   - univerzální koncept
   - prázdný koncept
2. $\mathcal{FL}$ = *Frame based language* - jazyk založený na rámcích umožňující vyjádřit:
   - průnik konceptů
   - univerzální omezení
   - omezení rolí
3. $\mathcal{EL}$ = *Existential language* - existenční jazyk umožňující vyjádřit:
   - průnik konceptů
   - existenční omezení

V následujícím přehledu je pak uvedený stručný soupis některých rozšíření těchto základních deskripčních logik včetně určení jejich expresivnosti [34]

- $\mathcal{F}$: vymezení dalších funkčních vlastností (např. kvantifikace unikátnosti)
- $\mathcal{E}$: určení existenčních omezení
- $\mathcal{U}$: sjednocení konceptů
- $\mathcal{C}$: komplexní vyjádření negace
- $\mathcal{H}$: vyjádření hierarchie rolí
- $\mathcal{R}$: vyjádření reflexivity a antireflexivity rolí, disjunkce rolí
- $\mathcal{I}$: vyjádření inverzních vlastností
- $\mathcal{N}$: omezení kardinality
- $\mathcal{Q}$: kvalifikované omezení kardinality
- $\mathcal{D}$: použití datatypových vlastností

Konkrétní deskripční logika se pak označuje kombinací písmen příslušné základní logiky a jejích rozšíření. Mezi typicky užívané deskripční logiky patří $\mathcal{ALC}$ a její další rozšířené expresivnější verze [2].

A.2 Příloha 2

A.2.1 Podtřídy třídy Územní jednotka

Územní prvek	Stát		
	Kraj		
	Okres		
	Správní obvod ORP		
	Správní obvod POÚ		
	Obec	Obec prvního stupně	
		Obec druhého stupně	
		Obec třetího stupně	
		Městys	
		Město	Statutární město
		Vojenský újezd	
	Městský obvod/městská část		
	Katastrální území		
	Základní sídelní jednotka		

Účelový územní prvek	Volební okrsek
	Statistický obvod

A.2.2 Podtřídy třídy Instituce

Státní správa	Dopravní inspektorát	
	Finanční správa	Finanční úřady Specializovaný finanční úřad Odvolací finanční ředitelství
	Matriční úřady	Matriční úřady Zvláštní matrika
	Česká správa sociálního zabezpečení	
	Úřad práce	Úřad práce – Kontaktní pracoviště Úřad práce – Státní sociální podpora
	Stavební úřady	
	Živnostenský úřad	

Samospráva	Krajské úřady		Odbory krajského úřadu
	Úřady obcí	Obecní úřad	Úřad obce prvního stupně Úřad obce druhého stupně Úřad obce třetího stupně
		Úřad městysu	
		Městský úřad	Odbor městského úřadu Úřad městského úřadu Úřad městského obvodu/části

Ostatní veřejná správa	Český báňský úřad	
	Český telekomunikační úřad	
	Drážní správní úřad	Drážní inspekce Drážní úřad
	Katastrální úřad	
	Pozemkový úřad	
	Rada České republiky pro rozhlasové a televizní vysílání	
	Ředitelství silnic a dálnic	
	Ředitelství vodních cest	
	Správa uprchlických zařízení	
	Správa úložišť radioaktivních odpadů	
	Správa základních registrů	
	Správa železnic	
	Státní plavební správa	
	Úřad pro civilní letectví	
	Úřad pro přístup k dopravní infrastruktuře	
	Úřad pro zastupování státu ve věcech majetkových	
	Úřad průmyslového vlastnictví	
	Zeměměřický úřad	

Sbory	Armáda ČR		Krajská vojenská ředitelství Úřady vojenských újezdů Vojenské lesy a statky ČR
	Celní správa ČR		Celní úřady
	Hasičský záchranný sbor		Krajské ředitelství HZS Požární stanice
	Policie ČR	Krajské ředitelství PČR	Územní odbor PČR Policejní stanice Městské ředitelství PČR Obvodní oddělení PČR
	Vězeňská služba ČR		

Inspektoráty a kontrolní úřady	Česká inspekce životního prostředí
	Česká obchodní inspekce
	Česká plemenářská inspekce
	Česká školní inspekce
	Inspektoráty práce
	Státní energetická inspekce
	Státní úřad inspekce práce
	Státní zemědělská a potravinářská inspekce
	Technická inspekce České republiky
	Zeměměřické a katastrální inspektoráty
	Energetický regulační úřad
	Nejvyšší kontrolní úřad
	Puncovní úřad
	Státní ústav pro kontrolu léčiv
	Úřad pro dohled nad hospodařením politických stran a politických hnutí
	Úřad pro mezinárodněprávní ochranu dětí
	Úřad pro ochranu hospodářské soutěže
	Úřad pro ochranu osobních údajů
	Ústav pro odborně technické zjišťování příčin leteckých nehod
	Ústav pro státní kontrolu veterinárních biopreparátů a léčiv
	Ústřední kontrolní a zkušební ústav zemědělský
	Úřad pro technickou normalizaci, metrologii a státní zkušebnictví

