

UNIVERZITA PALACKÉHO V OLOMOUCI

PŘÍRODOVĚDĚCKÁ FAKULTA

katedra GEOINFORMATIKY

obor GEOINFORMATIKA

Romana VOLKOVÁ

MAPOVÁ ONTOLOGIE

DIPLOMOVÁ PRÁCE

Vedoucí práce: Prof. RNDr. Václav SNÁŠEL, CSc.

Olomouc 2010

Čestné prohlášení

Prohlašuji, že jsem zadanou diplomovou práci řešila sama a že jsem uvedla veškerou použitou literaturu. Všechna poskytnuta i vytvořená digitální data nebudu bez souhlasu školy poskytovat.

V Liberci, 1. května 2010

.....

PODĚKOVÁNÍ

Na tomto místě chci poděkovat za poskytnutí cenných rad a studijních materiálů:

Prof. RNDr. Václavu Snášelovi, CSc., vedoucímu práce.

Ing. Zdeně Dobešové, Ph. D.

RNDr. Janu Brusovi

OBSAH

1.	CÍL PRÁCE.....	6
2.	ÚVOD DO PROBLEMATIKY.....	7
2.1.	Ontologie obecně	7
2.2.	Ontologie v kartografii a gis.....	12
3.	METODY A POSTUP ZPRACOVÁNÍ	14
3.1.	Data.....	14
3.2.	Algoritmus tvorby ontologie.....	16
3.3.	Použitý software	16
4.	POPIS ALGORITMU FCA	17
5.	POPIS A ZPRACOVÁNÍ VSTUPNÍCH DAT	20
5.1.	Výběr dat	20
5.2.	Příprava dat	21
6.	EXTRAKCE ONTOLOGIE Z DAT	24
6.1.	Tvorba diagramu ontologie	24
6.2.	Práce s ontologií	36
6.3.	Analýza diagramů ontologie	41
7.	VÝSLEDKY A JEJICH VYUŽITÍ V PRAXI	47
7.1.	Srovnání ontologií legendy a mapového pole ZM 10	47
7.2.	Změny v mapě rozhleden na základě analýz	48
7.3.	Přínosy nástroje Toscanal.....	50
8.	DISKUSE	51
9.	ZÁVĚR	53
10.	LITERATURA.....	54
11.	SUMMARY	56
12.	PŘÍLOHY	58

ÚVOD

Výsledná podoba a kvalita kartografického díla je z naprosté většiny závislá na znalostech, schopnostech a umu tvůrce mapy. Sice existují nástroje, které usnadňují tvorbu mapy, nabízí předdefinované znaky, barevné palety, nástroje pro tvorbu popisu atp., ale na rozhodování "co v mapě bude a v jaké podobě?" zůstává autor tak trochu sám.

Tato práce se zabývá studiem extrakce ontologie z mapových dat. Ontologie jednoduše řečeno reprezentuje objekty, koncepty a další entity, jejichž existence je v daném zájmovém prostoru předpokládána, a vztahy, které jsou mezi jednotlivými entitami udržovány (Gruber, 1993). Ontologie tedy nabízí určitý prostor pro jednodušší rozhodování na základě vizualizace vztahů a závislostí mezi jednotlivými zájmovými objekty.

V praktické části je řešeno, které vztahy a závislosti mezi objekty hledat, jak objekty definovat a třídit. Následně se práce zabývá samotnou tvorbou ontologie, hodnotí použitý software. V závěrečné části jsou shrnuty poznatky o možných přínosech ontologie v oblasti kartografie.

1. CÍL PRÁCE

Podle zadání magisterské práce je cílem diplomové práce návrh, algoritmizace a praktická realizace nástroje pro extrakci ontologie z mapových dat.

Při seznamování se s tématem vyplynulo na povrch, že k extrakci ontologie mapových dat lze přistupovat různými způsoby. Jednak lze ontologie vytvářet nad již hotovým mapovým dílem, pak nám ontologie může posloužit k hodnocení správnosti zvolených kartografických vyjadřovacích prostředků, rozčlenění prvků do legendy a může pomoci k posouzení, zda mapa svým obsahem splňuje účel, za kterým byla vytvořena. Druhý přístup je opačný – ontologie slouží jako jakási myšlenková mapa, pomáhá nám utřídit si myšlenky, správně rozhodnout, co musí vytvářená mapa obsahovat, roztrždit entity a správně je vyjádřit. V obou zmíněných přístupech lze na mapu i na vytvářenou ontologii hledět z různých úhlů pohledu. A právě to je stěžejním cílem této diplomové práce – ukázat, co všechno může extrakce ontologie z mapových dat znamenat, k čemu a komu má sloužit.

Zadání práce hovoří o návrhu, algoritmizaci a praktické realizaci nástroje pro extrakci ontologie z mapových dat. Vzhledem k tomu, že nástrojů pro extrakci ontologie existuje několik, jsou dokonce i volně dostupné, nebylo by rozumné vytvářet již vytvořené, proto není tato práce zaměřena na tvorbu a programování nástroje, který by dokázal z map ontologii extrahovat. Naopak se práce soustředí na zkoumání možností, jak může již existující nástroj či algoritmus být užitečný pro oblast kartografie, jak ho používat, protože v této oblasti na základě studia odborných článků nebyl zaznamenán žádný překotný výzkum.

2. ÚVOD DO PROBLEMATIKY

2.1. Ontologie obecně

2.1.1. Historie

Počátky ontologie je třeba hledat ve filozofii. Samotný termín pochází z řečtiny, složením slov *ontos* – *jsoucí* a *logos* – *výklad*. Websterův výkladový slovník [20] hovoří o ontologii jako o odvětví metafyziky vztahující se k povaze a vztahů v jsoucnu, popř. specifické teorii o povaze bytí nebo druzích existence.

Ve filozofii se o ontologii hovoří jako o teorii přirozenosti existence. (Gruber, 2009) [2].

Přestože princip ontologie pochází již z doby Aristotelovy (např. Aristotelova ontologie obsahovala primitivní kategorie, jako jsou hmota a kvalita, pomocí nichž mělo jít vysvětlit všechno, co existuje), pojem ontologie se zažil až v 17. století. Začal ho používat německý filozof Christian Wolff, který chtěl pojmem ontologie nahradit pojem metafyzika (metafyzika ve smyslu učení o bytí). Ontologie jako filozofická disciplína se zajímá o bytí ve smyslu jeho nejobecnějších určení, vlastností a projevů. V tomto smyslu se tedy příliš nevzdaluje původnímu (Aristotelovskému) významu pojmu metafyzika (ve smyslu první filosofie). Podle Slovníku filozofických věd [21] je ontologie "pouze věda o obecném bytí, tj. bytí abstraktním, nikoliv absolutním, začíná abstrakcemi a nevychází z kruhu, nespočívá na žádné solidní bázi a nemá jiný výsledek, než diskreditaci metafyziky, s kterou byla často směřována".

Termín se v 80. letech 20. století začal využívat v oblasti informatiky, hlavně v oblasti výzkumu umělé inteligence. Vědci našli použitelnost ontologie v práci s matematickou logikou a vytvořili teorie, že v oblasti umělé inteligence by mohly vzniknout ontologie v podobě aritmetického modelu, který umožní z určité části automatizovat logické myšlení. Ontologie byla

použita při utváření základních stavebních kamenů pro modelování světa. V modelech vystupují agenti (např. autonomní roboti), kteří za použití specifického modelu dokážou chápat pouze tu část světa, kterou jejich ontologie reprezentuje, tedy pouze entity z jeho ontologie budou pro agenta existovat. Někteří vědci, hledající inspiraci pro ontologii ve filozofii, vnímali výpočtovou ontologii jako část aplikované filozofie.

90. léta přinesla snahu vytvořit standardy interoperability. S touto snahou vznikly technologie, které již považovaly ontologii jako standardní komponentu znalostních systémů.

2.1.2. Definice ontologie

Samotným definicím je v této práci věnováno více prostoru, než je pro práci samotnou nezbytně nutné. Důvodem je nedostatek česky psaného textu s touto tematikou, záměrem tedy je více přiblížit čtenáři problematiku ontologie, její význam i využití v informačních technologiích.

Definice ontologie není tak jasná jako v případě jiných termínů z oblasti informatiky. V následujícím textu je předloženo několik definic, které si jsou podobné, ale nejsou totožné. Jediná definice, která se od ostatních liší, je definice ontologie jako filozofické disciplíny. I pro pochopení podstaty ontologie v oblasti informatiky je pochopení ontologie z hlediska filozofie důležité. Některé z definic jsou rozebrány a vysvětleny více do hloubky.

Nejrozšířenější definice:

- Ontologie je teorie existence – definice pro oblast filozofie.
- Ontologie je explicitní specifikace konceptualizace – definice obecně navržená pro aplikaci ontologie v oblasti výzkumu umělé inteligence. V překladu znamená „objekty, koncepty a další entity, jejichž existence je v daném zájmovém prostoru předpokládána, a vztahy, které jsou mezi jednotlivými entitami udržovány (Gruber, 1993) [3].
- Ontologie je teorie slovníků nebo konceptů použitých při tvorbě umělých systémů.

- Ontologie je skupina znalostí popisujících domény.

Další definice ontologie:

- Ontologie je informační konceptuální systém.
- Ontologie je popis formální sémantiky.
- Ontologie je reprezentace konceptuálního systému prostřednictvím logické teorie
 - a) charakterizována specifickými formálními vlastnostmi.
 - b) charakterizována pouze svým specifickým účelem.
- Ontologie je slovník používající logickou teorii.
- Ontologie je mimoúrovňová specifikace logické teorie.

2.1.2.1. ONTOLOGIE V OBLASTI FILOZOFIE

V oblasti filozofie je potřeba vnímat ontologii jako samostatnou filozofickou disciplínu, která se zabývá podstatou reality a její organizací. Pokouší se odpovědět na otázky typu „Co je to existence?“, „jaké objekty mohou existenci vysvětlit?“ atp.

Aristoteles definoval ontologii jako vědu o bytí, o jsoucnu. Ontologie nezkoumá a nevymezuje jednotlivé druhy, typy či bytosti, ale naopak, co jsou vlastnosti a jevy všem bytostem či druhům společné.

Od Aristotelovy ontologie, dnes nazývané obecná ontologie, se liší tzv. formální ontologie. Formální ontologie definuje oblast, která ovlivňuje okolnosti existence konkrétního jevu a individualizuje jeho požadavky a nároky, které danému jevu ze své podstaty vyhovují. Z toho vyplývá, že formální ontologie se více než existencí konkrétních entit zabývá důkladným popisem podob existence, tzn. jejich strukturními vlastnostmi. V praxi může být formální ontologie využita jako teorie odlišností, které mohou být použity nezávisle na stavu světa, tj. řeší rozdíly:

- mezi entitami ve světě (fyzické objekty, události, regiony, ...)
- mezi kategoriemi použitými k modelování světa (koncepty, vlastnosti, kvalita, stav, funkce, část, ...)

Z tohoto pohledu i formální ontologie jako vědecká disciplína může být relevantní ke znalostní reprezentaci světa a přinášet užitek. (*Obitko, 2001, s. 5*) [7].

2.1.2.2. ONTOLOGIE JAKO SPECIFIKACE KONCEPTUALIZACE

Definici poprvé použil Thomas R. Gruber ve své práci „Toward Principles for the Design of Ontologies Used for Knowledge Sharing (1993)“ [4], kde ontologii použil v kontextu sdílení znalostí. Podle Grubera explicitní specifikace konceptualizace znamená: ontologie je popis konceptů a vztahů, které mohou existovat pro agenty či komunity agentů (agent ve významu např. autonomní robot, viz historie ontologie).

Stěžejní význam ontologie je v tom, že pomocí ní lze znalosti sdílet a znovu použít. Ontologie je vlastně specifikace použitá pro tvorbu ontologických vazeb. Ontologická vazba je určitá dohoda zanesená do slovníku, pomocí něhož lze klást dotazy a získávat odpovědi a tím vytvářet výroky. Přičemž ontologická vazba je vytvořena konzistentním způsobem, který respektuje teorii specifikovanou ontologií. Agenti jsou naprogramováni tak, že použitím slovníku vytváří ontologie a ontologie jsou tvořeny tak, že znalosti v nich zanesené mohou být sdíleny i mezi agenty navzájem.

Základní část znalosti je založena na konceptualizaci. Konceptualizaci tvoří objekty, koncepty a další entity, jejichž existence je v dané zájmové oblasti předpokládána, a vztahy, které mezi těmito entitami fungují. Konceptualizace je abstrakce, v podstatě jednoduchý pohled na danou problematiku, který potřebujeme pro různé účely použít. Všechny znalostní báze, systémy založené na znalostní bázi nebo agenti na znalostní úrovni jsou vázány nějakou konceptualizací. Je-li znalost domény reprezentována deklarativním formalismem, sada objektů, která má být reprezentována, se nazývá univerzum diskurzu (neboli celková oblast úvah). Tato sada objektů spolu s popsitelnými vztahy mezi objekty je obsažena ve slovníku, který v podstatě sadu objektů

reprezentuje. Pomocí slovníku program založený na znalostní bázi dané znalosti popisuje.

Z výše zmiňovaného vyplývá, že v oblasti umělé inteligence představuje ontologie programu definování sady popisovaných entit. V takovéto ontologii definice entit asociují názvy entit v univerzu diskurzu (např. třídy, vazby, funkce a další) s textem popisujícím významy názvů a zároveň definice stanovují formální principy, které omezují interpretaci a použití těchto výrazů.

Významově stejnou definici, vyslovil i Borst (1997) [1]: Ontologie se používají jako datový model reprezentující určitou znalost nebo její část. Je to formální a deklarativní reprezentace, která obsahuje definici pojmů a definici vztahů mezi jednotlivými pojmy. Ontologie je slovníkem, který slouží k uchovávání a předávání znalosti týkající se určité problematiky.

2.1.2.3. ONTOLOGIE JAKO REPREZENTATIVNÍ SLOVNÍK

Slovník myšlený touto definicí může být specifikován pro různé domény či témata. Přesněji řečeno, nejedná se o slovník, který by sám od sebe fungoval jako ontologie, ale je konceptualizován tak, že obsah slovníku je určen k vytvoření ontologie. Tento přístup poskytuje velkou výhodu třeba při překladu slovníku z jednoho jazyka do druhého (např. z češtiny do angličtiny), protože vytvářená ontologie se konceptuálně nezmění.

