

Otakar ČERBA

## ANALÝZA DEFINIC POJMU „MAPA“

Čerba, O.: *Analysis of definitions of concept “map”*. Kartografické listy 2011, 19, 5 figs., 12 refs.

**Abstract:** Cartographic communication between experts as well as systems is very complicated at this time. The domain ontologies represent a tool to improve the communication on the basis of a description of essential cartographic concepts. This paper shows the prototype (early version) of ontology processed the selected definitions of the concept “map”.

**Keywords:** definition, map, domain ontology, reasoning, cartographic communication

### Úvod

Na loňské konferenci AutoCarto 2010 se autor stal přímým účastníkem diskuze, která se zabývala problematikou tematických map na internetu. Přednášející prezentoval systém pro tvorbu tematických map, přičemž jeden z dotazů obecnosti byl zaměřen na smysl celé práce – „Proč je zapotřebí dělat takové systémy, když tematické mapy zvládají produkty společnosti Google?“ Odpovědí bylo vyjmenování typů tematických map, které software od Google nedokáže vytvořit, nebo je generuje ve velmi zjednodušené nebo kartograficky chybné podobě, případně je potřeba, aby uživatel disponoval netriviálními znalostmi z oblasti programování. Jedná se například o komplikovanější verze kartodiagramů nebo aplikaci různých stupnic pro kartogramy apod. Reakce z publika na sebe nenechala dlouho čekat – „... to přece nejsou tematické mapy, ale mapy s obrázky.“

Proč dochází k podobným nedorozuměním, ačkoli existují definice základních pojmů z oblasti kartografie, včetně termínu „tematická mapa“? Důvody lze rozdělit do dvou základních skupin:

1. Existence většího počtu definic, které vyplývají z historického vývoje kartografie a také ze specifik jednotlivých národních kartografických škol.
2. Tzv. „laicizace kartografie“, kdy kartografické nástroje (včetně software a webových služeb) nepoužívají pouze kartografové (ve smyslu odborníci v oblasti kartografie), ale také nepoučení uživatelé, kteří postrádají základní informace, včetně definic.

Jakým způsobem je takové nesrovnalosti možné eliminovat? V minulosti to byly zcela jistě slovníky nebo encyklopedie, které obsahovaly vyčerpávající informace z daného oboru. Dnes je zapotřebí širší platforma, která dokáže potřebné informace sdělit nejen člověku, ale také komunikuje se zařízeními, jež zavádí automatizované procesy do kartografie (např. geografické informační systémy, kartografické expertní systémy nebo webové služby).

Pro tyto účely se jeví jako ideální nástroje ontologie a systémy umožňující jejich zpracování, včetně tzv. reasoningu (proces, který na základě existujících znalostí je schopný odvodit znalosti nové). Účelem tohoto článku je nejen představit ontologie jako nástroj umožňující sémantické zpracování dat informací a znalostí. V článku bude také ukázána ontologie, která zpracovává základní definice pojmu „mapa“ na základě pojmů, které tyto definice obsahují. Výsledkem procesu „reasoningu“ je vygenerovaná multihierarchická struktura (tzv. odvozená ontologie), která přiřadí jednotlivé definice k příslušným kategoriím (dotazům). Díky tomuto příkladu bude jednodušší identifikovat vlastnosti ontologií, jejich výhody a možnosti uplatnění v kartografii.

Článek, skládající se ze dvou základních částí kromě úvodu a závěru, obsahuje i možnosti využití a rozšíření ontologie v budoucnosti. Nejprve jsou představeny základní pojmy především z oblasti ontologické inženýrství, se kterými nemusí být čtenáři obeznámeni. Tato pasáž obsahuje také odkazy na zdroje, které sloužily jako inspirace pro navrhovanou ontologii. Další část popisuje vlastní ontologii – její tvorbu, strukturu a výsledné vyhodnocení.

---

Ing. et Mgr. Otakar ČERBA, Západočeská univerzita v Plzni, Fakulta aplikovaných věd, Katedra matematiky  
Univerzitní 8, 306 14 Plzeň, e-mail: cerba@kma.zcu.cz

## 1. Základní pojmy a zdroje informací

Hlavním produktem kartografie v minulosti byla, je a v budoucnosti, také z největší pravděpodobností, bude mapa. Kartografové i uživatelé map se proto vždy zabírali otázkami typu „Co to vlastně je mapa?“ nebo „Co odlišuje mapu od jiných způsobů grafického vyjádření?“. Z tohoto důvodu vznikly a stále vznikají nejrůznější definice tohoto termínu. Tyto definice se snaží slovo „mapa“ více či méně přesně uchopit, specifikovat a deklarovat charakteristické vlastnosti mapy.

