

UNIVERSIDADE FEDERAL DE SANTA CATARINA

## Construindo Aplicações na Web Semântica

### Introdução à Web Semântica

---

**Renato Fileto**  
fileto@inf.ufsc.br

Programa de Pós-graduação em Ciência da Computação – PGCC  
Departamento de Informática e Estatística – INE  
Centro Tecnológico – CTC



## Tópicos

- **O que é a Web Semântica?**
  - Objetivos
  - Ontologias
- **Padrões da Web Semântica**
  - XML / XML-Schema
  - RDF / RDF-Schema
  - OWL
  - SPARQL
  - RIF



## Motivação: Fenômenos linguísticos

- **Homonímia:** mesmo nome para coisas distintas (e.g., Tornado, Lula, São Paulo, ...)
- **Sinonímia:** nomes distintos para a mesma coisa (e.g., São Paulo time de futebol, o tricolor paulista, o campeão mundial de ...)
- **Hipernímia / Hiponímia (IS\_A):** generalização / especialização (e.g., animal, mamífero, primata, ...)
- **Meronímia / Holonímia (PART\_OF):** agregação (e.g., país, estado, cidade, ...)
- **Instanciação (TYPE\_OF)** (e.g., Santo(São Paulo))
- ...

## Web semântica (Web 3)

**"A Web semântica é uma extensão da Web atual, em que a informação recebe significado bem definido, habilitando os computadores e as pessoas a trabalhar em cooperação."**

Tim Berners-Lee, James Hendler, Ora Lassila  
*The Semantic Web*, Scientific American, May 2001



## Bases da Web semântica

- **Representação formal da semântica**
  - Ontologias
  - Padrões para representar conhecimento na Web
- **Anotações semânticas**
  - Descrições de recursos (dados e serviços) baseadas em ontologias
- **Inferência (processamento do conhecimento)**
  - Máquinas de inferência (*reasoners*) e agentes inteligentes

## Estrutura da Introdução

- **Motivação: O que é a Web semântica**
  - Objetivos
  - Ontologias, anotações semânticas, inferência
- **Padrões da Web Semântica**
  - RDF / RDF-Schema, OWL, SPARQL, RIF, ...
- **Desafios da Web semântica**
  - Criação de ontologias, anotações semânticas, inferência, articulação de ontologias, ...
  - Desenvolvimento de aplicações
    - **Buscas semânticas**

## Ontologia na Filosofia

- **Filosofia grega:**
  - **Ontos = ser, logos = ciência**
  - **Conceitualização compartilhada** de algum universo de discurso ou domínio de interesse
  - **Conjunto de conceitos** (e.g. entities, attributes, processes) com suas **definições e inter-relacionamentos**.
  - **Visão unificada** para resolver problemas envolvendo significado (e.g., buscas semânticas, integração de dados, interoperabilidade de sistemas, composição de serviços).
  - Carrega uma **visão de mundo**



## Ontologias na Computação



- **Ontologia:**
  - **Conceitualização compartilhada** de um domínio
  - **Explícita e formal** para permitir o processamento por máquinas
- **Pode ter diferentes formas:**
  - **Thesaurus** com relações semânticas entre termos (e.g., sinonímia, ...)
  - **Taxonomia**
  - **Diagrama de classes**
  - **Base de conhecimento**
    - Classes, propriedades e suas relações
    - Instâncias de classes

## Compartilhada e Explícita

- **Compartilhada:** um grupo de pessoas, em um certo domínio, deve concordar sobre o significado de entidades e relações expressas em uma ontologia
- **Explícita:** uma ontologia é uma representação concreta e manipulável de (parte de) uma conceitualização
- Uma conceitualização poderia ser implícita
  - e.g., existente somente na cabeça de alguém, ou
  - embutida em um módulo de software

## Formal

- Uma ontologia [explícita!] pode ter diversas formas, mas sempre inclui um vocabulário de termos e alguma especificação de seu significado. O grau de formalização de do vocabulário e dos significados especificados pode variar consideravelmente:
  - **Altamente informal:** expressa em linguagem natural
  - **Semi-informal:** expressa em uma forma restrita e estruturada de linguagem natural
  - **Semi-formal:** expressa em uma linguagem artificial e formalmente definida
  - **Rigorosamente formal:** termos meticulosamente definidos com semântica formal, teoremas e provas de propriedades como consistência (soundness) e completude.