Obitko ve své práci *Ontologies Description and Application* (2001) [7] uvádí k tomuto tématu příklad ontologii z oblasti elektroniky. Je třeba vytvořit ontologii na základě slovníku popisujícího pojmy tranzistor, operační zesilovače, elektrické napětí a vztahy mezi těmito elementy – operační zesilovače jsou druh elektronických zařízení, tranzistory jsou součástí operačních zesilovačů. Je jasné, že tvorbě ontologie a tedy i tvorbě slovníku a samotné konceptualizaci musí předcházet pečlivá analýza druhů entit a vztahů, které v doméně existují.

Navíc lze na ontologii hledět ze dvou úhlů – ontologie může být použita jako sada termínů reprezentujících konkrétní fakta v dané doméně. Nebo může

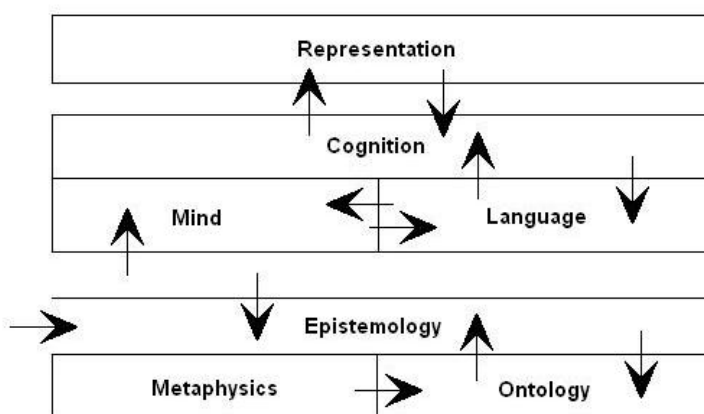
být na ontologii hleděno jako na základní sadu faktů, které jsou mezi všemi entitami v doméně sdílené.

V počátku této kapitoly je zmíněna i definice, kdy ontologie je skupina znalostí popisujících domény. Od ostatních se toto pojetí liší v tom, že ontologie je vnímána jako veškerá „svrchní“ znalostní báze, jejíž součástí je slovník popisující danou znalostní bázi.

2.2. Ontologie v kartografii a gis

V oblasti kartografie se s ontologiemi nesetkáváme tak často. Ontologie jsou mnohem více využívány, a tedy i zkoumány a bádány na poli informatiky, hlavně v oblasti internetu, vyhledávacích a katalogizačních služeb. Postupně používání ontologie proniká také do oblasti geoinformatiky a přes využití GIS se dotýká i aplikací v kartografii, konkrétně při tvorbě či úpravě mapových dat.

Johnathan Raper ve své knize Multidimensional Geographic Information Science (2000) [10] hovoří o ontologii jako o jedné základní náležitosti, kterou by se měl GIS ve světě zabývat a rozvíjet. Ontologie je metodologie, pomocí níž je obsah konceptualizovaného světa definován, uspořádaný a znamenán. Ontologii zahrnuje do systému dalších náležitostí, které jsou zásadní pro budování světového názoru (viz obr. 1)



Obr. 1: budování světového názoru podle Johnathana Rapera

Metody ontologie mají podněcovat definování a klasifikování jevů ve vědeckých disciplínách, a protože většina ontologických systémů používá pojmy vymezené v prostoru a čase, GIS často slouží jako předchůdce ontologie pro návrh nebo odhalení jevů ve vědeckém výzkumu. (Johnathan Raper, 2000).

Na Masarykově univerzitě v Brně vzniklo v posledních letech několik absolventských prací, které se zabývají ontologií v oblasti geografických informačních systémů.

Martin Paška ve své diplomové práci „Geovizualizačná ontológia v krízovom riadení“ (2007) specifikuje ontologii jako vědu, která se snaží vytvořit základ pro vyhledávání a sémantický překlad dat v prostorových informačních strukturách, což zrychlí sběr, analýzu a zpracování dat pocházejících z více zdrojů a majících rozdílnou syntaxi i sémantiku. [9]

Rudolf Šťastný ve své diplomové práci „Podmínky harmonizace map pro krizový management“ (2008) hodnotí možnosti aplikace ontologií při harmonizaci geografických informací. Hlavním bodem zmiňované práce je vytvoření ontologie pro export dat z databáze v případě krizové události (Šťastný, 2008). Autor danou ontologii vytváří z dat potenciálně využitelných hasičským záchranným sborem pomocí open – source software Protégé OWL. [14].

Dalším příkladem je disertační práce Tomáše Řezníka „Metadatový tok v krizovém řízení – od konceptu k implementaci“ (2008). Autor se v ní zabývá mimo jiné ontologií pro podporu prostorových datových infrastruktur, protože ontologie poskytují východisko k řešení interoperability mezi jednotlivými metadatovými standardy (Řezník, 2008) [11].

Ontologií na poli kartografie se zabývá výzkumný projekt „Inteligentní systém pro interaktivní podporu tvorby tematických mapy“, jehož řešitelem je Doc. RNDr. Jaromír Kaňok, CSc. spolu s dalšími pracovníky nejen katedry geoinformatiky Univerzity Palackého v Olomouci. Využití ontologie se diskutuje hlavně v problému zhodnocení teoretických východisek pro využití inteligentních systémů při asistované podpoře tvorby tematických map a

v problému návrhu báze kartografických znalostí použitelných při následném návrhu inteligentního systému. (*Dílčí zpráva za rok 2009*, [22]).

3. METODY A POSTUP ZPRACOVÁNÍ

3.1. Data

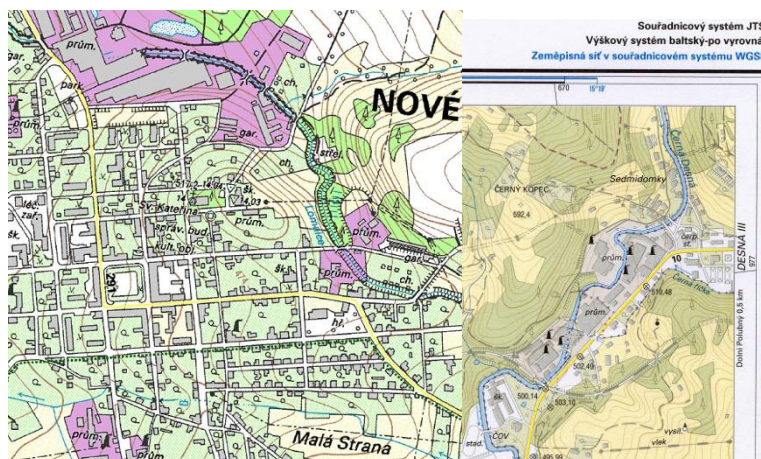
Vzhledem k již obecnosti zadání práce, kdy se hovoří pouze o mapových datech, byla vybrána dvě mapová díla, z nichž byly ontologie extrahovány a byl demonstrován jak jejich význam, tak i možný přístup k samotné extrakci.

3.1.1. Základní mapa 1 : 10 000

Základní mapa České republiky 1 : 10 000 (ZM 10) je základním státním mapovým dílem a je nejpodrobnější základní mapou středního měřítka. Zobrazuje celé území České republiky v souvislém kladu mapových listů, území České republiky je zobrazeno na 4533 mapových listech. Rozměry a označení mapových listů ZM 10 jsou odvozeny z mapového listu Základní mapy České republiky 1 : 50 000, rozděleného na 25 dílů.

ZM 10 obsahuje polohopis, výškopis a popis. Předmětem polohopisu jsou sídla a jednotlivé objekty, komunikace, vodstvo, hranice správních jednotek a katastrálních území (včetně územně technických jednotek), hranice chráněných území, body polohového a výškového bodového pole, porost a povrch půdy. Předmětem výškopisu je terénní reliéf zobrazený vrstevnicemi a terénními stupni. Popis mapy sestává z druhového označení objektů, standardizovaného geografického názvosloví, kót vrstevnic, výškových kót, rámových a mimorámových údajů. Obsahem mapových listů je i rovinná pravoúhlá souřadnicová síť a zeměpisná síť. Předměty obsahu mapy jsou znázorněny pouze na území České republiky. Míra generalizace polohopisu je na takové úrovni, že nedochází k rozsáhlejšímu spojování jednotlivých staveb do bloků a ke zjednodušování tvarů. Mapa tak poskytuje velmi podrobnou představu o zobrazovaném území.

Od roku 2001 se ZM 10 vyhotovuje digitální technologií ze Základní báze geografických dat České republiky (ZABAGED®) a databáze geografických jmen České republiky Geonames. V roce 2006 byla tato nová podoba ZM 10 dokončena pro celé území České republiky a je dále aktualizována. Tvorbu a aktualizaci ZM 10 zajišťuje Zeměměřický úřad. (převzato z <http://geoportal.cuzk.cz> [19])



Obr. 2: Ukázka Základní mapy 1 : 10 000

Kromě extrakce ontologie z obsahu mapového pole byla ontologie vytvářena i ze samotné legendy.

Konkrétně byl použit mapový list č. 03-14-18, jelikož jsou na něm zobrazeny prvky reprezentující krajinu, i komunikace, zástavba, průmyslové i zemědělské objekty.

3.1.2. Vytvářená mapa rozhleden

Pro tvorbu ontologie jako myšlenkové mapy potřebné při návrhu a tvorbě nového mapového díla byl navrhnut imaginární projekt tvorby mapy rozhleden. V tomto případě tedy nešlo o extrakci ontologie z hotového mapového díla, ale naopak pomocí ontologie logicky a uspořádaně shromáždit data a definovat jejich podobu tak, aby výsledná mapa obsahovala všechny důležité prvky a sloužila svému účelu. Vzhledem ke kladenému důrazu pouze na význam a tvorbu ontologií, nevznikla tato mapa fyzicky, nýbrž zůstala jen jakousi

imaginární představou „jak by to mělo vypadat, když se tvoří mapa (v tomto případě mapa rozhleden)“.

3.2. Algoritmus tvorby ontologie

Ze zadání nevyplývá přímo, jaký algoritmus má být při extrakce ontologie z mapových dat použit, ale na základě konzultace s vedoucím práce byla práce zcela zaměřena na tvorbu ontologií na základě teorie analýzy dat FCA (Formal Concept Analysis). A to i z důvodu, že existuje volně dostupný software, pomocí něhož lze tvořit ontologie založené přesně na tomto principu.

3.3. Použitý software

Při extrakci ontologie z mapových dat byl použit open – source software ToscanaJ. Jedná se o volně dostupný nástroj pro analýzu dat na základě teorie analýzy dat FCA (její popis je v následující kapitole). Software je vytvořen v jazyce Java a nahradil starší software Toscana vytvořený pro implementaci jazyka OWL (Web Ontology Language).

Projekt ToscanaJ vznikl spoluprací DSTC, Univerzity v Queenslandu a Technické univerzity v Darmstadtu a je k dispozici na adrese <http://toscanaj.sourceforge.net/>.

Nástroj má čtyři základní komponenty:

- ToscanaJ – komponenta prohlížeč vytvořených schémat.
- Elba – komponenta editor konceptuálních schémat v relačních databázích. Nabízí speciální nástroje jako např. export SQL skriptů.
- Siena – komponenta dosti podobná komponentě Elba (převážně díky sdílenému kódu). Siena edituje konceptuální schémata, která vychází z dat uložených v paměti.
- Lucca – experimentální editor vytvořený za účelem zužitkovat význam analýz SQL vět pro velmi předběžnou a intuitivní tvorbu databázových systémů.

Pro tuto práci byly použity pouze první dvě komponenty.

Při výběru software byly k dispozici v podstatě dvě varianty pro použitý software. Druhým, celkem rozšířeným a také open – source software, který přicházel v úvahu, byl program Protégé OWL Editor vyvíjený na ústavu Stanford Medical Informatics na Univerzitě v Manchesteru a dostupný na adrese <http://protege.stanford.edu>. Po nainstalování obou programů a jejich vyzkoušení a také vzhledem k zaměření celé práce na využití analýzy dat FCA, byl pro tuto práci vybrán program ToscanaJ.

4. POPIS ALGORITMU FCA

FCA (Formal Concept Analysis) je teorie analýzy dat, která identifikuje konceptuální strukturu mezi datovými sadami. Tyto struktury jsou graficky reprezentovány jako konceptuální mřížka dovolující analýzy komplexních struktur a objevení závislostí uvnitř dat. Jedná se o techniku konceptuálního shlukování s dobře vyvinutými matematickými podklady. Tato technika byla úspěšně použita v širokém okruhu aplikací v medicíně, psychologii, knihovnictví, softwarovém inženýrství a ekologii. Zároveň je to technika vhodná pro různé metody analýzy dat, hledání informací a objevování znalostí v databázích. (Obitko, Smid, Snášel, 2004) [8].

Hlavní charakteristiky této metody:

- koncepty jsou popsány vlastnostmi
- vlastnosti ovlivňují hierarchii konceptů, takže hierarchie není explicitně definována tvůrcem
- platí, že jsou-li vlastnosti různých konceptů stejné, pak jsou stejné i koncepty

Těmito třemi body se liší metoda FCA od běžného přístupu k tvorbě ontologií. Běžně totiž začíná tvorba ontologie návrhem hierarchie relevantních konceptů a základem ontologie je systematický vztah a jeho zařazení mezi koncepty. Tento postup pramení z postupu při navrhování objektově orientovaného nebo rámcového systému. Tohoto principu využívá i UML

(Universal Modeling Language). Modelování pomocí tohoto principu začíná návrhem hierarchie tříd či rámců. K těmto hierarchiím se přidávají atributy, přičemž atributy jsou dědičné podél zařazení relace. Tento postup vede ke dvěma zásadním problémům:

- vznikají tendence k vytváření hierarchií objektů s nejasným rozlišením. Pro modelování domény je zaváděna spousta objektů, ty jsou organizovány v systematickém uspořádání, ale nemají žádné jiné rozlišovací atributy. Z toho pramení problémy při sdílení informací.
- není jednoduché změnit třídy a jejich atributy, když už jsou jednou definovány (pouze tzv. redaktorování běžně nabízené v objektově orientovaných jazycích může tento problém odstranit).

A přesně těmto problémům se konstrukce ontologie pomocí analýzy FCA vyhýbá.