Ostatní instituce	Archiv	Bezpečnostní archiv	
		Archiv samosprávy	
		Specializovaný archiv	
		Státní archiv	Státní oblastní archiv Státní okresní archiv
	Bezpečnost	Národní úřad pro kybernetickou a informační bezpečnost Národní bezpečnostní úřad Bezpečnostní informační služba Státní úřad pro jadernou bezpečnost Státní ústav jaderné, chemické a biologické ochrany Státní ústav radiační ochrany Úřad pro zahraniční styky a informace	
	Cestovní ruch a místní rozvoj	Centrum pro regionální rozvoj České republiky Česká centrála cestovního ruchu	
	Doprava	Centrum služeb pro silniční dopravu Státní fond dopravní infrastruktury Řízení letového provozu ČR	
	Finance a obchod	Česká agentura na podporu obchodu Česká národní banka Evropské spotřebitelské centrum Správa státních hmotných rezerv Státní fond podpory investic Státní tiskárna cenin Státní zemědělský intervenční fond Agentura pro podporu podnikání a investic Finanční analytický úřad Hospodářská komora České republiky	
	Kultura a památková péče	Divadelní ústav Národní památkový ústav	
	Soudnictví a právo	Exekutorská komora ČR	Exekutorské úřady
		Finanční arbitr	
		Kancelář veřejného ochránce práv	
		Notářská komora ČR	Notáři
		Rejstřík trestů	
		Státní zastupitelství	Vrchní státní zastupitelství Nejvyšší státní zastupitelství Krajská státní zastupitelství Okresní státní zastupitelství
		Soudy	Vrchní soud Ústavní soud Nejvyšší soud Nejvyšší správní soud Krajské soudy Okresní soudy Městské soudy Obvodní soudy

	Školství	Centrum pro studium vysokého školství Centrum pro zjišťování výsledků vzdělávání-CERMAT Český institut pro akreditaci Národní pedagogický institut Národní vzdělávací fond	
	Výzkum a zkušební ústavy	Akademie věd ČR Centrum dopravního výzkumu Česká geologická služba Česká rozvojová agentura Český hydrometeorologický ústav Český metrologický institut Český statistický úřad Český úřad pro zkoušení zbraní a střeliva Český úřad zeměměřický a katastrální Elektrotechnický zkušební ústav Fyzikálně technický zkušební ústav Grantová agentura Institut pro kriminologii a sociální prevenci Státní zkušebna strojů Strojírenský zkušební ústav Technický a zkušební ústav stavební Technologická agentura Textilní zkušební ústav Ústav mezinárodních vztahů Ústav územního rozvoje Ústav jaderného výzkumu Řež Ústav pro studium totalitních režimů Ústav zemědělské ekonomiky a informací Výzkumný a vývojový ústav dřevařský Výzkumný ústav bezpečnosti práce Výzkumný ústav geodetický, topografický a kartografický Výzkumný ústav lesního hospodářství a myslivosti Výzkumný ústav meliorací a ochrany půdy Výzkumný ústav potravinářský Výzkumný ústav práce a sociálních věcí Výzkumný ústav rostlinné výroby Výzkumný ústav pro krajinu a okrasné zahradnictví Výzkumný ústav veterinárního lékařství Výzkumný ústav vodohospodářský T. G. Masaryka Výzkumný ústav zemědělské techniky Výzkumný ústav živočišné výroby	
	Zdravotnictví	Hygienická správa	Hlavní hygienik Krajské hygienické stanice
		Veterinární správa	Ústřední veterinární správa Krajské veterinární správy
		Státní zdravotní ústav	
		Ústav leteckého zdravotnictví	
		Zdravotní pojišťovny	
		Zdravotní ústavy	
		Nemocnice	

	Životní prostředí	Agentura ochrany přírody a krajiny
		CENIA, česká informační agentura životního prostředí
		Lesy ČR
		Ústav pro hospodářskou úpravu lesů
		Agrární komora ČR
		Státní fond životního prostředí
		Správa jeskyní ČR
		Správa národních parků