

Vyhľadávanie životných situácií prostredníctvom sémantiky

Vladimír Kuchár, FIIT STU, Datalan a.s.

Obsah

1. Aktuálny stav
2. Motivácia a ciele
3. Riešenie
4. Diskusia

Aktuálny stav

Chcem nájsť

Hľadať

Napríklad [Občiansky preukaz](#) alebo [Výpis z registra trestov](#)

Občan

Podnikateľ

 Bývanie


 Cestovanie

 Doprava

 Financie

 Kultúra

 Občan a štát

 Obrana a bezpečnosť

 Rodina a vzťahy

 Vzdelanie a šport

 Zamestnanie

 Zdravie

 Životné prostredie

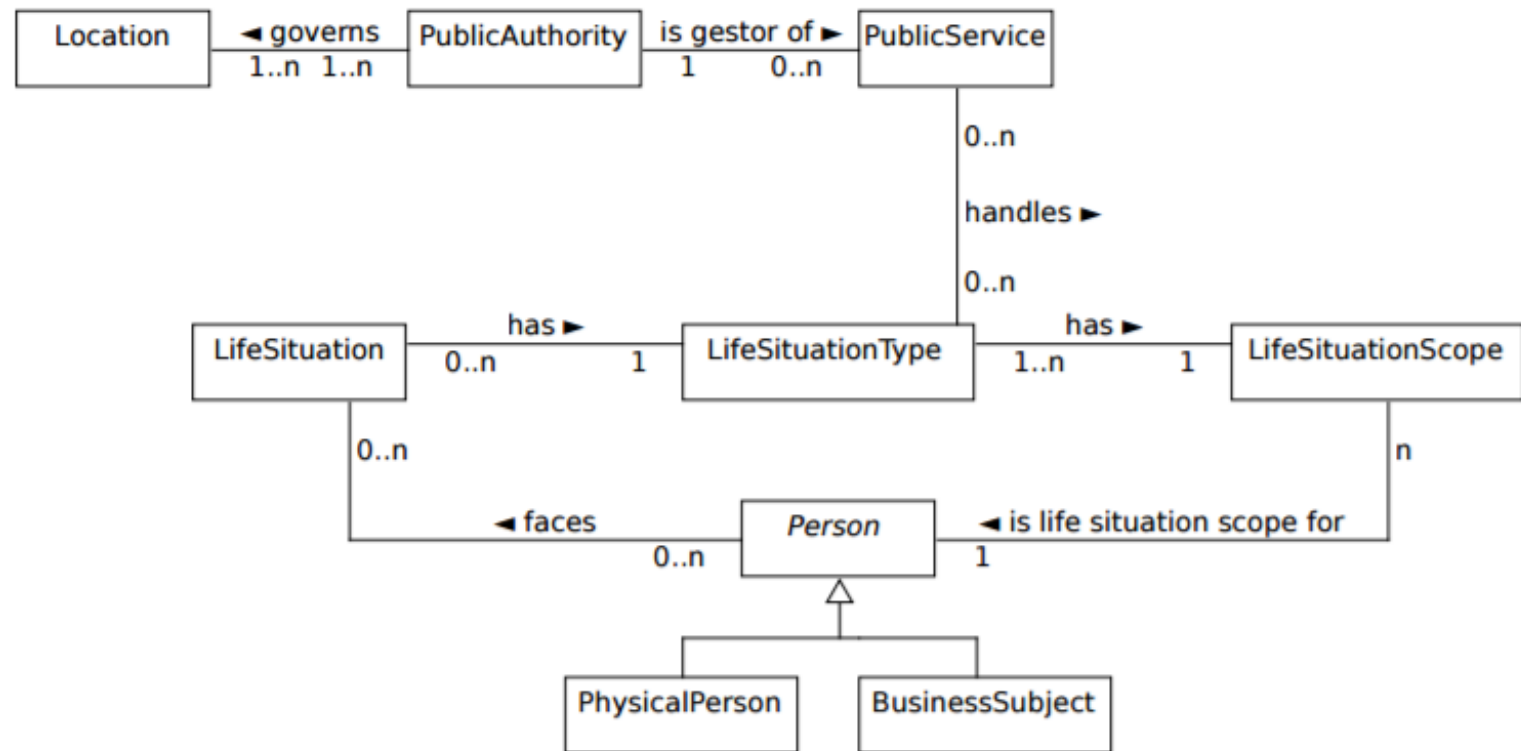
<https://www.slovensko.sk>

Motivácia a ciele

- Identifikovať dôležité koncepty
- Zachytiť vzťahy medzi konceptami
- Umožniť sémantické vyhľadávanie
- Demonštrovať vhodnosť sémantických technológií na tento účel