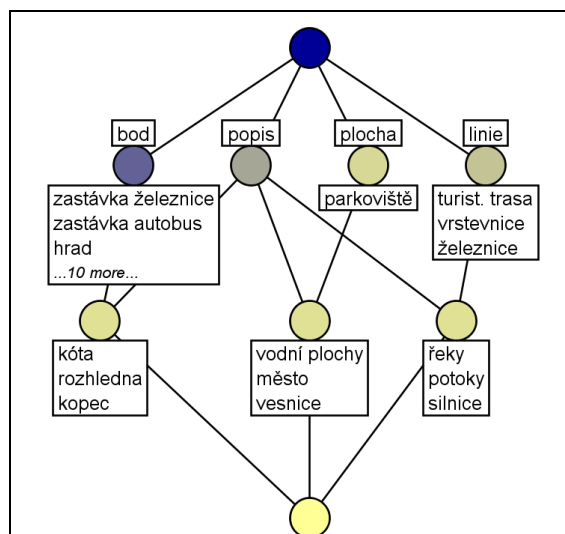
Tvorba ontologie, navrhované pomocí nástroje využívajícího analýzu dat FCA, probíhá v těchto krocích:

- na počátku je prázdná množina pojmů reprezentujících jednotlivé entity a jejich vlastností
- do tabulky pojmů, tzv. kontextové tabulky, se postupně přidávají všechny entity (objekty) a všechny jejich atributy, a to způsobem, kdy každý řádek je jedna entita a každý sloupec reprezentuje jeden atribut. Tabulka se vyplňuje tak, že se v buňce příslušné entity pouze zaznamená, zda entitě daný atribut přísluší, či nikoli. V podstatě se jedná o booleovskou logiku ano/ne. Viz. obrázek 3.

	sítě	hranice	porost, povrch a využití...	terénní reliéf	potépis	bod	linie	polygon
hranice porostu a poz...		x			x		x	
slyška (podtůk) nad 5...			x		x		x	
louka, pastviny			x		x			x
povrchová težba, lom, ...			x		x			x
ovocný sad, zahrada			x		x			x
okrasná zahrada, park			x		x			x
vínice			x		x			x
chmelnice			x		x			x
lesní půda se stromy			x		x			x
lesní půda s křovinami...			x		x			x
lesní půda s kosodřev...			x		x			x
lesní průsek			x		x		x	
osamělý strom			x		x	x		
lesík			x		x	x		
stromeradi, úzký pruh L...			x		x		x	
živý plot			x		x		x	
orná a ostatní půda, úc...			x		x			x
mocál, bažina			x		x			x
různé turistické objekty			x		x			x
různé průmyslové obje...			x		x			x

Obr. 3: Ukázka kontextové tabulky

- Po vyplnění tabulky pojmů již lze ontologii vizualizovat (viz. obr. 4) a s pomocí grafické podoby v ontologii odstraňovat chyby, přidávat či ubírat pojmy i atributy.



Obr. 4: Ukázka diagramu ontologie

- V případě, že dva pojmy (či více pojmů) zůstanou pohromadě, znamená to, že dané pojmy jsou definovány na základě naprosto shodných vlastností. V takovém případě by měly být oba pojmy spojeny do jednoho, nebo je naopak třeba najít další vlastnosti, které pojmy odlišují.

5. POPIS A ZPRACOVÁNÍ VSTUPNÍCH DAT

5.1. Výběr dat

Jak již bylo řečeno v kapitole 3.1 Data, pro extrakci ontologie byla vybrána dvě mapová díla – Základní mapa ČR v měřítku 1 : 10 000 (ZM 10) a modelový příklad tvorby mapy rozhleden.

5.1.1. Data ze ZM 10

Na tomto místě je potřeba přiznat, že zpracovávaný mapový list 03-14-18 nebyl zrovna nejšťastněji zvolen, protože neobsahuje velkou část objektů popsaných v legendě pro mapy ZM 10. Konkrétně na něm úplně chybí železnice. Nakonec ale bylo rozhodnuto, že to není tak velký problém, který by bránil zpracování ontologie z daného mapového listu pro účely této práce.

Pro extrakci ontologie z této mapy byly použity dva různé přístupy:

- Extrakce ontologie z mapového pole na základě toho, jak uživatel mapu vnímá, jaké vztahy a vazby mezi prvky mapy může vyzorovat.
- Extrakce ontologie z legendy na základě znakového klíče a hierarchii a uspořádanosti legendy.

5.1.2. Data z modelového příkladu

V případě simulace tvorby mapy nebyly uvažovány žádné konkrétní datové sady, které by byly pro tvorbu mapy vhodné či potřebné. V podstatě simulace byla založena na zodpovězení otázek „Pro koho bude mapa určena?“, „Co musí na mapě být, aby mapa splnila svůj účel?“, „Co je hlavním tématem?“ atp. Pomocí odpovědí na tyto otázky vznikl seznam všech objektů, které by nová mapa obsahovala.

Následně byly všechny objekty definovány, tedy ke každému byly nalezeny základní vlastnosti a význam, pomocí nichž šlo daný objekt identifikovat.

V další fázi bylo třeba vyřešit, jak důležité místo v mapě bude objekt mít, tedy zda bude hlavním tématem, či zda bude součástí mapového podkladu, a

jak daný objekt bude v mapě zobrazen (jaký mapový znak k němu bude přiřazen, jaké parametry bude znak mít).

5.2. Příprava dat

U všech tří příkladů extrakce ontologie bylo třeba se nejprve vypořádat s tím, z jakého úhlu pohledu bude na data hleděno. Nabízelo se několik možností – třídit data z hlediska grafické interpretace (např. podle parametrů znaku), nebo podle významu reprezentovaného prvku v reálném světě (např. zda se jedná o komunikaci, o průmyslový objekt, o výškopisný údaj atp. Další variantou bylo třídění podle typu znaku (bod, linie, plocha), v případě tematických map se nabízí třídění prvků podle důležitosti v mapě (zda se jedná o hlavní téma nebo prvek mapového podkladu)

Vzhledem k tomu, že nešlo určit, ze kterého směru je správné mapová data rozřazovat a třídit, vznikly pro každou extrakci kategorie a podkategorie, do nichž byly jednotlivé objekty rozřazovány. S takto rozřazenými objekty již bylo možno pracovat v programu ToscanaJ a vytvářet ontologie.

Navíc tím, že byly ontologie vytvořeny z mapových dat postupně definovaných z různých úhlů pohledu, vznikly zajímavé vazby mezi daty a ještě se tak zvětšil prostor pro zkoumání, které vazby jsou v mapových datech skutečně zásadní a co všechno můžou o mapě napovědět.

5.2.1. Mapové pole ZM 10

Při studiu mapového pole byly objekty na mapě tříděny do kategorií podle barvy znaku, typu znaku, obsahu mapy a významu objektů. Skutečný význam znaků nebyl porovnáván s legendou mapy, bylo pracováno pouze s vizuálním dojmem čtenáře s jeho vnímáním obsahu mapového pole.

Extrakce ontologie z již existující mapy sloužila ke kontrole, že s daty je nakládáno správně, že výsledky z analýzy ontologie odpovídají skutečnosti v mapě a skutečnosti v reálném světě. V závěru ale byly srovnány ontologie z mapového pole ZM 10 a z legendy té stejné mapy. Důvodem bylo zkoumání

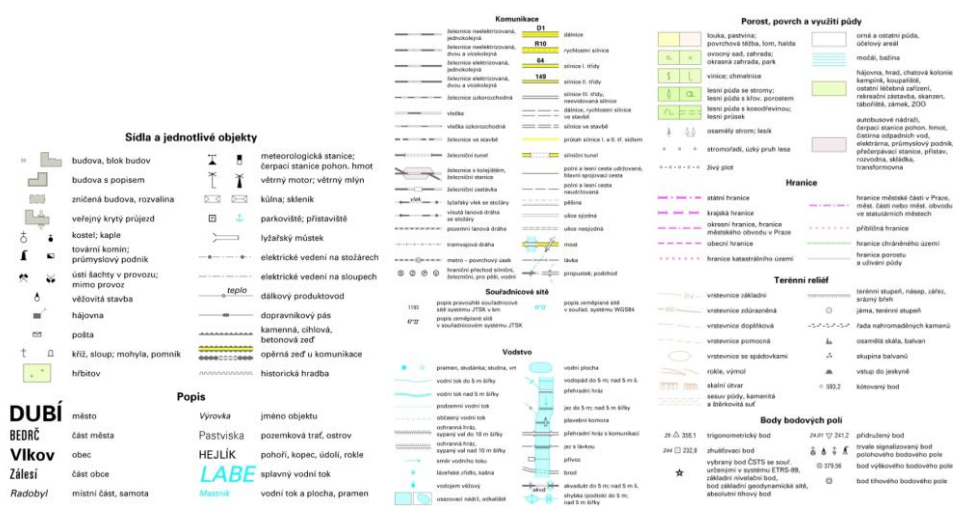
skutečnosti, zda si ontologie odpovídají, když si mají odpovídat obsah mapového pole a obsah legendy.

5.2.2. Legenda ZM 10

V případě hodnocení legendy Základní mapy 1 : 10 000 byla situace nejjednodušší. Nepředpokládalo se, že v tak významném mapovém díle by se vyskytovaly nějaké chyby, popř. by legenda neodpovídala skutečnosti v mapovém poli či by byly porušeny kartografické zásady tvorby správné legendy. Všechny čtyři kategorie, které byly definovány, tedy struktura legendy, obsah mapy, typ znaku a barva znaku, byly v podstatě z legendy převzaty.

Dá se vlastně říci, že třídění objektů uvedených v legendě mapy hodně pomohlo v ujasnění, jakým způsobem data zaznamenávat a jak dále s roztržiděnými objekty pracovat. Zároveň poznatky získané z hodnocení ontologie byly snadno porovnatelné se skutečností a snadněji byly hledány chyby provedené při zpracování dat.

Legenda není znázorněna na každém mapovém listu. Je dostupná na stránkách Českého úřadu zeměměřického a katastrálního ve formátu jpg obrázků. Odkazy jsou uvedeny v použité literatuře [18].



Obr. 5: Ukázka legendy ZM 10

5.2.3. Modelový příklad mapy rozhleden ČR

V případě simulace tvorby nové mapy předcházely úvahy, pomocí nichž se postupně skládal obraz obsahu výsledné mapy.

Celý proces se může zdát v některých svých částech zbytečný, zvláště když si tvůrce pokládá otázky, na něž bezpečně zná odpověď předem. Ale je to nejlepší způsob, jak získat jistotu, že žádná zásadní část nebyla vynechána a že ve výsledné mapě bude skutečně vše potřebné pro účel mapy.

V úvodu bylo řečeno, že ontologie může sloužit i jako myšlenková mapa. Přemýšlení nad nově vznikající mapou vypadalo následovně:

Základní motivace je vytvořit mapu rozhleden.

Přestože jednoznačná definice rozhledny neexistuje, následující popis je dostatečný a obecně přijímaný:

Rozhledna = vyhlídková stavba s vyhlídkovým prostorem vyvýšeným nad okolním terénem, přístupná veřejnosti.

Z daného rčení vyplývá několik zásadních informací:

- vyhlídková stavba se staví proto, že je z ní vidět na zajímavá místa, tedy nějaké *dominanty v okolí* => jeden z důvodů, proč rozhlednu navštívit.
- nad okolním terénem => v mapě by se měl objevit *výškopis*.
- přístupná veřejnosti => tuto veřejnost nejčastěji tvoří *turisté*.

A v úvahách je třeba pokračovat:

- dominanty v okolí – může být všechno možné – *hrady, zámky, jiné rozhledny, významné vrcholy, větší církevní stavby, věže* atp.
- výškopis – *vrstevnice* či jen *výškové kóty*? Bude záležet na měřítku. A je skutečně tak zásadní jen proto, že rozhledny většinou stojí na kopci? Pro turisty to ale může být zásadní.
- turisté – zcela obecně jsou tímto slovem označováni ti, již cestují za poznáním, oddechem, sportem či rekreací. => tzn. cestují vlakem (=> *železnice, železniční stanice a zastávky*), autem (=> *silnice, parkoviště*),

autobusem (=> *silnice, autobusové zastávky*), pěšky (=> *turistické trasy*), na kole (=> *cyklostezky*). Navíc turisté danou lokalitu nemusí znát => prvky pro lepší orientaci => *vodstvo* (=> významnější vodní toky a vodní plochy), *sídla*, další dominanty v okolí. A mohly by se jim hodit informace o občerstvení, infocentrech atd. atd.

Postupem času se nashromáždí takové množství objektů, které by bylo vhodné do mapy zanést, že není těžké se v tom ztratit a zapomenout, že primárně se tvoří mapa rozhleden, tedy že rozhledny by měly být tím prvním, co bude v mapě patrné.

Roztřídění objektů a vytvoření ontologie pomůže udržet v datech „pořádek“, nezapomenout na to nejdůležitější a odhalit případné nesrovnalosti či dokonce hrubé nedokonalosti vznikajícího konceptu budoucí mapy.

V případě této simulace byly objekty roztrženy do tří kategorií – význam objektu, typ znaku a důležitost objektu. Tyto tři kategorie stačily pro úroveň tvorby nové mapy v této práci. Pokud by došlo ke skutečné tvorbě mapy, bylo by potřeba přidat další kategorie týkající se konkrétního vytváření kartografických znaků, použití barvy, velikosti, tloušťky čar, fontu a velikosti písma atp.

6. EXTRAKCE ONTOLOGIE Z DAT

6.1. Tvorba diagramu ontologie

Jak bylo uvedeno v kapitole 4. popis algoritmu FCA, při analýze dat pomocí FCA se definují pojmy a jejich atributy, oboje se zanáší do kontextové tabulky a u každého pojmu je označeno, zda mu daný atribut náleží či nikoli.

Tato část se v programu ToscanaJ provádí pomocí v modulu Elba.

Při zakládání nového diagramu je možné vybrat typ databázového spojení. K dispozici je vložená DBMS, JDBC, zdroj ODBC, soubor Access a soubor

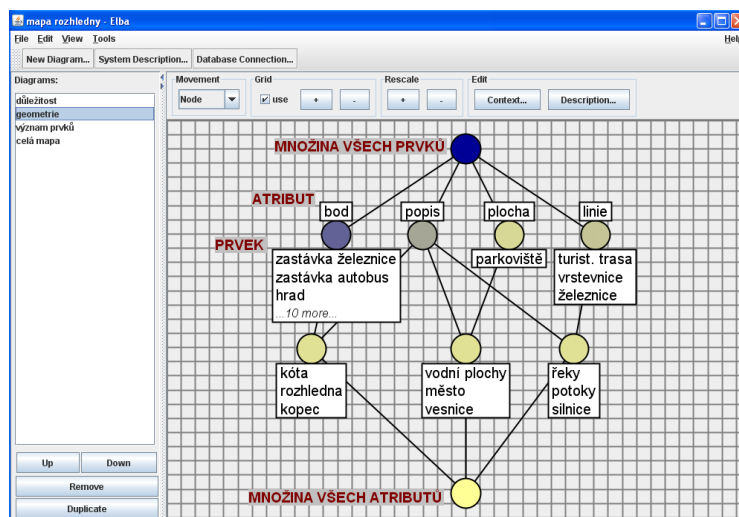
Excel (pouze s příponou .xls, nikoli .xlsx). Také ale nemusí být vybrána žádná databáze v případě, že uživatel chce vytvářet ontologii přímo, jako v případě této práce.