## Tipos de ontologias

- **Ontologia de domínio:** modela um domínio de conhecimento específico (e.g., medicina, biologia, engenharia, direito e legislação)
- **Ontologia aplicação:** modela objetos de uma aplicação (como um modelo de classes)
- **Ontologia de alto nível:** modela objetos comuns que são geralmente aplicáveis a uma grande variedade de situações (e.g., Dublin Core)
- **Meta-ontologia:** define meta-classes que determinam características de outras ontologias ou modelos em nível de abstração mais baixo



## Aplicações de Ontologias

- **Buscas semânticas**  
Ex.: Estado(São Paulo), Cidade(São Paulo)
- **Interoperabilidade**
- **Agentes móveis**
- **Reuso e composição de recursos**
- **Semantic Learning Objects**
- **Semantic Web Services**
- **Semantically enabled services & workflows**  
:

## Uses of ontologies

### Comunicação

entre pessoas e organizações

### Interoperabilidade

Entre sistemas

Componentes Reusáveis

Confiabilidade

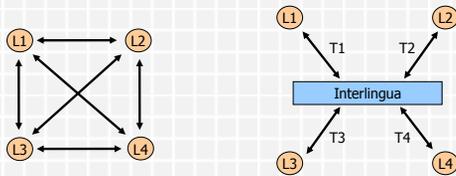
Especificação

Engenharia de Sistemas

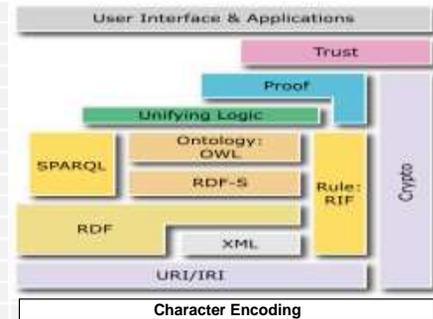
Outras distinções pode ser importantes dentro de cada categoria, tais como a natureza do software, tipo de usuários e domínio de aplicação.

## Ontologias como uma Inter-Lingua

- Ontologias pode ser usadas como suporte na tradução entre diferentes linguagens e representações de informação/conhecimento.



## Padrões da Web semântica



## Codificação de Caracteres

Exemplos de cabeçalhos de documentos XML com especificação do padrão de codificação de caracteres utilizados nesses documentos:

- `<?xml encoding="UTF-8"?'>`
- `<?xml encoding="UTF-16"?'>`
- `<?xml encoding="EUC-JP"?'>`
- `<?xml version="1.0" encoding="ISO-8859-1"?'>`

## URIs (Unified Resource Identifiers)

- <http://www.w3.org/Addressing/>
- <http://www.w3.org/Addressing/#background>
- <http://www.foldoc.org/?Uniform+Resource+Locator>
- <ftp://spy.secret@ftp.acme.com/pub/topsecret/weapon.tgz>
- <mailto:fred@doc.ic.ac.uk>
- <news:alt.hypertext>
- <telnet://192.0.2.16:80>
- <magic://a/b/c/d/ef>
- [ldap://\[2001:db8::7\]/c=GB?objectClass=one](ldap://[2001:db8::7]/c=GB?objectClass=one)
- <tel:+1-816-555-1212>
- <urn:oasis:names:specification:docbook:dtd:xml:4.1.2>
- <example://a/b/c/%7Bfoo%7D>
- <eXAMPLE://a/./b/..b/%63%7Bfoo%7d>



## XML-Schema

Um documento *XML-Schema* descreve os

- elementos,
- atributos,
- relacionamentos,
- etc.

permitidos em um ou mais documentos *XML*, isto é, define uma classe de documentos *XML* que adere a um conjunto de restrições estruturais e de dados.

