

TMSPARQL mapování

by Networked Planet

Připravil V. Ovčáčík

TM vs RDF

- Oba jsou standardy sémantického webu
- Topic Maps je určeno pro lidi. RDF pro stroje, až v druhé řadě pro lidi
- Zprvu se myslelo, že si navzájem konkurují
- V současnosti se zdá, že TM stagnují
- Pro SW by bylo nejlepší, pokud by se oba standardy propojily – musí se najít společný jazyk. Nejvhodnějším se zdá SPARQL

Mapování

- Rok 2009, společnost Networked Planet
- Cílem je otevření Topic Maps datových skladů pro Linked Data konzumenty
- Plně implementováno v komerční produktech Networked Planet
- Open source implementace – Nikunau
- Detaily v článku „[Making Topic Maps SPARQL](#)“
- Nejedná se o mapování tolog jazyka na SPARQL, ale o mapování TM na RDF (v rozsahu nutném pro dotazování)

Cíle

- Mapování musí být transparentní. V ideálním případě by uživatel neměl poznat, že SPARQL dotaz běží nad TM databází
- Otevření TM skladů bez úprav existujících dat
- Částečné nahrazení TML jazykem SPARQL

Módy

- Implicitní mód
 - vztahy jsou mapovány na jeden, popř. dva triple patterny
 - pouze pro binární vztahy
- Explicitní mód
 - využívá proprietárních zabudovaných predikátů: `rdf:type`, `tms:reifier`, `tms:player`, `tms:scope` apod.
 - na rozdíl od implicitního módu, umožňuje komplexnější pohled na TM (např. ternární vztahy, scope, reifikace atd.)
 - Všechny predikáty (kromě `rdf:type`) jsou v přímém rozporu s cíli mapování (neumožní transparentnost)

Vztahy

- Vztah v TM je definován nejméně 5 uzly, zatímco v RDF přesně 3 uzly
- Každý binární vztah lze přeložit na dva triple patterny
- Association type se v implicitním módu vždy ztratí, v explicitním módu může být zachován

`o:composed_by(o:Tosca : o:Work, o:Puccini : o:Composer)`

Association type

Role player

Role type

Role player

Role type

`o:Tosca o:Composer o:Puccini .`

Subject

Predicate

Object

`o:Puccini o:Work o:Tosca .`

Subject

Predicate

Object

Výskyty

- Podstatně jednodušší než vztahy, protože výskyt je tvořen právě 3 uzly.

o:date_of_birth(o:Puccini, \$BIRTHDAY)

Occurrence type

Topic

Value

o:Puccini o:date_of_birth ?BIRTHDAY .

Subject

Predicate

Object

Nikunau

- Součást projektu SesameTM vyvíjeného skupinou Topic Maps Lab (Lipská univerzita)
- Pro vyhodnocení SPARQL dotazů využívá Sesame engine
- Je možné jej použít pro jakýkoli TM backend podporující TMAPI
- Nasazení
 - Maiana (včetně SPARQL endpointu)
 - Ontopia (**sparql4ontopia** plugin; zatím bez endpointu)
- Nikunau má řadu omezení a bugů

Nedostatky Nikunau

- Nepřehledný, neokomentovaný a stěží rozšiřitelný kód
- V dotazech není možné uvést literály (mohou ale být součástí výsledku)
- Neimplementuje explicitní mód
- `rdf:type` v Nikunau není stejný jako `rdf:type` v RDF (ignoruje se vztah subclass-superclass)
- A další...

Závěrem

- Není možné současně naplnit všechny cíle mapování. Praktické využití je možné pro Topic Maps, jejichž ontologie dopředu počítá s mapováním a umožňuje vynechání explicitního módu při zachování přesnosti a úplnosti výsledků.
- Tutoriál na překlad dotazů z tolog do SPARQL:
http://nlp.vse.cz:8080/omnigator/models/topi_cmap_complete.jsp?tm=ItalianOpera.ltm