Druhým krokem je vytvoření kontextové tabulky, do níž se zanesou všechny objekty a všechny atributy (viz obr. 3). Během celé tvorby lze objekty i atributy libovolně přidávat a odebírat. Program nedovolí pojmenovat stejně dva atributy ani dva objekty, všechny musí být pojmenovány jedinečně.

Takto vytvořenou tabulku je potřeba vyplnit, tedy v každé buňce zaškrtnout, jestli atribut danému objektu přísluší či nikoli. Pokud ano, stačí buňku dvojklikem myši označit křížkem.

Hotovou tabulku již lze zobrazit do diagramu, pomocí tlačítka Create.

Na základě této tabulky vytvoří nástroj diagram složený z bodů, linií a popisů. Body představují objekty a atributy, přičemž v diagramu nejvýše položený bod reprezentuje množinu všech objektů a nejnižší položený bod množinu všech atributů. Linie propojují objekty a atributy tak, jak k sobě náležejí a jak je uvedeno v kontextové tabulce. Popisy nad body označují atribut, který k danému bodu přísluší, popisy pod body označují objekty příslušející k danému bodu. Pokud nad bodem není uveden žádný atribut, znamená to, že danému bodu přísluší objekty, které mají společný více jak jeden atribut. V takovém případě také k danému bodu vedou směrem od shora dvě a více linií (podle počtu společných atributů). Analogicky k tomuto, pokud pod bodem není uveden žádný objekt, znamená to, že žádný objekt nemá pouze ten atribut, který je nad daným bodem uvedený. Logicky tedy od tohoto bodu vede směrem dolů více jak jedna linie. Zároveň může nastat situace, kdy se v diagramu objeví bod, u něžž není uveden ani žádný objekt, ani žádný atribut. Znamená to, že žádný objekt neodpovídá takové kombinaci atributů, kterou daný bod reprezentuje, tedy každý objekt, který je s tímto bodem spojen, je ještě sám o sobě definován pomocí dalšího atributu. Viz obr. 6.



Obr. 6: Diagram ontologie v modulu Elba

Pro dobrou orientaci ve vytvořeném diagramu je ještě program vybaven nástroji určenými pro grafickou úpravu diagramu. Je možné volit barevné schéma pro výplň bodů, stejně jako možnost měnit velikost bodu podle množství příslušejících objektů. Podle množství spojovaných bodů je také možné měnit sílu linie. V modulu Elba se na pozadí zobrazuje mřížka, ke které se jednotlivé body přichytávají. Tuto mřížku je možno vypnout, ale také lze podle potřeby měnit její rozlišení. Zároveň lze nastavovat měřítko, neboli poměr mezi velikostí bodů a textu ke zvolené mřížce. Popis objektů i atributů lze vypnout.

Po vytvoření diagramu je potřeba nový diagram jednoduše řečeno rozplést a přerovnat, tak aby všechny vazby, objekty a atributy byly jasně čitelné. Pro toto rovnání je připraveno několik režimu ovlivňujících množství bodů, které se přesunou spolu s požadovaným objektem. Lze přesouvat samotný uzel (bod, k němuž linie vedou), nebo daný uzel a všechny uzly na něj napojené a to buď směrem k množině objektů, nebo směrem k množině atributů. K tomu je nabízena také jakási ideální varianta, kdy program sám rozhoduje, které další uzly se spolu s požadovaným uzlem přesunou.

Největší zjištěnou nevýhodou modulu Elba je skutečnost, že se při jakékoli změně v kontextové tabulce diagram vykreslí opět neuspořádaně a je potřeba jej „přerovnat“ znovu.

6.1.1. Vytvořené ontologie

V podstatě tvorbou výše popsaných diagramů se tvoří ontologie, tedy dávají se do souvislostí data – entity – na základě jejich vlastností, popř. na základě znalostí o daných entitách. Diagramy jsou grafickou podobou vzniklých ontologií.

Otázkou zůstává, jak by měla výsledná ontologie, tedy výsledný diagram vypadat, aby byl vypovídající, aby bylo možné na základě ontologie vysledovat závislosti a vazby mezi entitami, a tyto znalosti zužitkovat.

Ve všech třech příkladech (tedy při tvorbě ontologie z mapového pole ZM 10, z legendy ZM 10 a z „mapy rozhleden“) bylo vytvořeno více diagramů. V následujících odstavcích jsou diagramy popsány a přidána ukázka, všechny diagramy jsou přehledně zobrazeny v příloze 2 - 4.

6.1.1.1. Diagramy ZM 10 – mapové pole

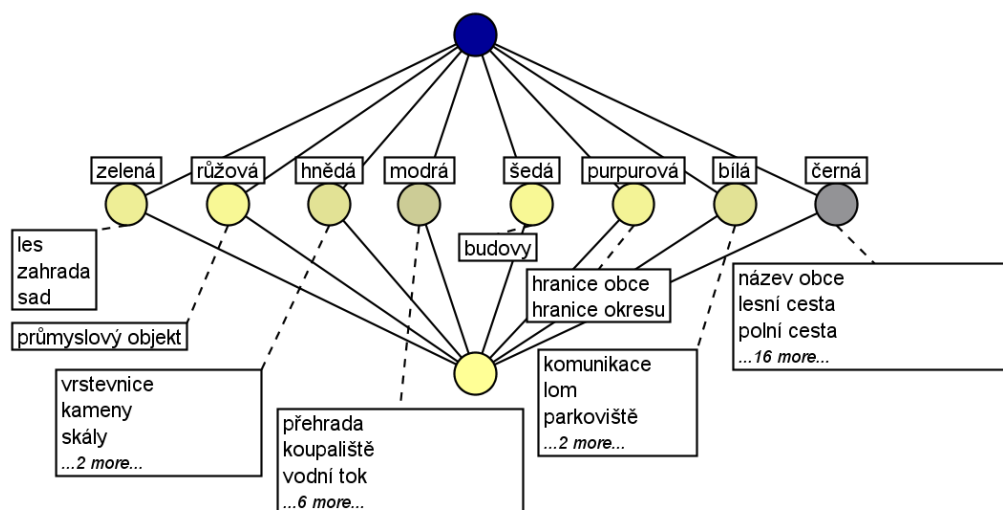
V tomto a následujícím odstavci jsou uvedeny tabulky a diagramy čerpající z jednoho mapového díla. Mapové pole i legenda ale byly zpracovávány samostatně, nebyla mezi nimi navzájem prováděna žádná kontrola. Cílem bylo zjistit nejen to, jak vypadají ontologie extrahované na základě čtení mapového pole a čtení legendy, ale také jak se mezi sebou tyto ontologie liší, jestli mezi nimi panuje shoda, či naopak rozpor. Výsledky tohoto porovnání jsou uvedeny v kapitole 7.

- **diagram „Barva znaku“** – V případě některých objektů bylo těžké určit, jestli jsou znázorněny jednou či více barvami, zvláště když byly znázorněny bílou barvou, tak zda černé okraje brát také jako barvu či nikoli. Nakonec černé okraje nebyly brány jako barva znaku, protože čtenář si při pohledu do mapy řekne, že např. silnice je znázorněna bílou, nikoli bílou linií s černým okrajem.

Tab. 1: ZM 10 – mapové pole – barva znaku

ATRIBUT	PŘÍSLUŠNÉ OBJEKTY
zelená	les; zahrada; sad
černá	název obce; lesní cesta; polní cesta; antropogenní terénní hrana; křížky zeměpisné sítě; symboly využití půdy; výškové kóty; body geodetických sítí; popisy bodů geodetických sítí; křížek, boží muka; lyžařský vlek; skála; vysílač; most, propustek; elektrické vedení; komín; věž; popis areálů; popisy významných míst
bílá	silnice; lom; parkoviště; hřiště; ostatní intravilán
hnědá	vrstevnice; kameny; skály; přírodní terénní hrana; popis vrstevnic
modrá	přehrada; koupaliště; vodní tok; název vodního toku; název přehrady; mokřady; šipka – směr vodního toku; podzemní vodní tok; pramen
růžová	průmyslový objekt
purpurová	hranice obce; hranice okresu
šedá	budovy

Vzhledem k tomu, že každému objektu byl přiřazen jeden atribut, není výsledný diagram nijak složitý:



Obr. 7: ZM 10 – mapové pole – barva znaku

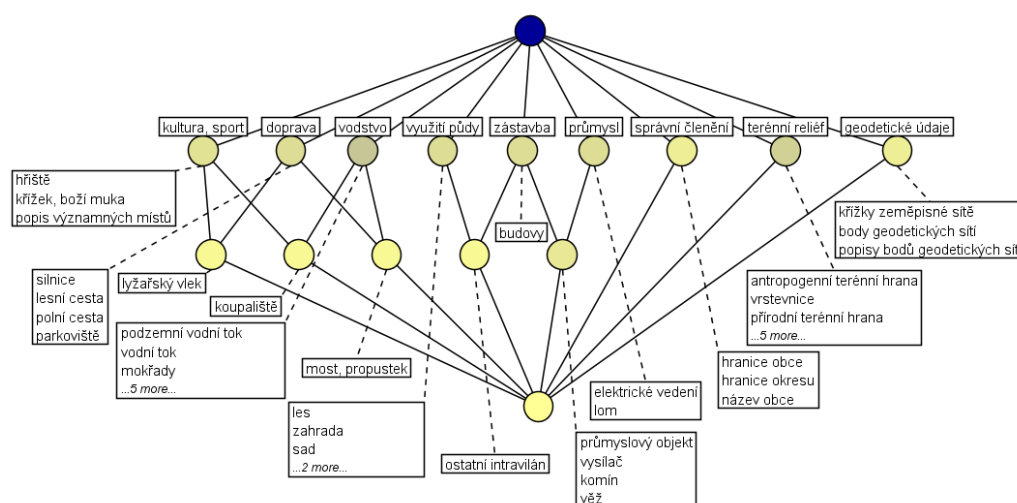
Diagram je též uveden v příloze 2.

- **diagram „Význam objektu“** – V tomto případě byly objekty rozděleny podle atributů definujících jejich význam, v podstatě na základě znalosti o tom, k čemu či v jaké oblasti objekty slouží.

Tabulka 2: ZM 10 – mapové pole – význam objektu

ATRIBUT	PŘÍSLUŠNÉ OBJEKTY
doprava	silnice; lesní cesta; polní cesta; most, propustek; lyžařský vlek; parkoviště,
vodstvo	podzemní vodní tok; přehrada; mokřady; most, propustek; koupaliště; pramen; název vodního toku; název přehrady; šipka – směr vodního toku
správní členění	hranice obce; hranice okresu, název obce
využití půdy	les; zahrada; sad; ostatní intravilán; popisy areálů; symboly využití půdy
terénní reliéf	antropogenní terénní hrana; vrstevnice; přírodní terénní hrana; skála; kameny; skály; výškové kóty; popis vrstevnic
kultura, sport	lyžařský vlek; hřiště; křížky, boží muka; koupaliště
průmysl	elektrické vedení; průmyslový objekt; lom; komín; věž; vysílač
zástavba	budovy; komín; věž; průmyslový objekt; ostatní intravilán; vysílač
geodetické údaje	křížky zeměpisné sítě; body geodetických sítí; popisy bodů geodetických sítí

Zde se již vyskytují objekty, které mají více jak jeden atribut, což je přehledně vidět z diagramu:



Obr. 8: ZM 10 – mapové pole – význam objektu

Diagram je též uveden v příloze 2.

- **diagram „Typ znaku“** – Objekty byly tříděny podle toho, zda jsou vyjádřeny bodem, linií, polygonem nebo zda představují jen popis jiného objektu. K těmto čtyřem přibýly ještě dva atributy – *symbol*, který je přiřazen objektům, které nerepresentují konkrétní entitu, ale pouze doplňují znak, kterým je entita vyjádřena. Atribut *ostatní* nesou pouze křížky zeměpisné sítě, kterým nešlo přiřadit jiný atribut, protože nerepresentují entitu fyzicky přítomnou ve skutečnosti, ale jsou jen součástí mapového listu. Diagram je uveden v příloze 2.

Tab. 2: ZM 10 – mapové pole – typ znaku

ATRIBUT	PŘÍSLUŠNÉ OBJEKTY
bod	budovy; most, propustek; koupaliště; pramen; křížek, boží muka; skála; vysílač; komín; věž; kameny; skály; body geodetických sítí
linie	hranice obce; hranice okresu; podzemní vodní tok; silnice; lesní cesta; polní cesta; antropogenní terénní hrana; lyžařský vlek; elektrické vedení; vrstevnice; přírodní terénní hrana; vodní tok
plocha	mokřady; přehrada; průmyslový objekt; lom; hřiště; parkoviště; les; zahrada; sad; ostatní intravilán
popis	název vodního toku; název přehrady; název obce; výškové kóty; popis bodů geodetických sítí; popis areálů; popis významných míst; popis vrstevnic
symbol	šipka – směr vodního toku; symboly využití půdy
ostatní	křížky zeměpisné sítě

- **diagram „Obsah mapy“** – Obsah ZM 10 převzatý ze stránek ČZÚK je rozdělen na polohopis, výškopis a popis. Stejně byly rozděleny i objekty. Diagram je uveden v příloze 2.

Tab. 3: ZM 10 – mapové pole – obsah mapy

ATRIBUT	PŘÍSLUŠNÉ OBJEKTY
polohopis	hranice obce; hranice okresu; podzemní vodní tok; silnice; lesní cesta; polní cesta; antropogenní terénní hrana; lyžařský vlek; elektrické vedení; vodní tok; křížky zeměpisné sítě; mokřady; přehrada; průmyslový objekt; lom; hřiště; parkoviště; les; zahrada; sad; ostatní intravilán; budovy; most, propustek; koupaliště; pramen; křížek, boží muka; skála; vysílač; komín; věž; kameny; skály; body geodetických sítí; šipka – směr vodního toku; symboly využití půdy
výškopis	vrstevnice; přírodní terénní hrana
popis	název vodního toku; název přehrady; název obce; výškové kóty; popis bodů geodetických sítí; popis areálů; popis významných míst; popis vrstevnic

- **diagram „Celá mapa“** - V tomto případě jsou všechny čtyři kategorie atributů spojeny do jedné kontextové tabulky a vytvořen jeden diagram. Diagram je uveden v příloze 2.

6.1.1.2. Diagramy ZM 10 – legenda

Tabulky a diagramy tvořené z legendy ZM 10 jsou nejobsáhlejší ze všech tří příkladů extrakce ontologie. Důvodem je fakt, že se v nich pracuje se všemi objekty, které se mohou vyskytnout na mapě ZM 10. Samotný vybraný mapový list 03-14-18 není takto obsáhlý a ani model tvorby mapy rozhlédlen neobsahuje tolik objektů.