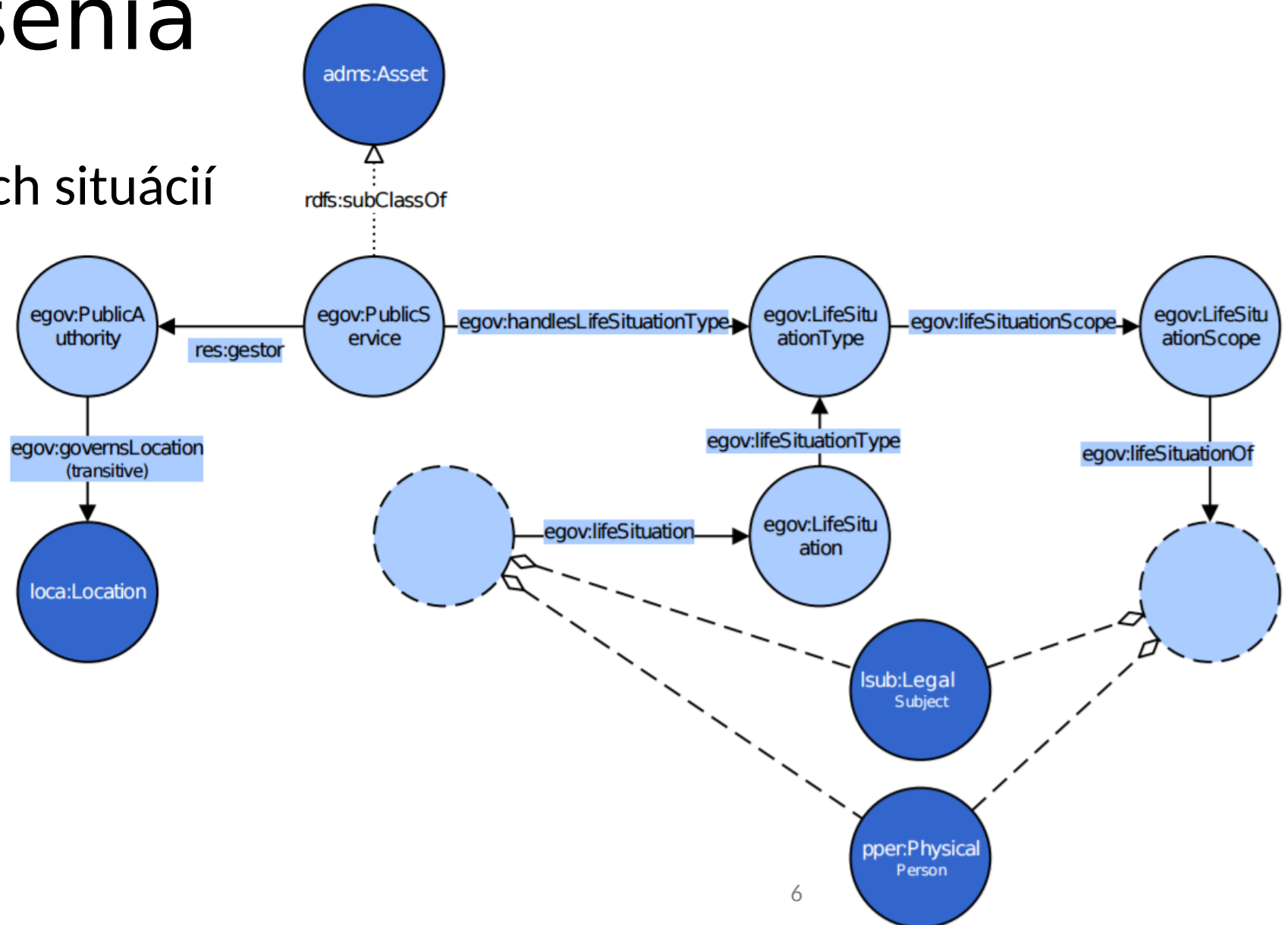
Návrh riešenia

- Konceptuálny model



Návrh riešenia

- Ontológia životných situácií



Vlastnosť *egov:governsLocation*

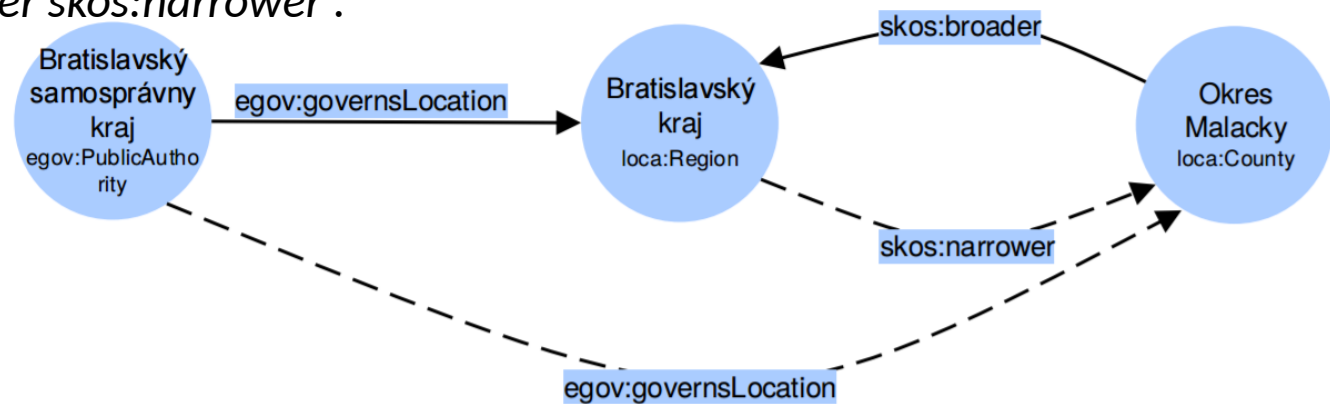
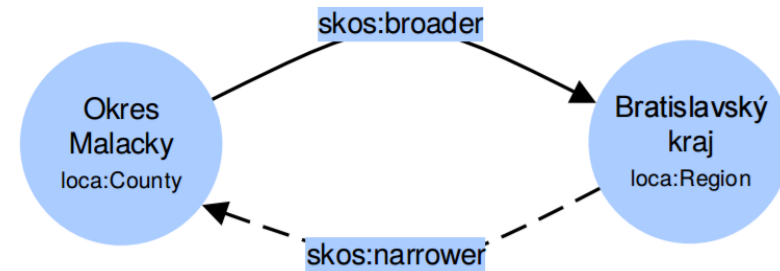
- Hierarchia v ontológií Location

-
-

- Využitie tejto hierarchie

$(P \text{ transitiveOver } Q), (x P y), (y Q z) \Rightarrow (x P z)$

egov:governsLocation *ptop:transitiveOver* *skos:narrower* .



