

Semântica na Web

Vocabulários

Carlos Bazilio

Depto de Computação
Instituto de Ciência e Tecnologia
Universidade Federal Fluminense

Motivação

- Imagine aplicações no contexto de saúde:
 - Profissionais usam termos para sintomas, doenças, tratamentos, ..
 - Empresas farmacêuticas usam termos para drogas, dosagens, alergias, ..
 - Sistemas de saúde usam termos para pacientes, sintomas, medicação, eficácia, ..
- É natural que pensemos no cruzamento destes dados para descobrirmos novas informações

Motivação

- Como garantirmos que este cruzamento se dará de forma correta, que as diferentes aplicações se referem ao mesmo conceito?

O que são Vocabulários?

- Definem conceitos e relacionamentos (termos) de uma área de domínio
- São usados para classificar termos, os quais poderão ser usados por aplicações específicas
- Classificação envolve detalhamento, restrição, ...
- Podem ser muito simples ou bastante complexos

O que são Vocabulários?

- Não há uma clara distinção entre vocabulários e ontologias
- Usualmente vemos o termo ontologia sendo utilizado para coleções complexas de termos
- Vocabulários são o “tijolo” para as técnicas de inferência na Web Semântica

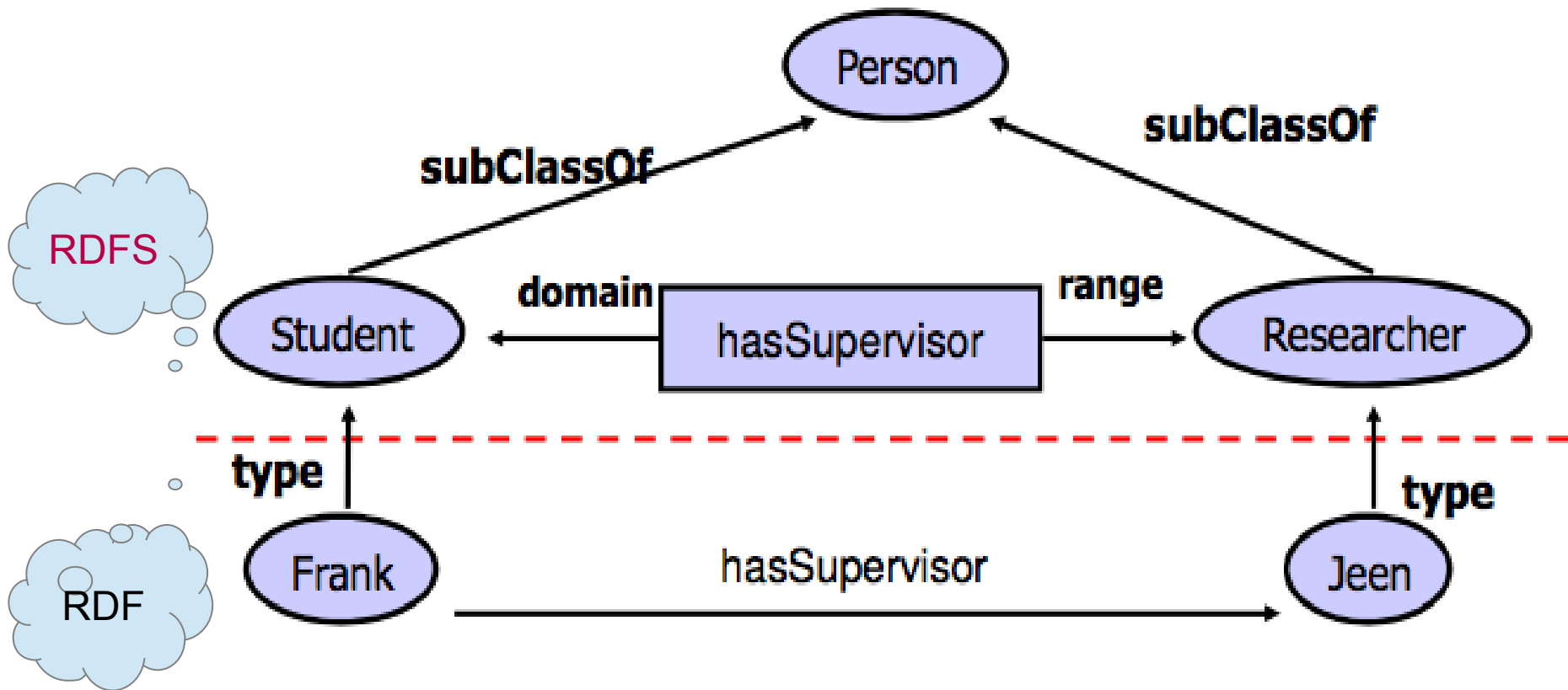
Quando são usados?