- **diagram „Struktura legendy“** – Nejjednodušší variantou, jak začít, bylo roztřídit objekty podle toho, jak jsou strukturované v legendě. V podstatě se jedná o pouhé přepsání legendy do kontextové tabulky.

Tab. 4: ZM 10 - legenda - struktura legendy

ATRIBUT	PŘÍSLUŠNÉ OBJEKTY
sídla a jednotlivé objekty	budova, blok budov; budova s popisem; zničená budova, rozvalina; veřejný krytý průjezd; kostel; kaple; tovární komín; průmyslový objekt; ústí šachty v provozu; ústí šachty mimo provoz; věžovitá stavba; hájovna; pošta; kříž, sloup; mohyla, pomník; hřbitov; meteorologická stanice; čerpací stanice PH; větrný motor; větrný mlýn; kůlna; skleník; parkoviště; přístaviště; lyžařský můstek; elektrické vedení na stožáru; elektrické vedení na sloupu; dálkový produktovod; dopravníkový pás; kamenná, cihlová, betonová zeď; opěrná zeď u komunikace; historická stavba
popis	město; část města; obec; část obce; místní část, samota; jméno objektu; pozemková trať, ostrov; pohoří, kopec, údolí, rokle; splavný vodní tok; vodní tok a plocha, pramen
komunikace	železnice neelektrizovaná, jednokolejná; železnice neelektrizovaná, dvou a více kolejná; železnice elektrizovaná, jednokolejná; železnice elektrizovaná, dvou a více kolejná; železnice úzkorozchodná; vlečka; železnice ve stavbě; železniční tunel; železnice s kolejištěm; železniční stanice; železniční zastávka; lyžařský vlek se stožáry; visutá lanová dráha se stožáry; pozemní lanová dráha; tramvajová dráha; metro – povrchový úsek; hraniční přechod silniční; hraniční přechod železniční; hraniční přechod pro pěší; hraniční přechod vodní; dálnice; rychlostní silnice; silnice 1. třídy; silnice 2. třídy; silnice 3. třídy, neevidovaná silnice; dálnice, rychlostní silnice ve stavbě; silnice ve stavbě; průtah silnice 1. a 2. třídy městem; silniční tunel; polní a lesní cesta udržovaná, hlavní spojovací cesta; polní a

	lesní cesta neudržovaná; pěšina; ulice sjízdná, ulice nesjízdná; most; lávka; propustek, podchod
souřadnicové sítě	popis S-JTSK; popis zem. sítě v S-JTSK; popis zem. sítě ve WGS 84
vodstvo	pramen, studánka; studna, vrt; vodní tok do 5m šířky; vodní tok nad 5m šířky; podzemní vodní tok; občasný vodní tok; ochranná hráz, sypaný val do 10m šířky; ochranná hráz, sypaný val nad 10m šířky; směr vodního toku; lázeňské zřídlo, kašna; vodojem věžový; usazovací nádrž; odkaliště; vodní plocha; vodopád do 5m šířky; vodopád nad 5m šířky; přehradní hráz; jez do 5m šířky; jez nad 5m šířky; plavební komora; přehradní hráz s komunikací; jez s lávkou; přívoz; brod; akvadukt do 5m šířky; akvadukt nad 5m šířky; shybka (potok) do 5m šířky; shybka (potok) nad 5m šířky;
porost, povrch a využití půdy	louka, pastvina; povrchová těžba, lom, halda; ovocný sad, zahrada; okrasná zahrada, park; vinice; chmelnice; lesní půda se stromy; lesní půda s křovinatým porostem; lesní půda s kosodřevinou; lesní průsek; osamělý strom; lesík; stromořadí, úzký pruh lesa; živý plot; orná a ostatní půda, účelový areál; močál, bažina; různé turistické objekty; různé průmyslové objekty;
hranice	státní hranice; krajská hranice; okresní hranice, hranice m. o. v Praze; obecní hranice; hranice kat. území; hranice městské části v Praze, části či obvodu ve statutárních městech; přibližná hranice; hranice chráněného území; hranice porostu a používání půdy
terénní reliéf	vrstevnice základní vrstevnice zdůrazněná; vrstevnice doplňková; vrstevnice pomocná; vrstevnice se spádovkami; rokle; výmol; skalní útvar; sesuv půdy, kamenitá a štěrková suť; terénní stupeň, násep, zářez, srázný břeh; jáma, terénní stupeň; řada nahromaděných kamenů; osamělá skála, balvan; vstup do jeskyně
body bodových polí	kótovaný bod; trigonometrický bod; zhušťovací bod; body geodetických základů; přidružený bod; trvale signalizovaný bod polohového bodového pole; bod výškového bodového pole; bod tíhového vodového pole

Každému objektu je přiřazen pouze jeden atribut, takže vznikl i přes množství objektů jednoduchý diagram. Ten je uveden v příloze 3.

Vzhledem k velkému množství objektů není již k následujícím diagramům místo vypisována tabulka, ale všechny tabulky jsou uvedeny v příloze 1 a všechny diagramy jsou uvedeny v příloze 3.

- **diagram „obsah mapy“** – stejně jako v případě extrakce ontologie z mapového pole ZM 10 byly i zde objekty rozlišeny podle toho, zda patří do polohopisu, výškopisu či popisu.

- **diagram „typ znaku“** – v tomto diagramu byly objektům přiděleny atributy *bod*, *linie*, *plocha* a *popis*. Atribut *popis* mají objekty, které mají atribut *popis* i v diagramu „obsah mapy“. Jinou variantou by bylo, kdyby v tabulce nebyly samostatně objekty reprezentující *popis*. Pak by se v tomto diagramu objevil atribut *popis* ve významu „objekt mající *popis*“.
- **diagram „barva znaku“** – Objekty byly definovány barvou černou, hnědou, zelenou, modrou, šedou, žlutou, bílou, růžovou a purpurovou. Stejně jako v případě mapového pole ZM 10 nebyly objektům přidělovány barvy, kterými byly kresleny např. okraje linií, pokud tím ovšem daný objekt nebyl odlišen od jiného objektu.
- **diagram „Celá legenda“** – V tomto případě jsou všechny čtyři kategorie atributů spojeny do jedné kontextové tabulky a vytvořen jeden diagram.

6.1.1.3. Diagramy „Mapa rozhleden“

V této části práce jsou uvedeny diagramy a tabulky, které vznikly na základě prvotního uvažování o obsahu mapy. V následujících kapitolách bude řešena analýza ontologií a v nich budou navrhovány změny podle toho, jaké chyby a nedostatky byly nalezeny. V této části tedy není schválně vůbec řešeno, zda jsou atributy objektům přiděleny správně či nikoli. Všechny diagramy jsou uvedeny v příloze 4.

- **diagram „Důležitost objektu“** – v tomto případě bylo uvažováno, které z objektů jsou ty nejdůležitější, protože ty by na výsledné mapě měly být nejvýraznější a nejméně generalizované. Atributy *zajímavost* a *pro dokreslení* nesou objekty, bez kterých by se mapa „obešla“, byla by čitelná. Ale zároveň jejich přítomnost mapu obohatí o užitečné a pro čtenáře potřebné informace. Mapa jen se stěžejními objekty by sice byla mapa, ale nikdo by si ji nekoupil...

Tab. 5: Mapa rozhleden - důležitost objektu

ATRIBUT	PŘÍSLUŠNÉ OBJEKTY
stěžejní	rozhledna; silnice; turistická trasa; železnice; vrstevnice; kóta; město
zajímavost	hrad; zámek; kostel; kaple; kopec; zastávka železnice; zastávka autobus; parkoviště; kemp; ubytovna; hotel; hospoda; restaurace; infocentrum
pro dokreslení	řeka; potok; vodní plocha; vesnice; samota

- **Diagram „Typ znaku“** – objekty byly roztríděny podle toho, zda budou vyjádřeny bodovým, liniovým nebo plošným znkem, popř. zda budou mít popis. V některých případech není na první zamyšlení jasné, zda bude lepší použít plošný či bodový znak. V takovém případě byla vybrána ta varianta, která se intuitivně zamlouvala více než ta druhá. Následná analýza totiž ukáže, zda by nebylo lepší některým objektům atribut a tedy i použitý znak změnit.

Tab. 6: Mapa rozhleden – typ znaku

ATRIBUT	PŘÍSLUŠNÉ OBJEKTY
bod	rozhledna; kóta; kopec; zastávka železnice; zastávka autobus; hrad; zámek; kostel; kaple; kemp; ubytovna; hotel; hospoda; restaurace; infocentrum; samota
linie	turistická trasa; řeka; potok; silnice; železnice; vrstevnice
plocha	parkoviště; vodní plocha; město; vesnice
popis	kóta; rozhledna; kopec; řeka; potok; silnice; vodní plocha; město; vesnice

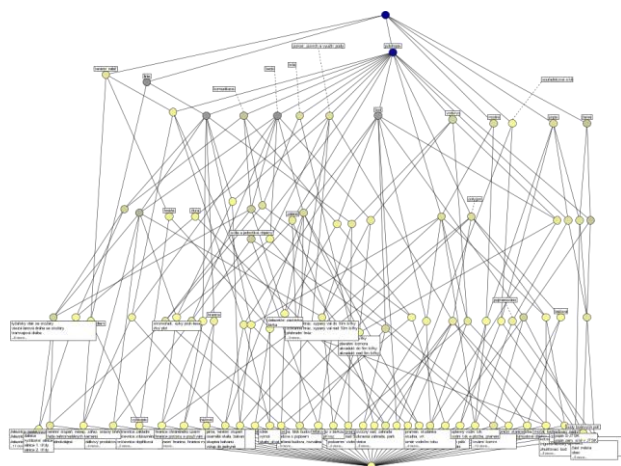
- **Diagram „Význam objektu“** – Tato kategorie je hodně odvislá od prvních uvažování o mapě (viz. kapitola 5.2. Příprava dat), kdy tvůrce přemýšlí o tom, pro koho uvažovaný objekt je, komu slouží, k čemu se využívá, co o něm víme, jak ho známe. Někdy se celá zamýšlená oblast dělí na podřadné objekty, někdy naopak jsou žádané objekty spojovány do společné „významové“ skupiny.

Tab. 7: Mapa rozhleden - význam objektu

ATRIBUT	PŘÍSLUŠNÉ OBJEKTY
pro turisty	rozhledna; parkoviště; turistická trasa; kopec; zastávka železnice; zastávka autobus; hrad; zámek; kostel; kaple; kemp; ubytovna; hotel; hospoda; restaurace; infocentrum
dominanta	rozhledna; kopec; hrad; zámek; kostel; kaple
výškopis	vrstevnice; kóta
vodstvo	řeka; potok; vodní plocha
doprava	parkoviště; turistická trasa; silnice; železnice; zastávka železnice; zastávka autobus
sídla	město; vesnice; samota

- **Diagram „Celá mapa“** – V tomto případě jsou všechny tři kategorie atributů spojeny do jedné kontextové tabulky a vytvořen jeden diagram.

Ve výsledku se ukázalo, že v případě ZM 10 je méně obsáhlých diagramů více vypovídajících než jeden diagram, který obsahuje vše. A to z toho důvodu, že jednak jeden „velký“ diagram obsahuje již i vazby svým způsobem nesmyslné, dává do souvislosti objekty a atributy, mezi nimiž snad již ani nelze souvislosti hledat, a zároveň obsahuje již takové množství objektů, atributů a vazeb, že diagram již nelze přehledně „rozplést“ a uspořádat (viz. obr. 8). V případě ontologie modelu mapy rozhleden se ale tento fakt neprokázal, v tomto případě je i kompletní „velký“ diagram přehledný a dá se s ním pohodlně pracovat.



Obr. 8: ontologie s velkým množstvím objektů a atributů

6.2. Práce s ontologií

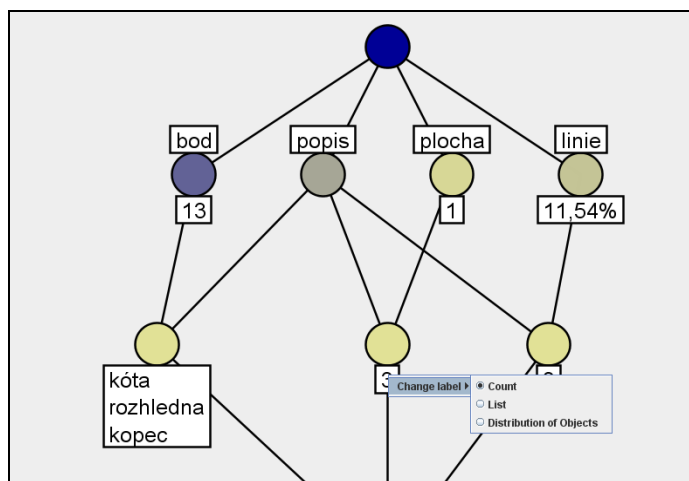
Nástroj ToscanaJ nabízí velkou škálu nástrojů, jak s diagramy pracovat. Zde je jejich funkčnost rozebrána do podrobností a demonstrována na výsledných diagramech.

6.2.1. Modul ToscanaJ

I přesto, že „velký“ digram ontologie mapy ZM 10 je v podstatě nepoužitelný, lze s dílčími diagramy pracovat a získávat stejné hodnotné informace. Pomocí programu ToscanaJ lze totiž pracovat se všemi diagramy týkajícími se jednoho zkoumaného jevu zároveň a lze hledat souvislosti mezi objekty napříč digramy. Možná, že bez této přednosti používaného programu by byl „velký“ diagram ontologie mapy ZM 10 přes své nevýhody lepším řešením, než více dílčích diagramů.