Jako hlavní zdroj definic termínu „mapa“ sloužily webové stránky Andrews' map definition<sup>1</sup> z roku 1998. Tyto stránky obsahovaly celkem 321 definic slova „mapa“. Definice jsou datovány od roku 1649 a jsou chronologicky řazeny. Nejnovější definice pochází z roku 1996, tedy vznikla ve stejném roce, kdy J. H. Andrews publikoval svůj článek (Andrews, 1996), ze které webové stránky vycházely. Definice jsou také obsaženy ve srovnávací studii (Bot, 2004). Pro výše uvedené výčet definic platí následující charakteristiky:

- Definice jsou uvedeny v angličtině. Mnoho z nich bylo skutečně uvedeno v originálním znění v angličtině, zbylé byly do anglického jazyky přeloženy.
- Definice obsahují stručná metadata – autora, rok publikování definice a zdroj (pokud jsou údaje známy).
- Velké množství definic bylo převzato z encyklopedií a slovníků, tedy z materiálů, které není možné považovat za odborné zdroje.

I přes výhrady uvedené v předchozím seznamu se jedná o unikátní zdroj, který přibližuje vývoj kartografie a jako vědeckého oboru i praktické činnosti v průběhu více než čtyř set let.

Další pojmy používané v tomto článku spadají do oboru softwarového inženýrství, sémantické analýzy a popisu dat, informací a znalostí. Jedná se o termín „ontologie“ a další související výrazy a dále a také o anglický výraz (zatím bez českého, případně slovenského překladu) „reasoning“, případně „reasoner“.

Ontologie jsou často chápány jako novinka, která se objevila před zhruba deseti lety v informačních technologiích. Ve skutečnosti se jedná o zhruba 2 500 let starou (Hay, 2006) vědní disciplínu (část filosofie), která dosáhla vrcholu v antickém Řecku, kde za významné představitele ontologie byli považováni např. Aristoteles nebo Platón. V rámci filosofie existují ontologie i v současnosti (Corazzon, 2010) a příslušné odkazy.

Ontologie do světa informačních technologií pronikly především díky Thomasi Gruberovi, který propojil ontologie jako filosofickou disciplínu studující bytí a jsoucno se světem informačních technologií, především datových modelů, a definoval ontologie jako „explicitní specifikace konceptualizace“ (Gruber, 1993). Tato definice byla později rozšířena Willemem Nico Borstem (1997), který ontologii považuje za „formální specifikaci sdílené konceptualizace“. R. Poli (2002) doporučuje důsledně rozlišovat přívlasky „formální“ a „formalizovaný“. Ontologie by měla být nejen formální (tj. měla by obsahovat jasně definované pojmy – sémantiku), ale také formalizovaná – jinými slovy výše zmíněné pojmy musí být zapsané pomocí přesně definovaného formátu (jazyka) – syntaxe. Pokud bychom se pokusili výše uvedené definice vyjádřit českými slovy, pak můžeme ontologii charakterizovat jako „jasný, zřetelný a přímo vyjádřený přesný výčet pojmů daného systému, včetně vztahů mezi těmito pojmy, za pomoci uzavřené množiny jazykových prostředků“.

Ontologie představují základní nástroj pro specifikaci slovníků pojmů a vztahů mezi koncepty na konkrétní (širší nebo užší) doméně (oblasti zájmu) a hrají klíčovou roli při popisu sémantiky dat (Dou et al., 2004). Z mnoha druhů ontologií pro tyto účely nejlépe vyhovují tzv. doménové ontologie (v současnosti zřejmě nejpoužívanější varianta ontologií), které se zabývají definicí pojmů a vazeb mezi těmito pojmy na konkrétní doméně (úzkém, uzavřeném okruhu znalostí a informací).

V našem případě budou základní doménu představovat definice termínu „mapa“ a výrazy (slova a slovní spojení) používané pro jeho definování. Mezi příklady doménových ontologií definovaných na základě oborů, které souvisí s kartografií, patří např. ontologie CartoOWL (Karam et

<sup>1</sup> Definice byly původně k dispozici na adrese <http://www.usm.maine.edu/~maps/essays/andrews.htm>. V současné době stránky nejsou funkční, ale na internetu existuje velké množství původních stránek. Je otázkou, jejíž řešení však není úkolem tohoto příspěvku, na kolik je re-publikování původního textu v souladu s autorskými právy.

al., 2010), ontologie připravovaná v rámci projektu CartoExpert – Inteligentní systém pro interaktivní podporu tvorby tematických map (Peňáz, 2010) nebo ontologie popisující doménu Land Use a Land Cover pro účely transformace příslušných klasifikačních systémů.

