

Produktové ontologie pro elektronickou komerci (včetně zadání pro práci na cvičeních)

Doc. Ing. Vojtěch Svátek, Dr.
Spolupráce Bc. Vlastimil Ovčáčík, Bc. Marek Dudáš

Letní semestr 2012

<http://nb.vse.cz/~svatek/rzzw.html>

Produktová ontologie

- Též označovaná jako „vertikální ontologie“ e-businessu
- Může zachycovat mj.
 - **Taxonomii** komodit určitého typu
 - Jejich **partonomii** (složení z částí, např. náhradní díly)
 - Soubor **parametrů** typicky používaných pro popis každé z komodit, a jejich hodnot

Vazby na jiné ontologie

- Typy komodit mohou odpovídat kategoriím z
 - Oficiálních **oborových taxonomií** (UNSPSC, eCI@ss, ebXML npod.)
 - Wikipedie, resp. z ní přímo odvozené **Product Type Ontology** (PTO), <http://productontology.org>
- Vrcholové kategorie pro parametry mohou odpovídat ontologii **GoodRelations** (GR)
 - <http://www.heppnetz.de/projects/goodrelations/>
 - Rozlišení kvantitativních vs. kvalitativních vlastností, integer/float...
 - Základní typy parametrů: výška, délka, hmotnost...
- Z **DBpedia** (<http://dbpedia.org/>) lze převzít
 - Hodnoty parametrů (např. typy použitých dílů)
 - URI instancí (produktů) - pro populární produkty

Vehicle Sales Ontology (VSO)

- <http://www.heppnetz.de/ontologies/vso/ns>
- Taxonomie typů vozidel
 - Resp. dopravních prostředků vč. plavidel (pragmatický přístup)
- Další třídy zastřešují hodnoty parametrů
 - Např. typ motoru
- Partonomie
 - Explicitně se na úrovni VSO nemodeluje
 - Ve vztahu k jiným, samostatně nabízeným typům produktů lze použít *gr:isAccessoryOrSparePartFor*
- Parametry
 - Modelovány několika způsoby, viz dále

Parametry ve VSO

- Lze definovat **objektovou vlastností** a **třídou**
 - v případě předdefinovaných hodnot a jejich intervalů (i neceločíselných), např. typ vozidla, paliva, převodovky...
 - třída může být specifická z VSO, např. *vso:EngineTypeValue*, nebo obecná z GR, např. *gr:QuantitativeValue*
- Lze definovat **datovou vlastností**
 - v případě otevřené množiny hodnot, např. datum výroby, textový popis stavu, VIN
 - ne pro kvantitativní vlastnosti, ty se zřejmě vždy chápou jako rozložitelné na enumerace hodnot/intervalů
- Vlastnosti obou typů
 - mají kardinalitu uvedenou v rámci *rdfs:label*
 - jsou deklarované jako podvlastnosti obecných vlastností z GR (např. *gr:quantitativeProductOrServiceProperty*)

Další vertikální ontologie

- Tickets Ontology
 - <http://www.heppnetz.de/ontologies/tio/ns>
- Mobile Device Ontology
 - <http://www.openvocabulary.info/ontologies/device.html>
- Consumer Electronics Ontology (CEO)
 - <http://www.ebusiness-unibw.org/ontologies/consumerelectronics/v1>
- **Reusable Energy Ontology (REO)**
 - Kolaborativní vývoj ve 4IZ440, ZS2011, dodatečné úpravy a doplňování V. Ovčáčík

Úkol pro nás

- Vytvořit obdobnou ontologii pro aktivní síťové prvky.
- Ontologie bude používat anglické pojmy v URI a anglické „labels“, české pojmy v českých „labels“
- Přímé využití GoodRelations, PTO a DBpedia
- Inspirace strukturou VSO a REO
- Předpokládaný rozsah:
 - pokud možno se přiblížit úplnému pokrytí typů produktů a parametrů z podkladových materiálů
 - řádově bude potřeba několik desítek až set entit (tříd, objektových a datových vlastností a individuí

Podklady a nástroje

- Textové podklady
 - Obsahují řadu informací, nás budou v tuto chvíli zajímat hlavně **typy** nabízených produktů a **parametry**, jakých mohou produkty nabývat
- Kolaborativní verze Protégé
 - Existuje verze klient-server a webová
 - Souběžné editování ontologie více návrháři
 - Vedení diskuse online
 - Kromě připomínek a návrhu úprav entit využijeme i volbu „Example“ – pro skutečné produkty daného typu

Organizační poznámky

- Jako klient se použije Protégé 3.4.8 - <http://protegeserver.cz/navod.html>
- Ontologie
 - Je nazvaná Networking Hardware Ontology
 - Je umístěná na serveru **protegeserver.cz**
 - Už obsahuje naimportovanou ontologii GoodRelations
 - Na entity z dalších ontologií (např. PTO) se lze odvolávat pomocí nadefinovaných prefixů
- Tvorba ontologie a diskuse nad ní bude probíhat v období 2.-20.4.
- Příspěvek jednotlivých studentů k ontologii je součástí hodnocení předmětu (max. 12 bodů)
 - Do úvahy se berou nejen přidání/úpravy entit, ale i zapojení do diskuse
- Podkladové materiály jsou rozděleny do 4 adresářů, kterým odpovídají ad hoc studentské týmy (s užší spoluprací uvnitř týmu)
 - Viz seznam *rozdeleni.txt* na disku G:

Klíčová návrhová rozhodnutí

- Objektové vs. datové vlastnosti pro vyjádření kvantitativních parametrů?
 - VSO zjevně preferuje objektové, i s hodnotami typu „hmotnost_360_kg“
- Lze dohledat příslušný pojem na Wikipedii? Pokud ne, jak přeložit?
- Co je seriózní parametr a co jen vágní reklamní tvrzení prodávající firmy?
- Stanovení *competency questions* – otázek, na které ontologie má mít odpovědi.
- Jak podrobná má ontologie být?
- Atd., bude doplněno dle praxe!

Další postup práce

- „Populování“ ontologie **fakty** o konkrétních produktech, firmách, lokalitách atd
 - Při něm se využijí „příklady“ produktů uložené v kolaboračních anotacích
- Linkování na **číselníky** ekonomických činností
- Porovnávání produktových nabídek s vypsánými **veřejnými zakázkami** ze systému ISVZUS
 - Využití RDF store na MFF UK