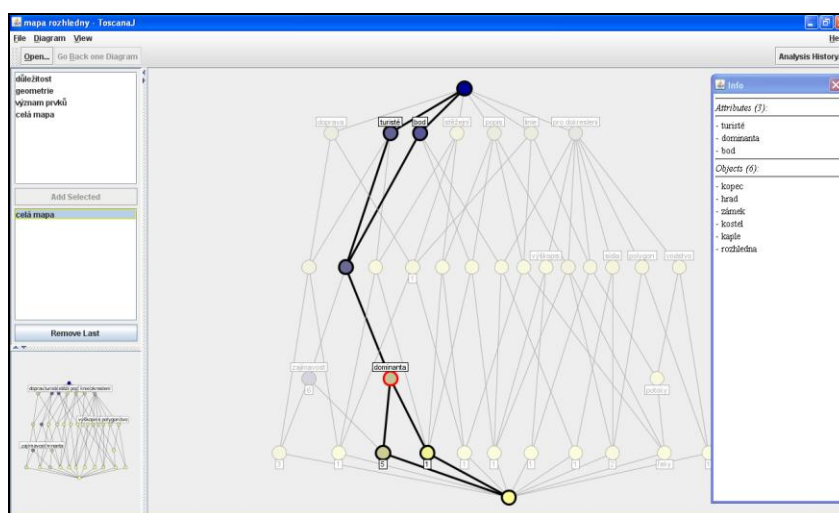
Komponenta ToscanaJ se od komponenty Elba liší v několika zásadních věcech:

- V tomto modulu již nelze diagramy upravovat. A to ani jejich grafickou podobu, ani jejich obsah. Vše je třeba připravit v modulu Elba.
- Zásadní jsou dvě textová a jedno grafické pole v levé části dialogového okna. V prvním z nich jsou vypsány všechny diagramy, které jsou v rámci souboru k dispozici. Pomocí tlačítka „Add Selected“ lze vybraný diagram přesunout do druhého pole. Při kliknutí na diagram v prvním nebo druhém poli se jeho grafická podoba zobrazí ve spodním grafickém poli. Celkové použití všech tří polí je vysvětleno v následující kapitole.
- U zobrazených bodů lze měnit povahu popisu objektu. Buď lze zobrazit názvy příslušejících objektů tak jako v modulu Elba, nebo lze zobrazit pouze jejich počet, nebo jejich procentuální zastoupení (příčemž 100 % jsou všechny objekty). Viz obr. 9.



Obr. 9: Změna popisu objektu v modulu ToscanaJ

- Přes příkaz „View -> Show Concept Information“ lze vyvolat okno Info, v němž se automaticky vypisují atributy a objekty toho prvku, na který je ukázáno myší. (obr. 10)
- V hlavním okně se zobrazuje aktuální diagram a v něm lze pomocí myši zvýraznit jen konkrétní část diagramu a to takovým způsobem, že se zvýrazní ty objekty, atributy a vazby, které souvisí s bodem, na nějž bylo poklepáno myší (obr. 10).
- Velký význam má i volba mezi tzv. *Flat Diagram* a *Nested Diagram* v záložce *Diagram*.



Obr. 10: Práce s diagramem v modulu ToscanaJ

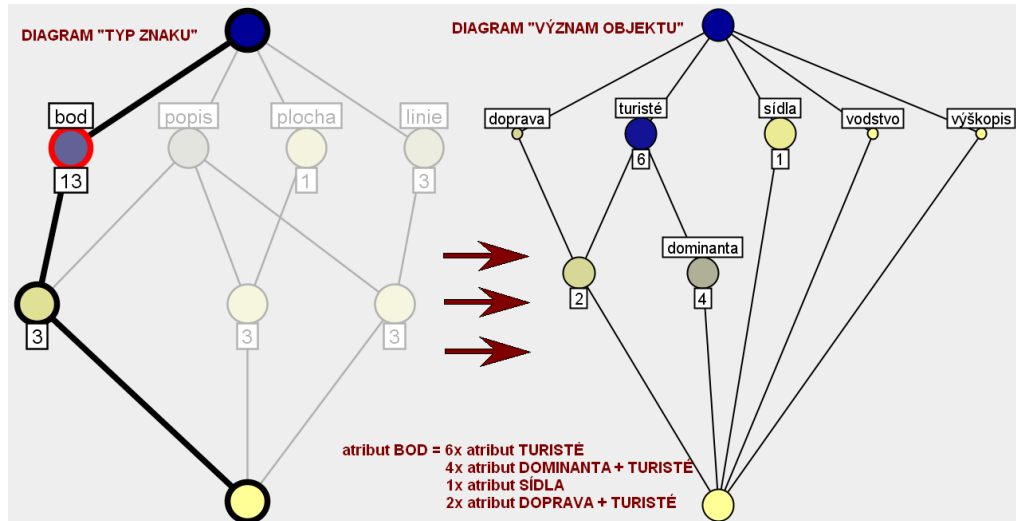
6.2.2. Volba „flat diagram“

Tento odstavec vysvětluje práci s diagramem, pokud je v záložce *Diagram* vybrána volba *Flat Diagram*.

Při spuštění modulu ToscanaJ a otevření konkrétního souboru (v tomto případě souboru *mapa rozhledny.csx*), se nabídnou v levém horním textovém okně čtyři diagramy. Pro zobrazení kteréhokoli z nich je třeba ho tlačítkem *Add Selected* přesunout do spodního textového okna. Pro analýzu dat napříč diagramy je potřeba všechny žádané diagramy přesunout do spodního okna.

Jedním kliknutím myši na konkrétní bod se zvýrazní vazby a body, v jejichž souvislosti je konkrétní bod umístěn v diagramu. Viz obr. 10.

Dvojklikem na konkrétní bod se první diagram přepne na další diagram (podle pořadí ve spodním textovém poli v levé části dialogového okna). V tomto následujícím diagramu jsou pomocí velikosti bodu zvýrazněny jen ty body, které odpovídají atributu, na nějž bylo kliknuto v předešlém diagramu. (viz obr. 11).



Obr. 11: práce s diagramy s volbou "flat diagram"

Jak je z obrázku patrné, atribut „bod“ má 13 objektů, přičemž šest z těchto objektů má i atribut *turisté*, čtyři objekty mají atributy *dominanta* a *turisté*, dva objekty mají atributy *doprava* a *turisté* a jeden objekt má atribut *sídla*.

Diagram, který vznikl dvojklikem na atribut *bod* v „prvním“ diagramu, má jako množinu objektů pouze objekty, které mají pouze atribut *bod*. Tedy všech třináct objektů tvoří množinu 100 %. Množina atributů obsahuje všechny atributy zahrnuté v diagramu významu prvků.

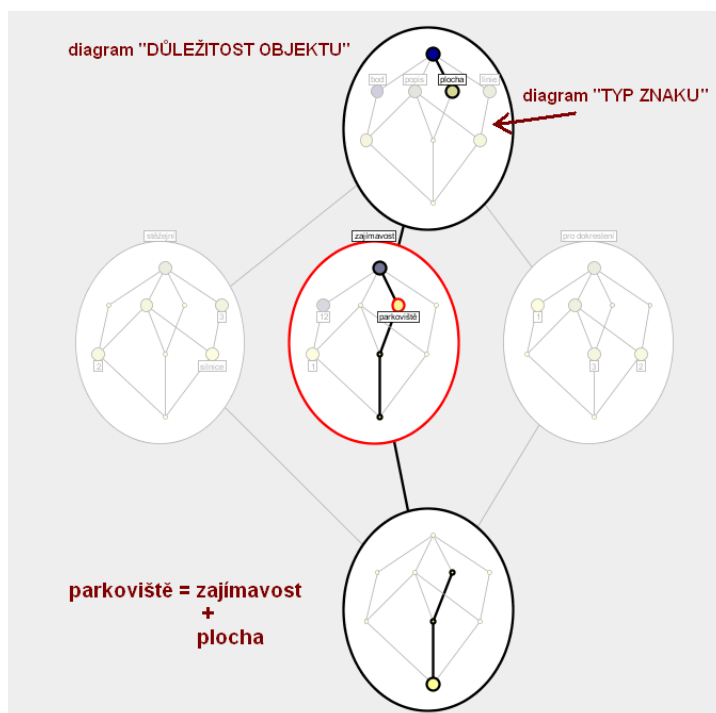
Z pohledu na diagram „typ znaku“ je ale patrné, že existují ještě tři prvky, které mají také atribut *bod*, ale zároveň atribut *popis*. Pro zobrazení i těchto tří prvků ve „druhém“ diagramu je potřeba vybrat volbu „*Filter: use all matches*“ v záložce „*Diagram*“. Pokud bude zaškrtnuta volba „*Filter: only exact matches*“, zobrazí se v druhém diagramu objekty obsahující pouze vybraný atribut (tak jako v příkladu na obrázku 11).

Podobná je i volba „*Show only exact matches*“ a „*Show all matches*“ v záložce „*View*“. Rozdíl je v tom, kolik objektů se zobrazí u každého atributu. V případě první volby se u atributu *turisté* zobrazí pouze šest objektů, tedy těch, které mají pouze atribut *turisté*. Ve druhém případě se u atributu *turisté* objeví dvanáct objektů, tedy k těm původním šesti přibudou objekty, které mají krom atributu *turisté* i atribut *dominanta* popř. atribut *doprava*. První volba je výhodná v tom, že každý objekt je v celém diagramu uveden pouze jednou, kdežto v případě druhé volby se objekty sčítají a opakují – objekt je uveden u každého atributu, který mu náleží. Druhá volba je výhodná v tom, že se při přechodu z jednoho diagramu do druhé neztratí žádný objekt, který má jako jeden ze svých atributů ten požadovaný. Která volba je vybraná, lze zjistit jednoduše podle toho, že v případě druhé volby je u nejvýše položeného bodu (tedy u bodu reprezentujícího všechny objekty) uveden počet těchto objektů (popř. jsou tam vypsány jejich názvy nebo je tam uvedeno 100 %, v závislosti na vybraném popisu objektů).

6.2.3. Volba „nested diagram“

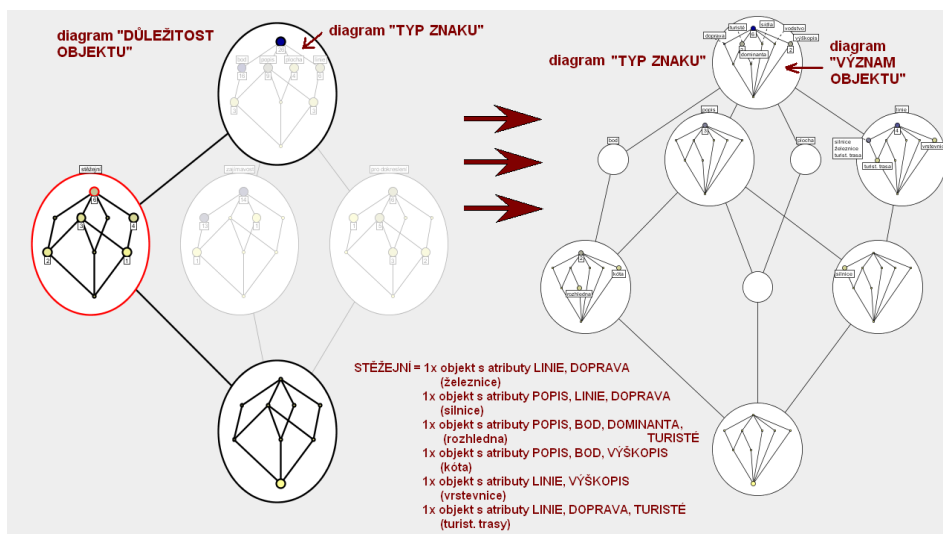
Po zaškrtnutí této volby místo „Flat diagram“ plní program svoji funkci naprosto stejně, jen vizualizace je odlišná.

Druhý diagram se sám jakoby vloží do prvního diagramu a v podstatě je okamžitě vidět souvislost mezi jednotlivými diagramy, tedy jak jsou objekty podle atributů v druhém diagramu rozmístěny v závislosti na atributy z prvního diagramu (viz obr. 12).



Obr. 12: práce s diagramy s volbou „nested diagram“

I v případě tohoto způsobu zobrazení lze na sebe navazovat jednotlivé diagramy (viz. obr. 13). Je však očividné, že pro složité a hodně obsáhlé ontologie je tento způsob zobrazení ne zrovna šťastný, hlavně z důvodu čitelnosti. Samozřejmě je možné takto na sebe navázat více diagramů, než dva.



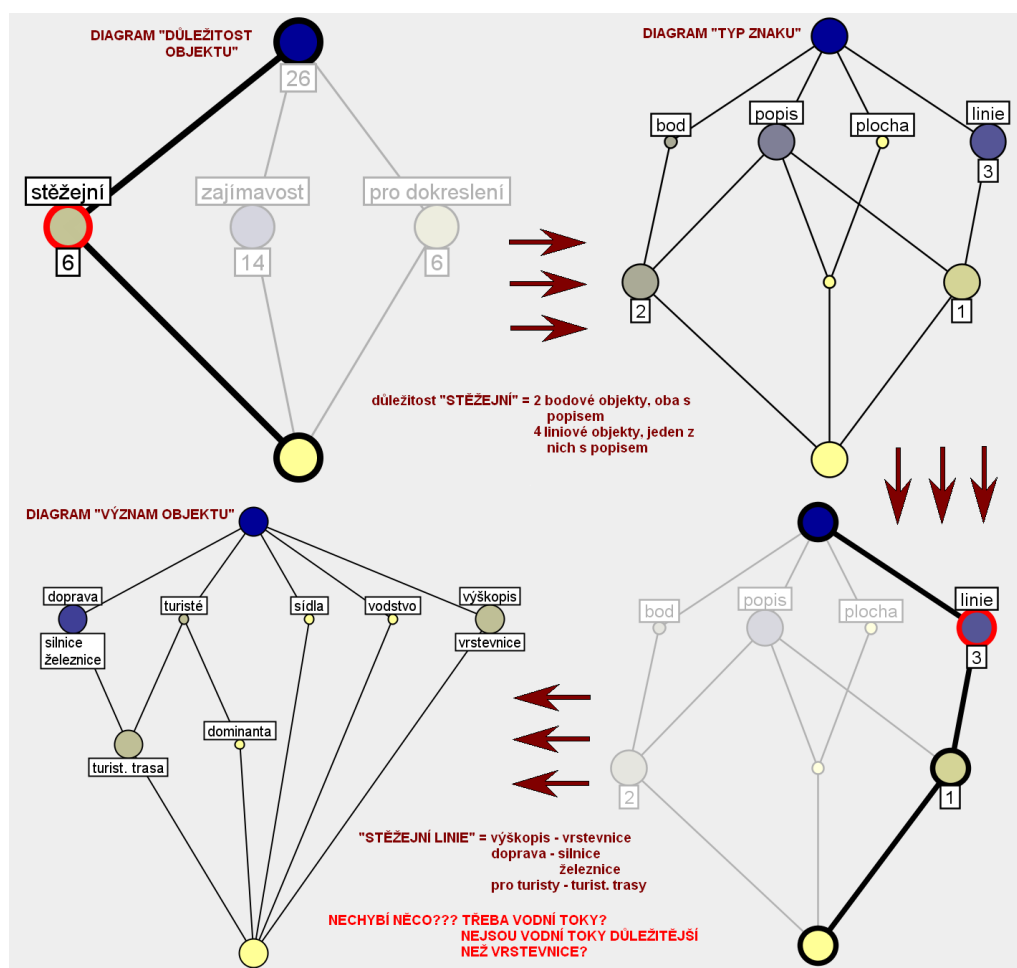
Obr. 13: práce mezi dvěma diagramy

6.3. Analýza diagramů ontologie

V této kapitole jsou popsány způsoby, jak z vytvořených diagramů získat informace užitečné pro tvůrce mapy. U každého způsobu je popsán modelový příklad vztažený k ontologii modelového příkladu tvorby mapy rozhleden.

