

# GISON: ontologická integrace urbanistických datových sad IPR

Jiří Čtyroký<sup>1</sup>, Petr Křemen<sup>2</sup>

<sup>1</sup>IPR Praha, <sup>2</sup>FEL ČVUT

---



IPR ———  
PRAHA



# Slovníky a ontologie

---

- Měli bychom co nejvíce používat třídy a vlastnosti definované existujícími slovníky.
- Někdy je ale vhodné nebo nutné zavést vlastní třídy či vlastnosti, tj. definovat vlastní slovník.
- Při definici vlastního slovníku je zásadní vysvětlit sémantiku (význam) nových tříd a vlastností pomocí existujících slovníků tam, kde je to možné.

# Urbanismus a vymezení pojmů

---

- Velké množství pojmů
- Paralelní sémantické zakotvení:
  - Právní (zakotvení v právních předpisech)
  - Zvykové (aktivně užívané komunitami)
- Komplexní vztahy mezi pojmy
- Význam vyžaduje vždy kontext
  - (např: stavba, komunikace)
- Rozdílné terminologie různých (profesních) komunit
  - (např: park vs. zeleň, brownfield vs. přestavbové území)



# Datové zdroje

---

- Pojmově ukotvené (významy pojmů jsou definovány)
  - Základní registry
  - Další informační systémy veřejné správy (např. ZABAGED, ISKN, VDB ČSÚ)
  - Popisná data regulativů územního plánu
  - Objekty nazvané shodnými pojmy nemusí mít v různých IS totožný význam ani vymezení!
- Pojmově neukotvené (významy pojmů jsou definovány nedostatečně)
  - Digitální technická mapa
  - jevy Územně analytických podkladů
  - Interní podkladová a analytická data

# Datové zdroje

---

- Obtížné vzájemné propojování
  - Rozdílné definice pojmů
    - Věcně shodné objekty mohou být definovány rozdílnými pojmy
    - Shodné pojmy mohou být využívány pro určení rozdílných objektů
  - Rozdílné významové (a logické) hierarchie pojmů
- Pojem může být reprezentován objektem na jakékoli úrovni hierarchie datového modelu:
- Téma\*
  - Datová sada (datová třída/třída objektů)
    - Podtřída objektů = definovaná společnou hodnotou domény atributu (z číselníku)
      - Objekt = definovaný unikátním identifikátorem (vlastností)

\* *Může existovat několik paralelních tematických thesaurů*

# Řešené problémy

---

- Hledání dat
  - Hledání datových sad/tříd objektů
  - Hledání objektů v datových sadách/třídách objektů

Např:

najdi datové sady ploch bydlení

najdi datové sady lesů

- Hledání území požadovaných vlastností

Např:

najdi území pro bydlení v Praze 6

Kolik rodinných domů se nachází do 1 km od lesa?

# Příklad: plochy bydlení

---

Pojmy: Bydlení ↔ Obytné

Bytové ↔ Rodinné domy ↔ Činžovní vily

Smíšené

IPR nedisponuje samostatnou datovou sadou „plochy bydlení“, ale existuje vymezení v:

- Třída: Současný stav využití území (UAP)
  - Podtřídy: BD - bytové domy, BQ - polyfunkční rodinné a bytové domy, BRR - rodinné domy, BRV - činžovní vily
- Třída: Plán využití ploch - funkční plochy (územní plán)
  - Podtřídy: OB - čistě obytné, OV - všeobecně obytné, SMJ - smíšené městského jádra, SV- všeobecně smíšené

# Ontologie vs. taxonomie

---

Ontologie poskytují způsob, jak *tvorit pojmy, definovat jejich význam a propojovat je tak, aby byly **pochopitelné, sdílené a přepoužitelné.***

Vyžadují zamyslet se nad významem pojmů a vazeb, popsat je a formalizovat. Jdou dál než klíčová slova, či taxonomie.