- *XML-Schema* tem uma sintaxe *XML*
- *XML-Schema* é mais robusta, versátil e poderosa que *DTD (Document Type Definition)*

## Um Documento *XML-Schema*

(WaterBal.xsd)

```
<schema xmlns="http://www.w3.org/2001/XMLSchema" *
<element name="WaterBal" type="WaterBalType"/>
<attribute name="location" type="string"/>
<attribute name="latitude" type="Latitude"/>
<attribute name="longitude" type="Longitude"/>
<element name="WeatherData" type="AgregValues"/>
<complexType name="AgregValues">
  <attribute name="Date" type="DateType"/>
  <sequence>
    <element name="Temperature" type="decimal"/>
    <element name="AvgRainfall" type="decimal"/>
    <element name="PotET" type="decimal"/>
    <element name="RealET" type="decimal"/>
    <element name="Stored" type="decimal"/>
    <element name="WaterDeficit" type="decimal"/>
    <element name="WaterExcess" type="decimal"/>
  </sequence>
</complexType>
</element>
</schema>
```

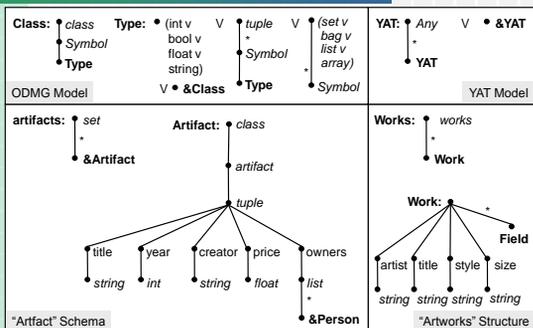
## Validação de documentos XML

- **Documento bem-formado (well-formed):** satisfaz as restrições de formação expressa na especificação do XML (<http://www.w3.org/TR/2004/REC-xml-20040204>)
- **Documento válido:** satisfaz as restrições (elementos, atributos, aninhamentos, tipos, etc.) expressas em uma especificação de esquema XML em DTD ou XSL (*XML-Schema*) que é associada a esse documento XML

## Heterogeneidade semântica em XML

```
<object id="a1" class="artifact"> <work>
<tuple> <artist> Monet </artist>
  <title> Nymphs </title> <name> Nymphs </name>
  <year> 1897 </year> <style> Impressionist </style>
  <creator> Monet </creator> <size> 21 x 61 </size>
  <price> 10,000,000 </price> <cplace> Givern </cplace>
  <owners refs="p1,p2,p3"> </work>
</tuple>
</object>
<object id="p3" class="person">
<tuple>
  <name> Claudia </name>
  <age> 17 </age>
  <history>
    Painted with
    <tech> Oil on canvas </tech>
    in ...
  </history>
</tuple>
</object>
```

## XML como Padrão de representação de dados



## RDF – Resource Description Framework

Uma linguagem e modelo padronizados para expressar conhecimento na Web semântica.

Um comando (*statement*) é uma tripla da forma:

**Recurso:** qualquer coisa referenciada por uma URL

**Propriedade:** qualquer propriedade de um recurso

**Valor:** um literal ou outro recurso

RDF-Schema define classes de recursos, suas propriedades (que podem ser relacionamentos com outras classes) e possíveis valores.

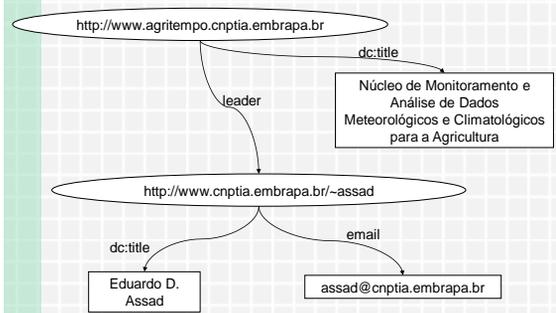
Os formatos para representar ontologias na Web semântica (e.g., DAML+OIL, OWL) são extensões do RDF.