- Integração de dados entre diferentes datasets
- Também são úteis para fornecer informação extra que possa gerar novos relacionamentos
- Por exemplo, febre e estado febril podem ser equivalentes ou não, dependendo da aplicação em questão

Definição de Vocabulários

- A W3C disponibiliza um série de linguagens para a definição de vocabulários
 - RDF e RDFS (RDF Schemas)
 - Simple Knowledge Organization System (SKOS) - *Taxonomias*
 - Web Ontology Language (OWL)
- A escolha depende da complexidade e do rigor exigido na especificação dos dados de uma aplicação

RDF Schema [2]



RDF Schema

Person rdf:type #Class

Student rdf:type #Class

 rdfs:subClassOf #Person

hasSupervisor rdf:type #Property

 rdfs:range #Researcher

 rdfs:domain #Student

Frank rdf:type #Student

 hasSupervisor #Jeen

```
<?xml version="1.0"?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns="http://www.myspace.fake/supervisions#">

  <rdf:Description rdf:ID="Person">
    <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  </rdf:Description>

  <rdf:Description rdf:ID="Student">
    <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
    <rdfs:subClassOf rdf:resource="#Person"/>
  </rdf:Description>

  <rdf:Description rdf:ID="hasSupervisor">
    <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Property"/>
    <rdfs:range rdf:resource="#Researcher"/>
    <rdfs:domain rdf:resource="#Student"/>
  </rdf:Description>

  <rdf:Description rdf:ID="Frank">
    <rdf:type rdf:resource="#Student"/>
    <hasSupervisor rdf:resource="#Jeen"/>
  </rdf:Description>
</rdf:RDF>
```

SKOS

- Simple Knowledge Organization System
- Ontologia para descrição de vocabulários controlados, taxonomias e enciclopédias (*thesaurus*)

SKOS – Exemplo [3]

```
ex:Computer rdf:type skos:Concept;  
skos:prefLabel "Computer"@en;  
skos:prefLabel "Computador"@pt;  
skos:altLabel "Ordenador"@es ;  
skos:broader ex:Laptop ;  
skos:related ex:Software .
```

```
ex:Laptop rdf:type skos:Concept;  
skos:prefLabel "Laptop"@en;  
skos:prefLabel "Portatil"@es;  
skos:narrower ex:Computer ;  
skos:closeMatch ex:Netbook ;  
skos:exactMatch ex2:Portatil.
```

Fornecendo mais Detalhes

- Como dizemos que 2 conjuntos de indivíduos são disjuntos?
- Como restringimos cardinalidade nos relacionamentos?
- Como definimos relações inversas, transitivas?

Ontologia

- Conceitualização compartilhada de um domínio de interesse
- Visão de IA:
 - Ontologia = Teoria lógica (base de conhecimento)
- Visão de BD:
 - Ontologia = Modelo conceitual

Ontologia - Exemplo

class-def animal	% animals are a class
class-def plant	% plants are a class
subclass-of NOT animal	% that is disjoint from animals
class-def tree	
subclass-of plant	% trees are a type of plants
class-def branch	
slot-constraint is-part-of	% branches are parts of some tree
has-value tree	
max-cardinality 1	
class-def defined carnivore	% carnivores are animals
subclass-of animal	
slot-constraint eats	% that eat any other animals
value-type animal	
class-def defined herbivore	% herbivores are animals
subclass-of animal, NOT carnivore	% that are not carnivores, and
slot-constraint eats	% they eat plants or parts of plants
value-type plant OR (slot-constraint is-part-of has-value plant)	

OWL

- Padrão da W3C para representar ontologias na web
- Com ela podemos definir:
 - Classes + hierarquia
 - Propriedades, subpropriedades
 - Relações entre classes (herança, disjunção, equivalência, ..)
 - Restrições de propriedades (tipo, cardinalidade)
 - Características de propriedades (transitiva, reflexiva, ...), ...

OWL

- Esta definição permite:
 - Classificação mais precisa
 - Inferência de informações
 - Verificação de consistência
- Expressões avaliadas sob uma ontologia usualmente são descritas em DL [4]
- DL (*Description Logic*) é um fragmento da Lógica de Primeira Ordem
- Protégé [5] é o editor mais utilizado para a construção de ontologias

Estudos de Caso em Web Semântica

<http://www.w3.org/2001/sw/sweo/public/UseCas>

Referências

- [1] www.w3.org/standards/semanticweb
- [2] “Semantic Web”, Riccardo Rosati, 2012/2013
- [3] http://semanticweb.com/introduction-to-skos_b33086, Juan Sequeda
- [4] en.wikipedia.org/wiki/Description_logic
- [5] <http://protege.stanford.edu/>