6.3.1. Hledání špatně přiřazených atributů

Na obrázku 14 jsou pro praktickou ukázkou dány do souvislosti všechny tři dílčí diagramy z modelu tvorby mapy rozhleden. Z dvaceti šesti objektů je šest objektů považováno za stěžejní prvky v budoucí mapě. Z těchto šesti jsou dva objekty bodové i s popisem a čtyři objekty liniové, z nichž jeden bude mít popis. V posledním diagramu „význam prvku“ již byly zobrazeny názvy jednotlivých objektů (což samozřejmě lze udělat v kterémkoli diagramu během celého procesu). Bylo zjištěno, že tři liniové objekty mají atribut *doprava* a jeden má atribut *výškopis*. Zároveň je jasné patrné, že atribut *vodstvo* nemá žádný objekt typu *linie*, který bude v nové mapě stěžejní. A tvůrce mapy se může začít ptát, jestli je to v pořádku, když vodstvo by mělo tvořit kostru mapového podkladu... Samozřejmě nelze tvrdit, že jen díky ontologii přišel autor na chybu ve svém prvotním uvažování, ale rozhodně je to nástroj, který pomůže podobné chyby odhalit.



Obrázek 14 analýza diagramů

Možná by autor odhalil svoji chybu již ve fázi, kdy by dal do souvislosti diagram „důležitost“ a diagram „typ znaku“ a místo počtu objektů by si u daných atributů nechal zobrazit názvy objektů. Ale krokem, kdy byl přidán diagram „význam prvku“, naprosto viditelně bylo ukázáno, že chybí stěžejní prvek s atributem vodstvo. Takto mohly být vizualizovány objekty s jinými libovolnými kombinacemi atributů.

6.3.2. Změny v kontextové tabulce

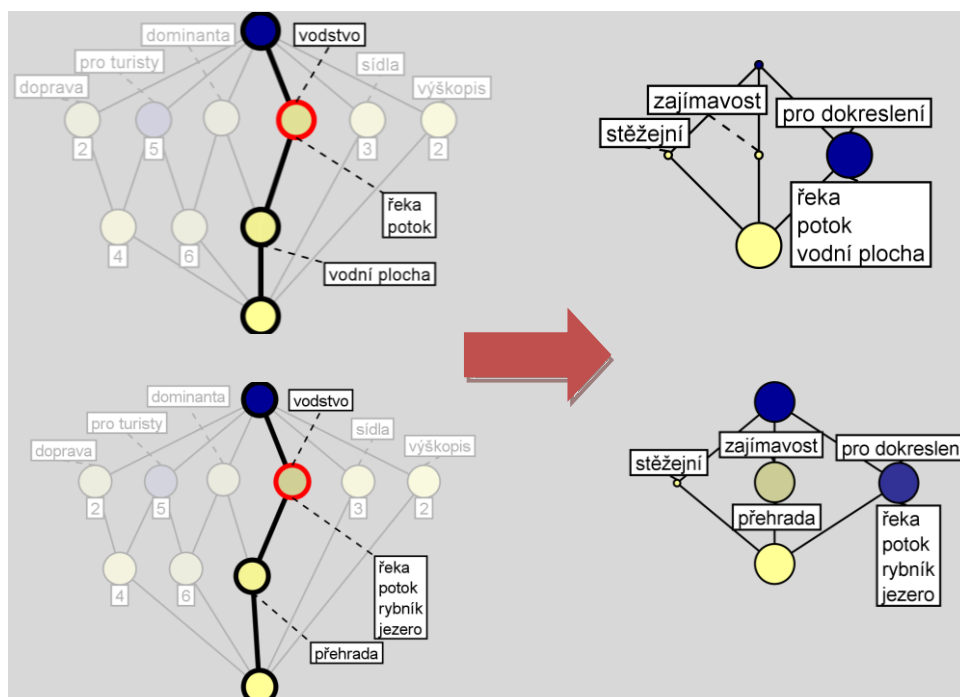
I při analýzách vzniklých diagramů je možno provádět změny v kontextové tabulce a zkoumat jejich význam. Jen pro připomínku – všechny změny je potřeba provádět v modulu Elba, analýzy v modulu ToscanaJ.

Jak se změny projeví a jaký mohou mít význam, je nastíněno v následujícím odstavci.

Úvahy tvůrce dovedly k tomu, že mezi *dominanty* (tedy objekty, které jsou zajímavé při pohledu z rozhledny) patří i *přehrady*. Doposud ale tvůrce pracoval pouze s objektem *vodní plochy*.

Tedy přichází dotaz, zda všechny vodní plochy budou dominanty, nebo zda jen velká vodní díla. Protože pokud by to měla být jen velká vodní díla (např. přehrady), bude potřeba objekt *vodní plochy* rozdělit např. na *přehrady*, *rybníky*, *jezera*. Takže jeden objekt zanikne, ale vzniknou tři nové. Je patrné, že všechny budou vyjádřeny plošným znakem. Dominantou bude jen objekt *přehrada*. Význam objektů se nemění, všechny budou mít atribut *vodstvo*. Zbývá tedy rozhodnout o důležitosti jednotlivých objektů. Objekt *vodní plochy* měl doposud atribut *pro dokreslení*, tedy spadal mezi ty nejméně důležité objekty. Když byl objekt *přehrada* zahrnut mezi *dominanty*, měl by také mít atribut *zajímavost*. Zbylé dva objekty mohou zůstat s atributem *pro dokreslení*. Zároveň na atribut *dominanta* se váže atribut *pro turisty*. Všechny změny jsou patrné na obrázku 15. V podstatě se jen změní počty objektů u některých atributů, protože nebyl přidán žádný objekt, který by měl oproti ostatním originální kombinaci atributů.

Jedna změna na sebe váže další změny. Pokud se nejedná pouze o přejmenování jednoho atributu v jedné kontextové tabulce, zasáhne změna všechny ontologie daného jevu.



Obr. 15: projev změny objektu či atributu

6.3.3. Stejně definované objekty

Jak bylo řečeno v kapitole 4. Popis algoritmu, pokud v diagramu u některého z bodů zůstane více objektů, znamená to, že dané objekty jsou definovány naprosto stejně a buď by měly být sloučeny do jednoho, nebo by měly být popsány dalším – rozdílným – atributem. Tato situace nastala hned několikrát ve všech třech případech. Jakmile taková situace nastane, je potřeba začít uvažovat, jestli a proč je žádoucí takto stejně definované objekty zanechat jako samostatné objekty, nebo je sloučit do jednoho.

Např. v ontologii celé mapy rozhleden „zbyly“ dva objekty s atributy *popis*, *pro dokreslení*, *sídla*, *polygon*. Jedná se o objekty „vesnice“ a město“. Pro dodržení výše zmiňovaného pravidla je potřeba učinit jednu z následujících možností:

- přidat nějaký atribut, který objekty odliší – např. typ či velikost použitého písma. Ale v takovém případě by do objektové tabulky přibyl

atribut, který by měl být vyplněn u všech objektů. Tuto variantu tedy nelze volit prvoplánově pouze pro vyřešení problému se dvěma shodně definovanými objekty.

- změnit některý atribut u jednoho z objektů – např. u objektu *vesnice* zaměnit atribut *plocha* za *bod*, nebo u objektu *město* zachovat popis, ale u objektu *vesnice* nikoli. Nabízí se samozřejmě další i varianty.
- sloučit oba objekty do jednoho objektu – vzhledem k tomu, že žádný jiný objekt nemá atribut *sídla*, bylo by možno oba objekty nahradit právě nadřazeným objektem *sídla*.

Další konkrétní příklad lze opět uvést na ontologiích mapy rozhleden. V diagramu „celá mapa“ je patrné, že celkem pět objektů má atributy *pro turisty*, *bod*, *zajímavost* a *dominanta*. Konkrétně se jedná o objekty *hrad*, *zámek*, *kostel*, *kaple*, *kopec*. V takovém případě je potřeba zvážit, zda skutečně nechat objekty samostatně, nebo zda je sloučit do jednoho objektu. Významově samozřejmě nejde o totožné objekty, ale třeba by pro účel mapy stačilo, aby např. *kostel* a *kaple* byly reprezentovány jedním znakem vyjadřujícím *církevní stavby*. Nebo samozřejmě lze zvolit variantu nenahrazovat dva objekty jedním označením, ale oddělit je od sebe čárkou a vyjádřit jedním znakem, takže by v mapě byl znak pro objekt *hrad*, *zámek* a znak pro objekt *kostel*, *kaple*.

6.3.4. Vybáženost použitých znaků

Díky dělení objektů podle typu znaku lze analyzovat, zda jeden typ znaku nebude na mapě zastoupen nepoměrně více než jiný a mapa nebude přetížena.

Konkrétně je tato úvaha opět demonstrována na ontologii mapy rozhleden. Při pohledu na diagram „typ znaku“, je vidět, že z celkových dvaceti šesti objektů bude 17 objektů vyjádřeno bodovým znakem, 3 objekty plošným znakem, 6 objektů liniovým znakem, k tomu 9 objektů bude mít popis. Hrozí

tedy, že výsledná mapa bude přeplněná bodovými znaky, navíc každému bodovému znaku bude třeba vymyslet konkrétní grafickou podobu.

Daný problém by šel řešit minimálně dvěma způsoby:

- Jednou variantou je přehodnotit některé objekty a místo bodového znaku použít pro ně plošný znak. V případě této sady objektů se žádný takový vhodný objekt nevyskytuje. Naopak by stálo za úvahu, zda bude reálné *vesnice* zobrazovat plošným a ne bodovým znakem.
- Druhou variantou je zmenšit množství bodových znaků způsobem popsáním v předchozí podkapitole.
- A další variantou může být přidání nových objektů, které budou reprezentovány plošným nebo liniovým znakem. Samozřejmě s touto variantou je třeba zacházet opatrně, aby pro změnu mapa nebyla přetížena a neobsahovala nadbytečné množství znaků. Ale v modelovém příkladu by např. mohl být přidán objekt *les* – jednalo by se o plošný znak a zároveň by mohl být i dobrým orientačním prvkem pro uživatele mapy. V podstatě se dá říct, že objekt *les* na mapě doposud chyběl, že to nebude objekt jen pro vyplnění volného místa.

7. VÝSLEDKY A JEJICH VYUŽITÍ V PRAXI

Výsledkem práce je mimo jiné soubor diagramů, které reprezentují ontologie extrahované z obsahu dvou mapových děl – jednoho skutečného, státního mapového díla Základní mapy v měřítku 1 : 10 000, a jednoho imaginárního, který reprezentuje model tvorby nové mapy rozhleden. Všechny vytvořené diagramy ontologií jsou v přílohách 2 – 4.

Hlavním cílem ale bylo zhodnotit možnosti, které skýtá používání a tvorba ontologie v oblasti kartografie a mapových dat.

7.1. Srovnání ontologií legendy a mapového pole ZM 10

Obě ontologie byly tvořeny z jedné mapy a jedním čtenářem mapy. Zároveň během extrakce ontologie nebyla mezi oběma částmi mapy prováděna žádná kontrola. Tento přístup byl snadný, protože mapové listy ZM 10 neobsahují legendu každý zvlášť, ale existuje jedna legenda pro celé mapové dílo. Z toho ale také vyplývá, že neexistuje legenda, která by obsahovala pouze ty entity, které obsahuje mapové pole vybraného mapového listu. Nelze tedy vyloučit, že v ontologii extrahované z mapového pole ZM 10 některé entity nechybí.

Při srovnání obou ontologií byly zjištěny tyto skutečnosti:

- ontologie mapového pole je mnohem jednodušší a obsahuje mnohem méně objektů než ontologie legendy. To je dáno tím, že mapové pole obsahuje méně objektů než celá legenda ZM 10. Další příčina je vysvětlena v následujícím odstavci.
- objekty z mapového pole byly čtenářem mapy popsány jednodušeji a obecněji, než jsou uvedeny v legendě. Příkladem můžou být *body geodetických sítí* v ontologii mapového pole a rozčlenění bodů bodových polí na *kótovaný bod*, *trigonometrický bod*, *zhušťovací bod*, *body geodetických základů*, *přidružený bod*, *trvale signalizovaný bod*, *polohového bodového pole*, *bod výškového bodového pole*, *bod*

tíhového bodového pole. Dalším příkladem jsou vrstevnice, které čtenář v mapě vnímá pouze jako *vrstevnice*, kdežto v legendě jsou rozděleny na *vrstevnice základní, zdůrazněná, doplňková, pomocná, se spádovkami*. Tato neshoda je ale pochopitelný a rozhodně nevypovídá o tom, že by si ontologie navzájem neodpovídaly.

- při extrakci objektů z mapového pole byly popisy brány jako samostatné objekty, stejně jako tomu je v legendě. Rozdíl je pouze v tom, že z mapového pole byly jako samostatné objekty extrahovány i popisy vrstevnic, bodových polí, výškových kót, a různých průmyslových areálů.
- barvy znaků byly vnímány podobně v mapovém poli i v legendě. Jediný vážnější rozdíl je rozlišení šedé a černé v legendě i v případě objektů, které v mapovém poli byly považovány za černé. Typy znaků byly vnímány a objektům přiřazeny stejně v obou případech.
- v legendě nejsou uvedeny popisy průmyslových a jiných objektů, které jsou v mapě popsány.

Celkový výsledek porovnání obou ontologií je takový, že ontologie si neodporují, nebyl v nich shledán žádný zásadní rozpor. Zároveň ale poukazuje na to, že čtenář není schopen bez legendy mapu číst tak důkladně jako s legendou, nerozlišuje a nečlení významově podobné objekty, naopak tyto objekty při čtení považuje za stejné a označuje je nadřazenými výrazy. V některých případech bez legendy nebo bez předešlé znalosti není čtenář schopen správně určit konkrétní význam objektu. Porovnání ontologií v podstatě potvrdilo fakt, že legenda je nedílnou součástí mapy a je nezbytná pro správné a úplné čtení mapy.

7.2. Změny v mapě rozhlédn na základě analýz

Výsledky všech provedených analýz jsou uvedeny vždy u dané analýzy, tedy v podkapitolách 6.3.1 – 6.3.4.