## Sintaxe XML do RDF

```
<?xml version="1.0"?>
<rdf:RDF xmlns:RDF="http://www.w3.org/RDF/RDF/"
  xmlns:dc="http://dublincore.org/dc">
  <rdf:Description about="http://www.agritempo.cnptia.embrapa.br">
    <dc:title> Núcleo de Monitoramento e Análise de Dados
    Meteorológicos e Climatológicos para a Agricultura </dc:title>
    <leader>
      <rdf:resource="http://www.cnptia.embrapa.br/~assad"
    </leader>
  </rdf:Description>

  <rdf:Description
    about="http://www.cnptia.embrapa.br/~assad">
    <dc:title> Eduardo D. Assad </dc:title>
    <email> Assad@cnptia.embrapa.br </email>
  </rdf:Description>
</rdf:RDF >
```

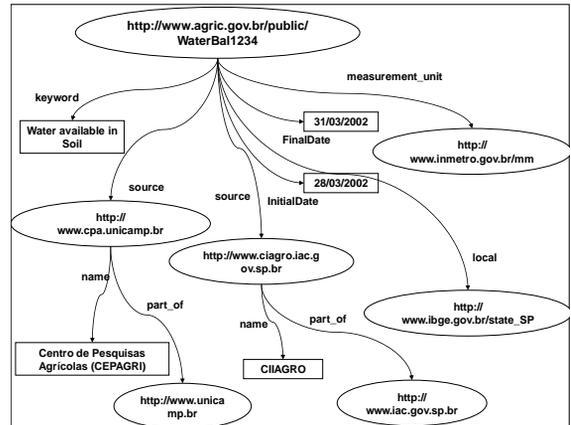
## Estrutura do conhecimento em RDF



## Metadata in RDF

Water Balance (same place and institution)

```
<?xml version="1.0"?>
<rdf:RDF xmlns:RDF="http://www.w3.org/RDF/RDF/"
  xmlns="http://agric.gov.br/DocStd/">
  <rdf:Description about="http://www.agric.gov.br/public/WaterBall1234">
    <Source>
      <rdf:resource="http://www.cepagri.unicamp.br"
      <rdf:resource="http://www.ciagro.iac.gov.sp.br"
    </Source>
    <InitialDate> 28/03/2002 </InitialDate>
    <FinalDate> 31/03/2002 </FinalDate>
    <keyword> Water available in Soil </keyword>
    <local>
      <rdf:resource="http://www.ibge.gov.br/state_SP"
    </local>
    <measurement_unit>
      <rdf:resource="http://www.inmetro.gov.br/mm"
    </measurement_unit>
  </rdf:Description>
</rdf:RDF>
```



## RDF-Schema

```
<rdf:about="&AgricZoning;Country" rdfs:label="Country">
  <rdfs:subClassOf df:resource="&AgricZoning;TerritorialDivision"/>

  <rdf:about="&AgricZoning;State" rdfs:label="State">
    <rdfs:subClassOf df:resource="&AgricZoning;TerritorialDivision"/>

  <rdf:Property rdf:about="&AgricZoning;statesOfCountry"
    a:minCardinality="1"
    rdfs:label="statesOfCountry">
    <rdfs:domain rdf:resource="&AgricZoning;Country"/>
    <rdfs:range rdf:resource="&AgricZoning;State"/>
    <a:inverseProperty rdf:resource="&AgricZoning;countryOfState"/>
  </rdf:Property>
```

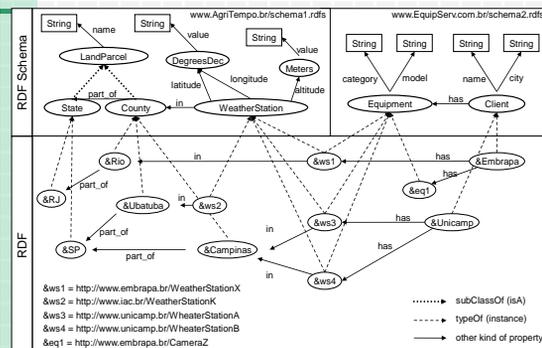
## RDF-Schema

```
<rdf:Property rdf:about="&AgricZoning;countryOfState"
  a:maxCardinality="1"
  a:minCardinality="1"
  rdfs:label="countryOfState">
  <rdfs:range rdf:resource="&AgricZoning;Country"/>
  <rdfs:domain rdf:resource="&AgricZoning;State"/>
  <a:inverseProperty
    rdf:resource="&AgricZoning;statesOfCountry"/>
</rdf:Property>
```