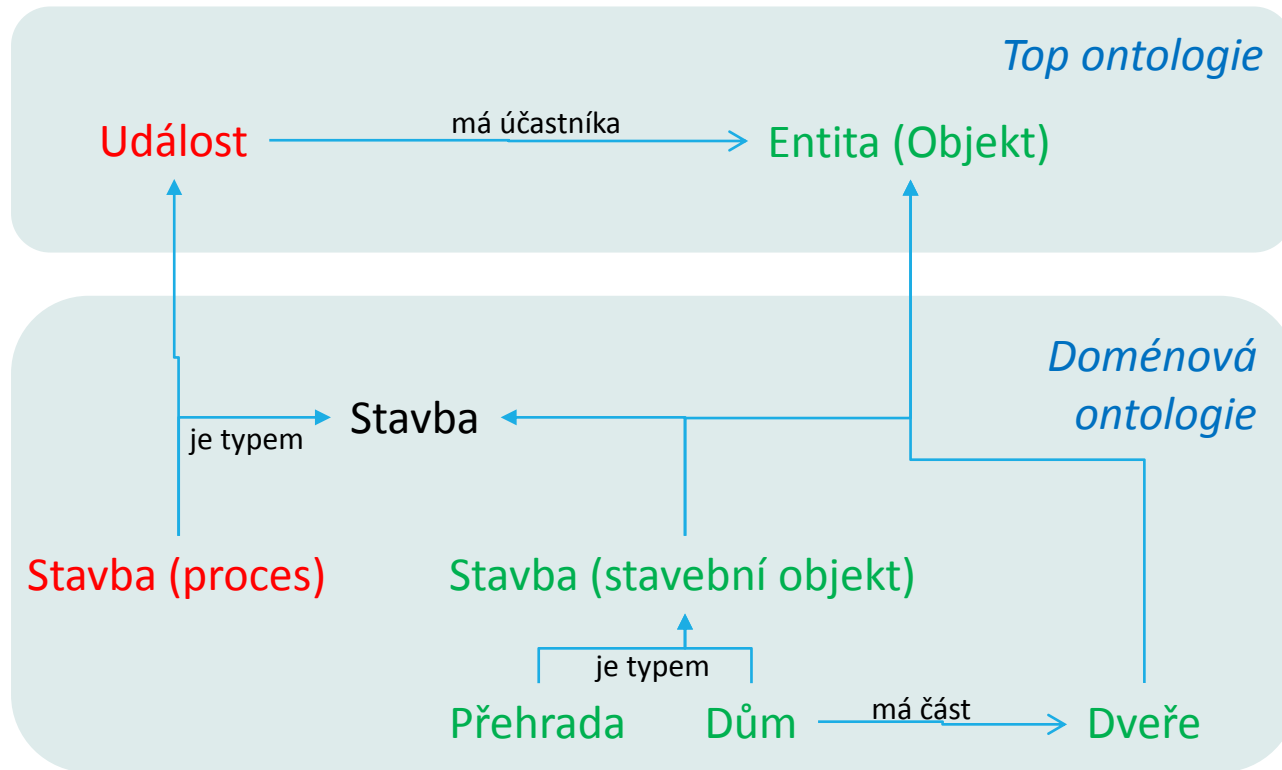
- Stavba ← široký význam (objekt, staveniště, proces)
- Přehrada
- Dům ← široký význam (obydlí, stavba)
- Dveře ← není specializací nadřazeného pojmu, ale jeho součástí