Posledním pojmem, který se zaslouží krátké představení, je termín „reasoning“. Jedná se o proces, který završuje výstavbu ontologie. Vlastní ontologie totiž nejsou určeny jen pro deklaraci explicitního popisu, včetně struktury, ale na základě tohoto popisu by měly být získávány nové informace a znalosti. K tomu slouží právě „reasoning“, resp. příslušné nástroje – „reasonery“ (např. Hermit, FaCT++ nebo Pellet) – schopné analyzovat ontologickou strukturu a na jejím základě a pomocí logických operátorů vytvořit nový, tzv. odvozený model (obr. 1 verzus obr. 4).

## 2. Ontologický systém pojmu „mapa“

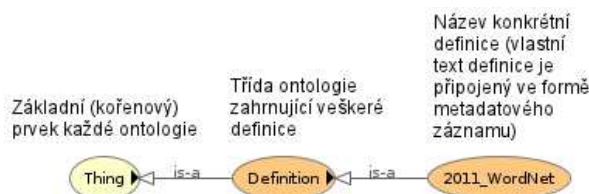
V současnosti existuje velké množství (často se vzájemně překrývajících) nejrůznějších metodik pro tvorbu ontologií. Pro naše účely byla zvolena kombinace metodiky MENTHODOLOGY (Fernandez et al., 1997) a postupu uvedeného v publikaci (Pundt, 2007).

Vlastní tvorba ontologie představuje proces složený z následujících kroků:

1. Získání úvodních informací (typu jaký bude účel ontologie, pro koho je ontologie určena, jaké bude omezení domény apod.) o budoucí ontologii.
2. Kongnitivní mapování domény, které spočívá v nalezení, popisu a definování klíčových pojmů (konceptů), jejich struktury a vztahů mezi nimi.
3. Transformace informací získaných v bodě 2 do ontologického systému – formalizace ontologie.

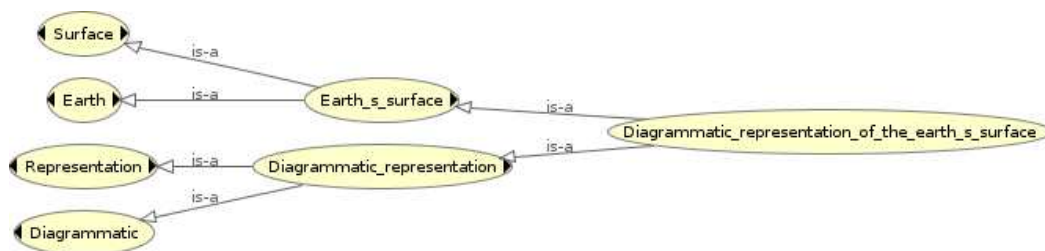
Ontologie zabývající se definicemi termínu mapa (resp. jeho anglického překladu „map“) se skládá ze dvou hlavních částí. První část popisující koncepty je tvořena třemi typy konceptů:

- Definition (jednotlivé definice /současná verze obsahuje 32 definice pokrývající 5 století/, které obsahují v názvu rok publikování definice a autora nebo zdroj definice (obr. 1) a v rámci anotace je uvedeno plné znění příslušné definice; definice jsou popsány pomocí objektových a datových vlastností),
- Term (hierarchická struktura obsahující v první úrovni základní termíny, především substantiva a adjektiva, a jejich spojení, která představují vnořené koncepty – například koncept „Flat representation“ tvoří podtřídou konceptů „Flat“ a „Representation“) – viz (obr. 2), který ilustruje strukturu definice (například výraz „earth's surface“ se skládá z pojmů „surface“ a „Earth“, přičemž při tvorbě hierarchie byly úmyslně zanedbávány přechody mezi slovními druhy), která se objevuje v systému WordNet,
- Question (množina konceptů substituující dotazy provedené nad ontologií, např. rozdělení definic podle staletí nebo obsah klíčových slov či spojení).