Transforácia dát do RDF

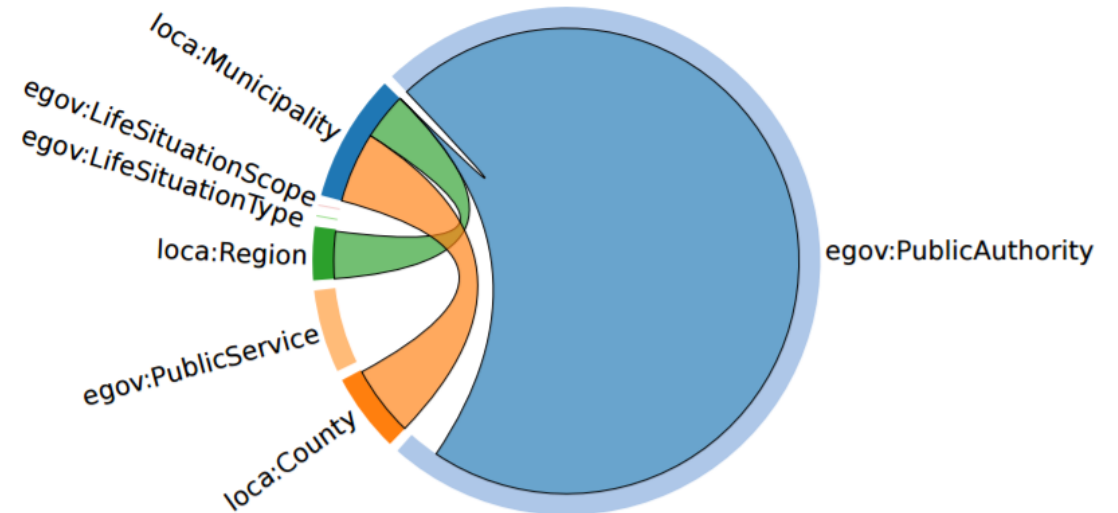
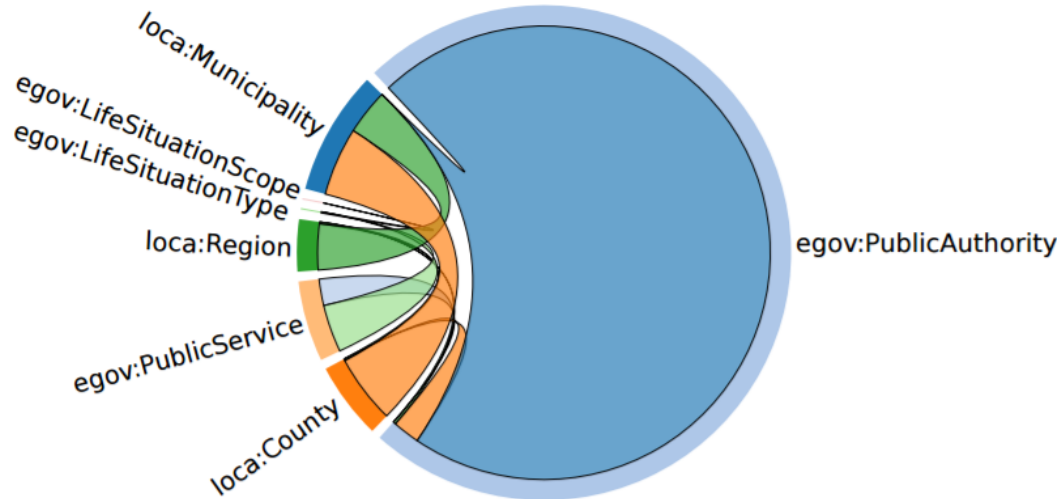
- Číselník obcí, okresov, krajov
 - Nekonzistentné používanie identifikátorov
 - 12096 explicitných tripletov
- Dataset životných situácií a dáta o povinných osobách
 - Životné situácie a eGovernment služby štátnej správy
 - Dáta o povinných osobách z MetaIS
 - Typ
 - Opis typu
 - TYP_A1
 - ministerstvá
 - TYP_A2
 - ostatné ústredné orgány štátnej správy
 - TYP B
 - Generálna prokuratúra SR, Najvyšší kontrolný úrad SR, iné štátne orgány
 - TYP C1
 - obce
 - TYP C2
 - vyššie územné celky
 - TYP D1
 - Kancelária NR SR, Kancelária prezidenta SR, ...