**Propojená data = sdílení dat**

**Ontologie = sdílení významu dat**



# Ontologie jsou modelem znalostí



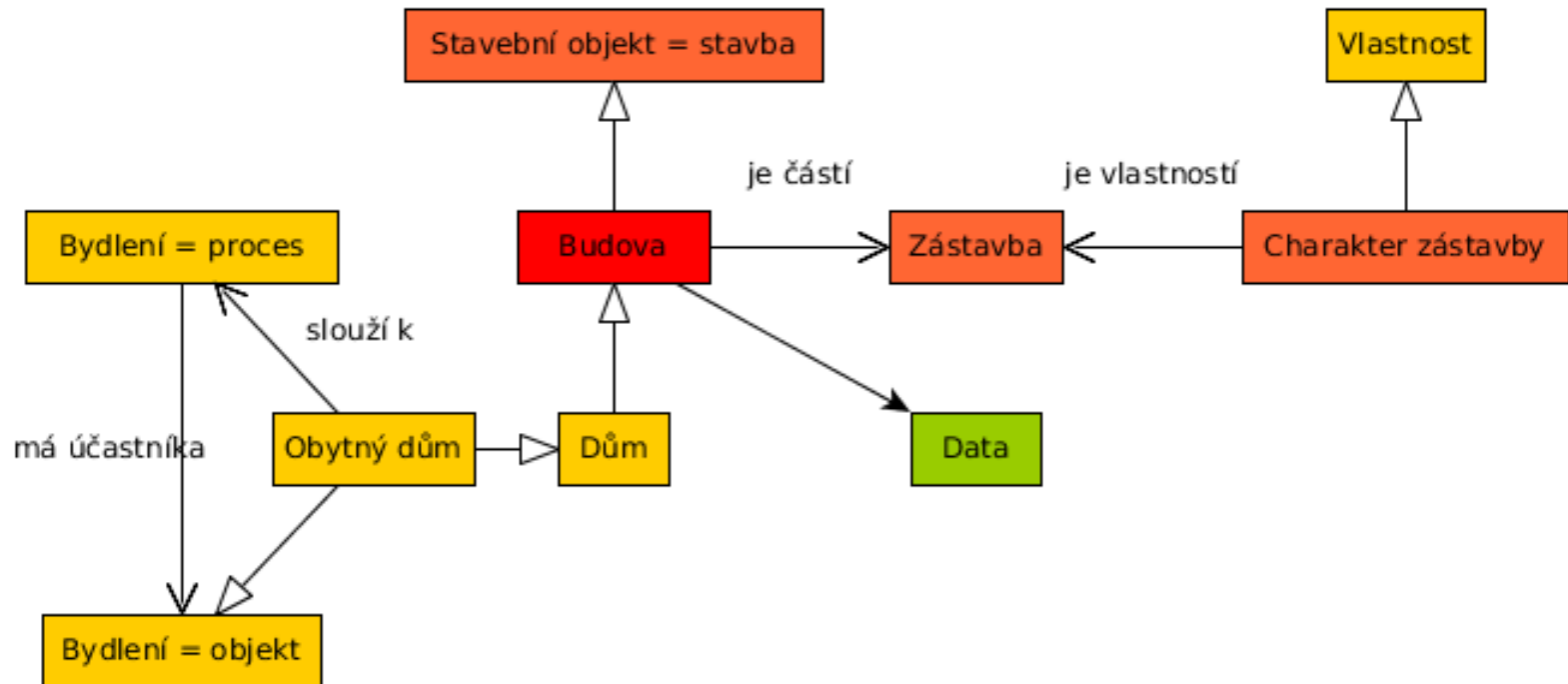
Metodologie  
(UFO)

Top-level ontologie  
(UFO-A, UFO-B)

Technologie  
(RDF(S), OWL, SWRL, SPARQL)

# Popis datových sad IPR

- Každá datová sada (dosud zpracováno pět datových sad) je popsána **ontologickým modelem** dat v ní obsažených, např. datová sada *Podlažnosti* (viz níže).
- Ontologické modely datových sad jsou významově propojeny pomocí integrační ontologie.



# Distribuované dotazy

---

- Příklad: *Které pozemky mají být dle územního plánu **použity pro bydlení** a **dosud tak použity nejsou** ?*
  - Zodpovězení dotazu vyžaduje informace ze dvou datových sad
    - Funkční využití území (FVU)
    - Současný stav využití území (SSVU)
1. Překlad dotazu do integrační ontologie
  2. Nalezení datových sad, které umí zodpovědět části dotazu
  3. Vyhodnocení částí dotazu nad vybranými datovými sadami
  4. Integrace výsledků dotazu