Ontologie vytvářejí struktury mezi entitami na základě definovaných vlastností a vztahů mezi nimi a pomáhají tak odhalit závislosti a souvislosti mezi entitami. Toho lze v kartografii využít několika způsoby:

- Při definování typu znaku, kterými budou jednotlivé objekty znázorněny, lze pomocí ontologie snadno určit, zda některý z typů znaku nepoměrně nepřevažuje nad ostatními znaky. Pak by celá mapa byla nevyvážená. Tvorbou ontologie lze tuto případnou závadu odhalit v počátku tvorby mapy, nikoli až po pracném vytvoření všech znaků a jejich vykreslení do mapového pole.
- Při definování významu objektu se v podstatě dává základ uspořádání legendy, protože objekty jsou rozděleny do nadřazených kategorií se společným významem. I v této fázi ontologie pomáhá odhalit problematické objekty, které např. mohou patřit do více kategorií.
- Pro tvorbu tematické mapy je nezbytné, aby ve výsledku samotné téma mapy bylo skutečně to nejdůležitější, na první pohled patrné a jasné. Tvorbou ontologie si autor pomáhá udržet v používaných objektech pořádek, má stále kontrolu nad tím, zda objektů již v mapě nebude moc, nebo zda nebude hlavní téma zastíněno některým méně důležitým objektem. Snadněji se kontroluje, zda objekty mapového podkladu jsou skutečně jen v mapovém podkladu. Kromě uspořádanosti legendy tedy ontologie napomáhá i správně hierarchizaci legendy.
- Definováním objektů lze zamezit zbytečné nebo naopak nedostatečné generalizaci reálného světa. Na základě výsledných ontologií je autor nucen přemýšlet nad tím, zda objekty potřebuje mít na mapě samostatně znázoněné, nebo naopak zda bude lepší sloučit je do jednoho nadřazeného objektu.

7.3. Přínosy nástroje ToscanaJ

Největší výhodou software ToscanaJ je jeho dostupnost, protože se jedná o open – source software volně stažitelný z internetu.

Nástroj je velmi jednoduše použitelný, není nutné, aby uživatel nejprve vytvořil databázi v některém jiném programu a s touto databází pak vstupoval do modulu Elba. Na druhou stranu i tato varianta je možná, do programu je možné načíst data z několika různých databází.

Tvorba samotných ontologií je založena na booleovské logice, zda daná entita (v programu používán výraz objekt) má či nemá danou vlastnost (v programu používán výraz atribut). Na základě takto přiřazených atributů program sám vykreslí strukturu vložených dat. Výsledný diagram znázorňuje všechny objekty a tributy a propojení mezi nimi a reprezentuje ontologii dané datové sady.

Výsledné diagramy lze exportovat do několika grafických formátů.

8. DISKUSE

Výsledky práce ukazují, že ontologie mají v kartografii své místo. Je to užitečný nástroj, jak jednoduše a rychle třídit a hledat konkrétní data bez ztráty informace o vazbách a vztazích mezi daty. Vše je usnadněno grafickou podobou ontologie. Nejedná se pouze o databázi, ve které uživatel vidí svá data jen ve formě tabulek. Předností ontologie je právě podoba diagramů, kde jsou vazby a souvislosti vidět na první pohled. Programy pro tvorbu ontologie navíc obsahují nástroje pro zvýraznění požadovaných vazeb, atributů a objektů.

Součástí této práce je tvorba ontologie zvlášť z mapového pole a zvlášť z legendy Základní mapy ČR v měřítku 1 : 10 000. Bylo zjištěno, že pro kartografii nemá tento přístup valného využití, pouze možná jako kontrolní nástroj pro zkoumání kvality legendy a samotné kresby v již hotových mapách. Mnohem přínosnější práce nad již existující mapou určitě bude tvorba ontologie z objektů, které jsou vyčteny z mapového pole, ale jejich konkrétní význam je podepřen a potvrzen výkladem daného znaku v legendě. Protože legenda je v mapě přesně z důvodu lepší a správné čitelnosti mapy, bez legendy se čtenář mnohdy jen dohaduje o skutečném významu daného znaku.

V této práci se jistě nepodařilo postihnout všechny oblasti, v nichž může být používání ontologie pro tvorbu map přínosná. Řeší se zde vyváženost bodových, liniových a plošných znaků, správně uspořádaná legenda či dostatečná generalizace objektů. Vzhledem k tomu, že tvorba nové mapy v této práci ustrnula ve fázi modelu tvorby mapy, nejsou zde již řešeny možnosti využití ontologie při samotném návrhu znakového klíče a jeho konkrétní podoby. Pak autor může ontologii využít při řešení tvaru, velikosti, barvy znaku, při hledání správného fontu a velikosti písma pro dané popisy atp.

Dalším využitím ontologie může být práce s topologií, která zde není vůbec řešena. Jednalo by se tedy o prostorové vztahy a vazby, ne pouze o vlastnosti kartografických znaků a význam objektů v reálném světě. Příkladem může být

uvažování o vodním toku – v této práci byl vodní tok uvažován jako linie kreslená modrou barvou, s popisem, významově součástí vodstva. Stejně tak vodní plocha byla plocha kreslená modrou barvou, s popisem, významově součástí vodstva. Ale již nebylo řešeno, že vodní plocha musí vždy ležet na vodním toku. Popř. že vodní tok vždy teče „z kopce dolů“, čemuž by měly odpovídat vrstevnice znázorňující terénní reliéf. Takto definované vazby a vztahy vytvoří omezení, která musí být nezbytně dodržena při tvorbě mapy. Konkrétně ontologie pracující s topologií objektů může mít velký přínos nejen pro oblast kartografie, ale i pro oblast geografických informačních systémů, které pracují s prostorovými daty.

9. ZÁVĚR

Diplomová práce mapová ontologie se zabývá extrakcí ontologie z mapových dat, konkrétně z mapového pole a legendy Základní mapy ČR v měřítku 1 : 10 000 a z modelového příkladu tvorby nové mapy rozhleden.

Práce se ve své praktické části zabývá nejen tvorbou ontologií z konkrétních map, ale zároveň hodnotí přínosy použití ontologií v oblasti kartografie a vhodnost použitého software pro tvorbu ontologií.

Pro extrakci ontologie byla použita analýza dat FCA (Formal Concept Analysis) a open - source software ToscanaJ, který je založený na FCA. Použití tohoto software se ukázalo jako velmi vhodné řešení, protože nástroj je jednoduše a intuitivně ovladatelný, výsledky poskytuje rychle a přehledně. Zároveň obsahuje dobré funkce pro tvorbu, správu a editaci ontologií. Tento nástroj může být dobrou a jednoduchou pomůckou pro každého, kdo při tvorbě mapy chce využít výhody práce s ontologiemi.

Použití ontologií může mít pro každého tvůrce kartografických děl velký význam. Díky grafické vizualizaci všech souvislostí mezi objekty budoucí mapy lze lépe odhalovat chyby v uvažování a v plánovaném použití kartografických znaků, vyjadřovacích prostředků i v použití konkrétních datových sad či dat. Analýza ontologie může pomoci odhalit nevyváženost mapy vlivem použití nepoměrně většího množství jednoho typu znaku než ostatních. Lze z ní hodnotit dostatečnost generalizace a podrobnost objektů. V případě tematických map pomáhá kontrolovat, zda téma mapy skutečně bude stěžejní, nebo zda se již neztrácí ve spoustě dalších objektů, které mají být do mapy zaneseny.

Výsledkem je sada diagramů reprezentujících ontologie extrahované z vybraných map. Diagramy jsou uvedeny v přílohách práce ve formátu jpg. Zároveň jsou na přiloženém CD, odkud je možno s nimi bez omezení v

programu ToscanaJ pracovat. Program ToscanaJ je také přiložen. Součástí příloh je i poster dokumentující jednotlivé fáze práce.

10. LITERATURA

- [1] BORST, W. N.: *Construction of engineering ontologies for knowledge sharing and reuse*. [Ph.D. thesis]. University of Twente, Enschede, 1997, 243 s.
- [2] GRUBER, T.: *Ontology*. In *Encyclopedia of Database System*. 2009. [cit. 2009-11-09]. Dostupné z <<http://tomgruber.org/writing/ontology-definition-2007.html>>
- [3] GRUBER, T.: *A Translation Approach to Portable Ontology Specifications*. In *Knowledge Acquisition*. Knowledge Systems Laboratory, 1993, s. 199 – 220
- [4] GRUBER, T.: *Toward Principles for the Design of Ontologies Used for Knowledge*. Kluwer Academic Publisher, 1993
- [5] HOŘEJŠÍ, M.: *Systém pro vytváření ontologií*. [Bakalářská práce] Masarykova univerzita v Brně, Fakulta informatiky, 2007, 24 s.
- [6] MORAVEC, P., OBITKO, M., SMID, J., SNÁŠEL, V.: *WordNet Ontology Based Model for Information Retrieval*. The 2004 International Conference on Communications in Computing CIC 2004/PSMP: 21. – 24. dubna 2004, Las Vegas, Nevada, USA, CSREA Press ISBN 1-932415-36-X, s. 288 – 294
- [7] OBITKO, M.: *Ontologies – Description and Applications*. Report č. GL 126/01. Gerstner Laboratory for Intelligent Decision Making and Control Series of Research Reports. 2001, dostupné z: <<http://cyber.felk.cvut.cz/gerstner/reports/GL126.pdf>>
- [8] OBITKO, M., SMID J., SNÁŠEL, V.: *Ontology Design with Formal Concept Analysis*. The 2004 International Conference on Communications in Computing CIC 2004/PSMP: 21. – 24. dubna 2004, Las Vegas, Nevada, USA, CSREA Press ISBN 1-932415-36-X, s. 202 – 210
- [9] PAŠKA, M.: *Geovizualizačná ontológia v krízovom riadení*. [Diplomová práce] Masarykova univerzita v Brně, Přírodovědecká fakulta. Geografický ústav, 2007, 81 s.

- [10] RAPER, J.: *Multidimensional Geographic Information Science*. New York, Taylor & Francis Inc, 2000, 237 s.
- [11] ŘEZNÍK, T.: *Metadatový tok v krizovém řízení: od konceptu k implementaci*. [Disertační práce] Masarykova univerzita v Brně, Přírodovědecká fakulta. 2008, 128 s.
- [12] SMID, J., OBITKO, M., SNÁŠEL, V.: *Semantically Based Knowledge Representation*. The 2004 International Conference on Communications in Computing CIC 2004/PSMP: 21. – 24. dubna 2004, Las Vegas, Nevada, USA, CSREA Press ISBN 1-932415-36-X, s. 274 – 280
- [13] ŠMAJS, J., KROB, J.: *Úvod do ontologie*. Brno, VMU, 1994. Dokument dostupný z <<http://www.phil.muni.cz/fil/eo/skripta/index.html>>
- [14] ŠŤASTNÝ, R.: *Podmínky harmonizace map pro krizový management*. [Diplomová práce] Masarykova univerzita v Brně, Přírodovědecká fakulta. Geografický ústav, 2008, 72 s.
- [15] VOŽENÍLEK, V.: *Diplomové práce s geoinformatiky*. Vydavatelství Univerzity Palackého v Olomouci, Olomouc, 2002, 31 s.
- [16] ToscanaJ [počítačový program]. Ver. 1.6, 2006 [citováno 2010-01-06]. Dostupné z <<http://sourceforge.net/projects/toscanaj/files/>>
- [17] Základní mapa ČR, list 03-14-18, 1 : 10 000. Praha, ČUZK, 2006.
- [18] *Základní mapa ČR 1 : 10 000 z dat ZABAGED – mapové značky*. ČUZK, poslední revize 21. 5. 2010 [cit. 2010-07-01]. Dostupné z <http://www.cuzk.cz/Dokument.aspx?PRARESKOD=998&MENUID=0&AKCE=DOC:30-ZU_MAP_ZNAC>
- [19] *Český úřad zeměměřický a katastrální* [online]. Poslední aktualizace 21. 5. 2010 [cit. 2010-07-25]. Dostupné z <http://www.cuzk.cz/Dokument.aspx?PRARESKOD=998&MENUID=0&AKCE=DOC:30-ZU_zmsm>
- [20] *Dictionary and Thesaurus - Merriam-Webster Online* [online]. [cit. 2009-11-09]. Dostupné z <<http://www.merriam-webster.com/>>.
- [21] *Dictionnaire des sciences philosophique*, Paříž 1875, s. 1219.
- [22] *Inteligentní systém pro interaktivní podporu tvorby tematických map*. Dílčí zpráva za rok 2009 [online]. [cit.2010-07-12]. Dostupné z http://cartoexpert.comuf.com/Dilci_zprava_2009.pdf

11. SUMMARY

The diploma thesis called Map Ontology was solid at the Department of Geoinformatics of Palacky University in Olomouc.

The aim of diploma thesis is design, algorithm development and implementation of a tool for an ontology extraction from map data. There are many tools for the ontology extraction and some of them are freeware, therefore this thesis is not focused on creating and program coding of a tool for ontology extraction. This thesis focuses on a research of different ways and possibilities, how to use already existing tool or algorithm in Cartography.

The theoretical part of this thesis deal with a definition of the ontology and its infiltration from philosophy to information technologies. There are mentioned different works about ontology in the sphere of geoinformatic and cartography in more detail.

The practical part focuses on the ontology extraction from two maps – from the Base map of the Czech Republic with scale 1 : 10 000 and from the model example of creating new watchtower map. In the first case the ontology is caused by the visual perception of the map by the reader and by searching for the connections between entities in the map and in the legend of the map. Particular entities are divided and classified according to the characteristics of map symbols and the type of map symbol. Also there is used the dividing of the entities in accordance with the legend of map content of map. In the second case is the ontology caused by logical and organized data collecting and by definition of the information so that the final map contains all important entities and is instrumental purpose. The entities are defined and classified according to the entity meaning, importance and type of map symbol.

The ontology is extracted by a method that is based on Formal Concept Analysis (FCA). It is a theory of data analysis which identifies conceptual structures among data sets. In the concrete there was used open – source

software ToscanaJ. ToscanaJ is a freeware tool for data analysis on the basis of FCA.

The result of the thesis is a set of diagrams which represents the ontology extracted from the Base map of the Czech Republic with scale 1 : 10 000 and from the model example of creating watchtower. At the same time the benefit of using the ontology in the sphere of cartography is evaluated. Particularly usage of the tool ToscanaJ and the method FCA within extraction of the ontology from map datasets.

12. PŘÍLOHY

1. Tabulky objektů a atributů k diagramům ontologie legendy ZM 10
2. Diagramy ontologie mapového pole ZM 10
3. Diagramy ontologie legendy ZM 10
4. Diagramy ontologie modelového příkladu tvorby mapy rozhleden
5. CD-ROM (volná)
6. poster (volná)