Obr. 1 Původní model (informace zadané explicitně) definice uvedené v systému WordNet

Druhá část ontologie obsahuje objektové a datové vlastnosti, které slouží k popisu a následnému vyhodnocení konceptů („reasoningu“). V současné verzi ontologie jsou k dispozici tři objektové vlastnosti – containTerm, isSmilar a isSynonym. ContainTerm přiřazuje každé definici konkrétní sekvenci pojmů, ze kterých se definice skládá – obr. 2 a 3. Obr. 3 ukazuje logický výraz přiřazující pomocí logických operátorů „and“ a „or“ a existenčních a univerzálních kvantifikátorů „only“ a „some“ konkrétní definici její obsah ve formě strukturovaného pojmu. Objektové vlastnosti isSmilar a isSynonym (obě vlastnosti jsou v testovací fázi) měly by sloužit k propojení významově příbuzných termínů a k nalezení příbuzných definic. V ontologii je implementována jedna datová vlastnost – year (ke každé definici přiřazuje rok jejího vzniku nebo první publikace).



Obr. 2 Struktura pojmu „Diagrammatic representation of the earth's surface“, který obsahuje definice uvedené v systému WordNet

```

● (containTerm some Diagrammatic_representation_of_the_earth_s_surface)
  and (containTerm some Diagrammatic_representation_of_the_part_of_earth_s_surface)
  and (containTerm only
    (Diagrammatic_representation_of_the_earth_s_surface
    or Diagrammatic_representation_of_the_part_of_earth_s_surface))
  
```

Obr. 3 Popis pojmu, který obsahuje definice uvedené v systému WordNet, pomocí vlastnosti containTerm

Pro vytvoření ontologie byl použitý open-source software Protégé (verze 4.1), kterou vyvíjí kalifornská univerzita Stanford University School of Medicine. Protégé je označován jako editor ontologií a expertní systém. Tento software má několik nesporných výhod (např. podpora jazyka Web Ontology Language 2.0 /OWL/, pravidelné aktualizace, nezávislost na platformě, rozsáhlá dokumentace nebo uživatelské prostředí), kterého ho staví na přední místa mezi podobnými programovými produkty.

Jak bylo zmíněno v předchozích odstavcích, vytvoření vlastní ontologie představuje pouze první krok. Následovat musí tzv. „reasoning“, tedy proces, který na základě informací explicitně definovaných v ontologii vytvoří informace nové. V případě ontologie popisující definice termínu „mapa“ byl „reasoner“ (konkrétně FaCT++ 1.52) použit pro hledání odpovědí na koncepty ve třídě Questions. Aktuální verze obsahuje následující dotazy:

- Q\_Century – celkově pět konceptů, které jsou deklarovány příslušností definic k jednotlivým stoletím (obr. 4). Dotaz probíhá na základě posuzování minimální a maximální hodnoty datové vlastnosti year. Výsledky dotazu vizualizované v obr. 4 lze přeložit tak, že definice uvedené v systému WordNet obsahuje pojmy „part“, „representation“ a „earth's surface“, přičemž poslední pojem se skládá ze dvou dalších výrazů („Earth“ a „surface“, které společně s výše uvedenými termíny patří do skupiny klíčových slov pro definici pojmu „mapa“. Z obr. 4 je možné dále zjistit, že „earth's surface“ patří mezi sousloví (výrazy vzniklé kombinací jiných pojmů) a že definice podle WordNet byla vytvořena ve 21. století.
- Q\_Top\_Terms – rozdělují jednotlivé definice podle to, zda obsahují některý z nejčastěji se vyskytujících termínů, přičemž jedna definice může být přiřazena k více termínům (vytvoření multihierarchie). Nejfrekventovanější termíny definic byly získány analýzou textu všech definic pomocí online nástrojů Textalyser (<http://textalyser.net/>) a Online-Utility.org Online Text Analysis Tool (<http://www.online-utility.org/text/analyzer.jsp>). Před vlastní analýzou musely být zdrojové texty ontologie upraveny pomocí XSLT stylů – došlo k výběru textů definic a odstranění elementů jazyka OWL. Po eliminaci slov, která byla z hlediska obsahu ontologii marginální (předložky, spojky apod.), byly vybrány nejfrekventovanější pojmy všech definic (v pořadí podle četnosti) – „surface“, „earth“, „representation“, „part“ a „graphic“<sup>2</sup> – obr. 4. Všechny termíny byly zobrazeny pomocí tzv. „cloudu“ (online nástroj Wordle.net), který ukazuje nejčastěji používaná slova v grafické podobě (obr. 5).

<sup>2</sup> Zajímavé je, že tento výraz se, na rozdíl od předchozích termínů, začíná objevovat v definicích až koncem první poloviny 20. století.