Transforácia dát do RDF

- Typy
 - TYP_A1, TYP_A2, TYP_B, TYP_D1, TYP_D2, TYP_D3, TYP_D4
 - TYP_C1
 - TYP_C2
- Objekt predikátu `egov:governsLocation`
 - priamo na inštanciu Slovenskej Republiky s URI <https://data.gov.sk/id/un-country/703>
 - príslušná dedukovaná inštancia obce povinnej osoby typu `loca:Municipality`
 - príslušná dedukovaná inštancia kraja povinnej osoby typu `loca:Region`
- - Z 2985 povinných osôb sa nepodarilo určiť *egov:governsLocation* pre 214
 - obce so zhodným názvom bez vyplneného okresu
 - chyby v diakritike
 - Cabaj-Čápor s povinnou osobou Čápor
 - mestá Bratislava a Košice

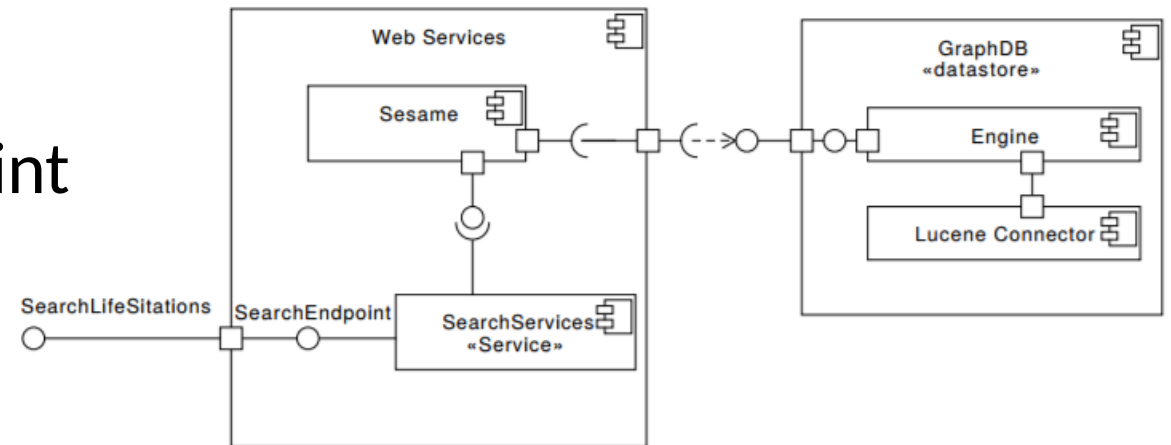
Výsledky inferencie

- počet explicitných tripletov • 52366
- počet odvodených tripletov • 193283
- počet tripletov celkovo • 245649
- pomer nárastu počtu tripletov • 4,691



Implementácia vyhľadávania

- Graph DB 8 Free
- RDF4J / Sesame
- Náš SearchService, SearchEndpoint
-
-
- „pobyt v Lozorne“
 - > Trvalý pobyt/Stahovanie, 6 služieb
 - > Prechodný pobyt, 3 služby
 - > Pobyt v zahraničí, 7 služieb



Ďalšia práca

- Rozšíriť ontológiu aby zachytávala aj povinné osoby, na ktoré je prenesený výkon verejnej moci a ich vzťah k nadriadenej povinnej osobe
- Rozšíriť ontológiu, aby korektne zachytávala vzťah mesta Bratislava a Košice s ich mestskými časťami
- Vylepšiť implementáciu vyhľadávania, slovníky
- Možnosť pridať iné rôzne služby a filtre vyhľadávania

Recap

- Aktuálny stav
- Motivácia
- Riešenie
-
-
-
-

Ďakujem za pozornosť