## Trecho de RDF

```
<AgricZoning:Country rdf:about="&AgricZoning;pais_55"
  AgricZoning:nameBR="BRASIL" rdfs:label="BRASIL">
  <AgricZoning:officialRegionsOfCountry
    rdf:resource="&AgricZoning;regof_1"/>
  <AgricZoning:officialRegionsOfCountry
    rdf:resource="&AgricZoning;regof_2"/>
  <AgricZoning:officialRegionsOfCountry
    rdf:resource="&AgricZoning;regof_3"/>
  <AgricZoning:officialRegionsOfCountry
    rdf:resource="&AgricZoning;regof_4"/>
  <AgricZoning:officialRegionsOfCountry
    rdf:resource="&AgricZoning;regof_5"/>
  <AgricZoning:statesOfCountry rdf:resource="&AgricZoning;estado_11"/>
  <AgricZoning:statesOfCountry rdf:resource="&AgricZoning;estado_12"/>
  :
  <AgricZoning:metroAreasOfCountry
    rdf:resource="&AgricZoning;metro_5201"/>
  :
</AgricZoning:Country>
```

## Descrições em RDF de recursos para a coleta de dados científicos



## OWL – Ontology Web Language

- Estende o RDF com vocabulário padronizado e construções para definir:
  - Escopo local de propriedades
  - Disjunção de classes
  - Combinações de classes
  - Restrições de cardinalidade
  - Características especiais de propriedades (e.g., transitividade, simetria, anti-simetria, propriedades inversas)
  - :

## The 3 flavors of OWL

- OWL Lite**
  - Expressividade restrita (exclui: classes enumeradas, disjunção, cardinalidade arbitrária, ...)
  - Fácil de entender e usar
- OWL DL**
  - Equivalente à lógica descritiva (*DL = Description Logics*)
  - Ainda permite eficiência e computabilidade
- OWL Full**
  - Compatível sintática e semanticamente com RDF, embora mais poderosa
  - Pode ocasionar inferências indecidíveis

## Disjunction and Equivalence of Classes

```
<owl:Class rdf:about="&associateProfessor">
  <owl:disjointWith rdf:resource="&assistantProfessor"/>
</owl:Class>

<owl:Class rdf:ID="faculty">
  <owl:equivalentClass
    rdf:resource="&academicStaffMember"/>
</owl:Class>
```

## Inverse properties

```
<owl:ObjectProperty rdf:ID="teaches">
  <rdfs:range rdf:resource="&course"/>
  <rdfs:domain rdf:resource="&academicStaffMember"/>
  <owl:inverseOf rdf:resource="&isTaughtBy"/>
</owl:ObjectProperty>
```

## Sintaxe abstrata para OWL

```

Class(Person partial
restriction (hasChild allValuesFrom(Person)))

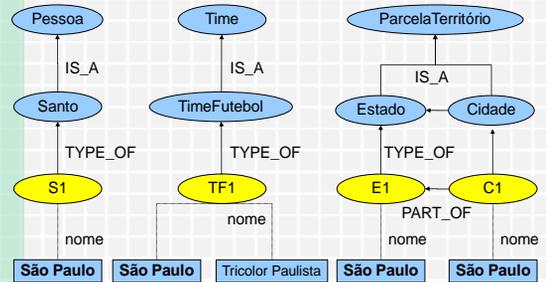
Class(Parent complete
Person
restriction (hasChild someValuesFrom(Person)))

ObjectProperty(hasChild)

Individual (John type(Person)
value(hasChild Mary))