5. Vytvoření aplikace (nadstavby ontologie), která bude využívat výsledky ontologie pro sémantické vyhledávání.

6. Doplnění dalších termínů a jejich definic.

Ontologie v pokročilejším stádiu by měla sloužit ke zlepšení kartografické komunikace ať už ve formě „inteligentního slovníku“ nebo tezauru. Může také představovat nástroj (nikoli řešení) podporující analýzy vedoucí ke změnám základních definic oboru kartografie. Na základě nově získaných znalostí bude možné deklarovat klíčové pojmy, které se v souvislosti s mapou a kartografií obecně objevovaly od minulosti po současnost. Důležité bude také sledování změn ve výskytu klíčových pojmů, které odráží vývoj kartografie a změny v pohledu na její úlohu ve společnosti i v systému vědních oborů. V neposlední řadě je zapotřebí zmínit úlohu definice mapy v procesu komunikace mezi kartografy navzájem a především mezi kartografy a automatickými systémy.

Výhodou navrhovaného řešení je aplikační nezávislost a otevřenost ontologií, které umožňují implementaci výsledné odvozené ontologie do libovolného konkrétního řešení, např. kartografických expertních systémů nebo generických ontologií (ontologií, které překračují hranice jednotlivých domén). Dále je možné popsaný postup znovu využít pro zpracování jiných pojmů a témat z oblasti kartografie a také různých taxonomických systémů.

### Literatúra

- ANDREWS, J. H. (1996). What was a map? The lexicographers reply. *Cartographica*, 33, 4, s. 1-11.
- BORST, W. N. (1997). *Construction of engineering ontologies for knowledge sharing and reuse*. Enschede (University of Twente).
- BOT, K. (2004). *42 or 363 Definitions of cartography*. (The Institute for Infinitely Small Things). [cit. 2011-04-20]. Dostupné na: <[www.lulu.com/items/volume\\_26/417000/417228/1/print/009\\_42.pdf](http://www.lulu.com/items/volume_26/417000/417228/1/print/009_42.pdf)>
- CORAZZON, R. (2010). *Birth of a new science: the history of ontology from Suárez to Kant*. [cit. 2011-04-20]. Dostupné na: <<http://www.ontology.co/history.htm>>
- DOU, D. et al. (2004). Ontology translation on the semantic web. In Spaccapietra S. et al. (Eds.): *Journal on Data Semantics II*, LNCS 3360, s. 35-57, Berlin Heidelberg (Springer-Verlag).
- FERNÁNDEZ, M. et al. (1997). METHONTOLOGY: From ontological art towards ontological engineering. In *AAAI Technical Report SS-97-06*.
- HAY, D. C. (2006). Data modelling. RDF & OWL – part one: an introduction to ontologies. In *The data administration newsletter – TDAN.com*. [cit. 2011-04-20]. Dostupné na: <<http://www.tdan.com/view-articles/5025/>>
- GRUBER, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5, 2, s. 199-220.
- KARAM, R. a kol. (2010). Cartographic integration on mobile devices from several providers' LBS by means of map symbol ontology. In *WebMGS 2010*, Como, Italy.
- PEŇÁZ, T. (2010). Konstrukce ontologického modelu pro aplikační využití znalostí z domény tematické kartografie. In *Geografie pro život ve 21. století*. Ostrava (Ostravská univerzita v Ostravě), s. 259-265.
- POLI, R. (2002). Ontological methodology. *International Journal Human-Computer Studies*, 56, s. 639-664.
- PUNDT, H. (2007). From idea toward ontology. In *10th AGILE International Conference on Geographic Information Science*. (Aalborg University).

### S u m m a r y

#### Analysis of definitions of concept “map”

Domain ontologies can help to eliminate the vagueness of cartographic communication. The example mentioned in this paper introduces the prototype of domain ontology processing selected definitions of the essential cartographic concept “map”. Ontology is composed of three main parts:

- Classes divided into Definitions (including their description through properties and metadata), Terms (and their combinations) and Questions (classification of definitions based on used terms and other properties).
- Object properties interconnecting the classes Definitions and Terms.
- Data properties describing the definitions.

Ontology enables to detect the most usual definitions – these definitions contain four from five of the most frequented words in the definitions (the word were acquired by text analyses) – as well as the most extraordinary definitions (without above-mentioned terms).

The conclusion identifies the future steps necessary to further development of ontology and its application.

Fig. 1 Asserted model of the definition from WordNet.

Fig. 2 The structure of concept used in the definition WordNet.

Fig. 3 Description (through property containTerm) of the essential term contained in definition WordNet.

Fig. 4 Inferred model of the definition from WordNet.

Fig. 5 Cloud visualisation of the most frequented words in definitions of the term “map”.

**Recenzoval:**

**Doc. Ing. Milan HÁJEK, CSc.,**

**Zohor**