```

## Busca em uma ontologia



## SPARQL (recomendação W3C)

Exemplo de expressão em SPARQL:

```

SELECT ?concept, ?property, "São Paulo"
WHERE {
    ?concept property:hasProperty ?property
    FILTER(?property, "name")
}

```

## Linguagens de Regras (Prolog)

### Rules:

- parent(?x, ?z) :- father(?x, ?z) v mother(?x, ?z)
- sibling(?x, ?y) :- parent(?x, ?z) ^ parent(?y, ?z)

### Knowledge base:

- father(\_Maria, \_João)
- father(\_Pedro, \_João)
- father(\_Ana, \_João)
- ...

### Query:

- sibling(\_Maria, ?x)
  - ⇒ Yes !!!
  - ⇒ ?x = \_Pedro
  - ⇒ ?x = \_Ana
  - ...

## RIF (Rule Interchange Format)

- A buyer buys an item from a seller if the seller sells it to the buyer.
- John sells LeRif to Mary.*
- The fact *Mary buys LeRif from John* can be logically derived by a modus ponens argument.

```

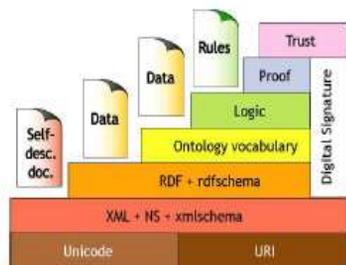
Document(
Prefix(cpt <http://example.com/concepts#>)
Prefix(ppi <http://example.com/people#>)
Prefix(bks <http://example.com/books#>)
Group ( Forall ?Buyer ?Item ?Seller (
    cpt:buy(?Buyer ?Item ?Seller) :
    cpt:sell(?Seller ?Item ?Buyer)
)
    cpt:sell(ppi:John bks:LeRif ppi:Mary)
)
)

```

## Controvérsia sobre pilha de padrões

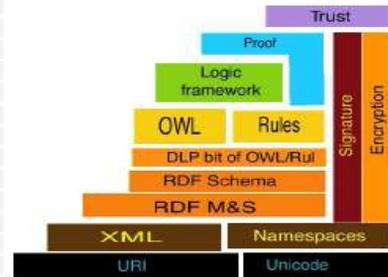
- Peter F. Patel-Schneider. A Revised Architecture for Semantic Web Reasoning. Third Workshop on Principles and Practices of Semantic Web Reasoning. Dagstuhl, Germany, September 2005. LNCS 3703, Springer Verlag, 2005.
- Ian Horrocks, Bijan Parsia, Peter F. Patel-Schneider, and James Hendler. Semantic Web Architecture: Stack or Two Towers? Third Workshop on Principles and Practices of Semantic Web Reasoning. Dagstuhl, Germany, September 2005. LNCS 3703, Springer Verlag, 2005.
- Peter F. Patel-Schneider. Building the Semantic Web Tower from RDF Straw. Nineteenth International Joint Conference on Artificial Intelligence. Edinburgh, Scotland, August 2005.

## Proposta inicial (2000)



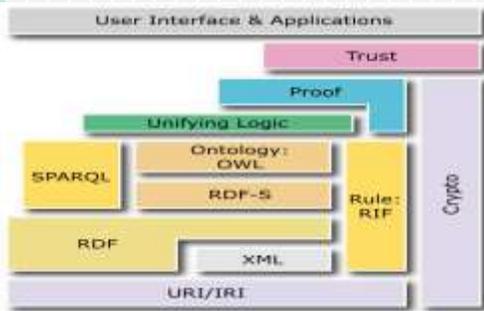
<http://www.w3.org/2000/Talks/1206-xml2k-tbl/>

## Proposta regras(2005)



<http://www.w3.org/2005/Talks/0511-keynote-tbl/>

## Proposta atual



## Como está a Web semântica hoje?

- Área de pesquisa muito ativa
- Visa estender o papel dos computadores no suporte a diversas atividades humanas
- Usa ontologias e anotações semânticas para catalogar, recuperar e compor dados e serviços de processamento na Web

## Perguntas?

