

# Buscas semânticas: panorama e tendências atuais

minicurso ERBD 2010

Renato Fileto  
fileto@inf.ufsc.br

Alunos Colaboradores: Caio Stein D'Agostini  
csd@inf.ufsc.br  
Karina Fasolin  
kfasolin@inf.ufsc.br  
Wanderson Rigo  
wander@inf.ufsc.br



Programa de Pós-graduação em Ciência da Computação - PPGCC  
Universidade Federal de Santa Catarina - UFSC

## Conteúdo

- **Motivação**
- **Web semântica**
  - Proposta
  - Padrões
  - Desafios
- **Buscas semânticas**
  - Definição do problema
  - Abordagens e ferramentas
  - Tendências
- **Conclusões**



## 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)**
  - 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

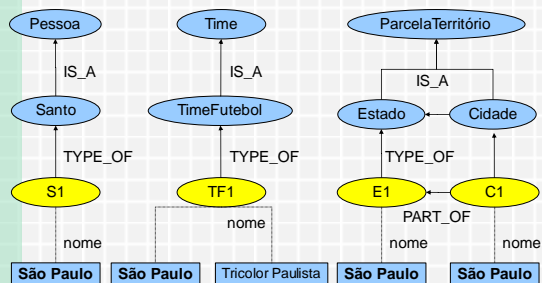
## Compartilhada e Explícita

- **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.

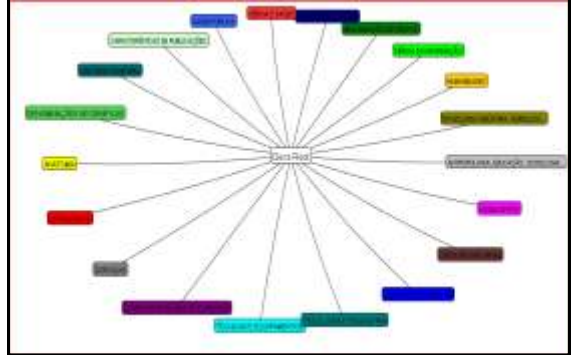
## Trecho de ontologia 1



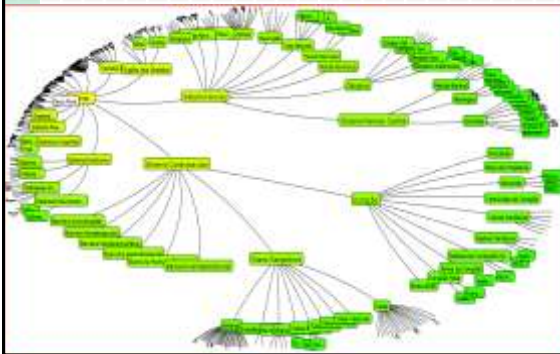
### Trecho de ontologia 2: Travel



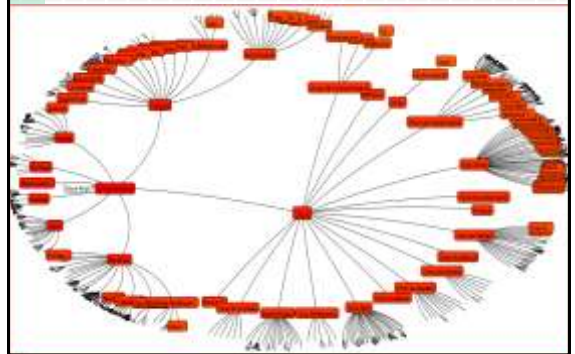
### Exemplo de ontologia 3: DeCS (Descritores em Ciências da Saúde)



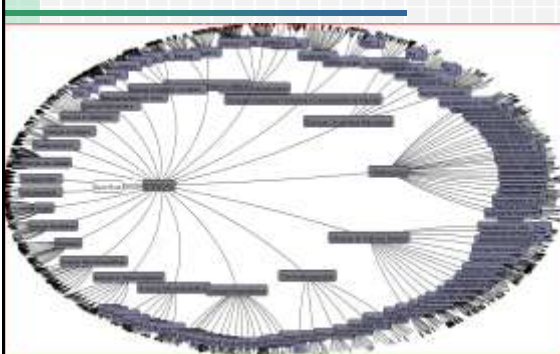
### Exemplo de ontologia 3a – DeCS (anatomia)



### Exemplo de ontologia 3b – DeCS (doenças)



### Exemplo de ontologia 3c – DeCS (doenças)



### 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

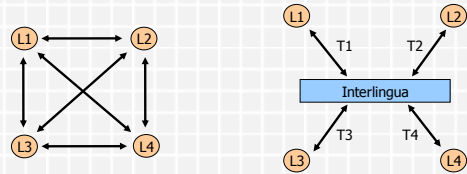
**Interoperabilidade**  
Entre sistemas

Componentes Reusáveis      Confiabilidade  
Especificação  
**Engenharia de Sistemas**

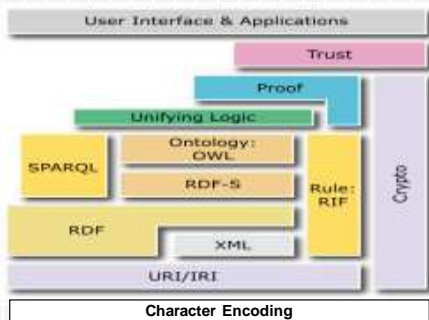
Outras distinções pode ser importantes dentro de cada categoria, tais como a natureza do software, tipo de usuários e domínio de aplicação.

## 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/elf>
- [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>

## NameSpaces

Exemplos de trechos de documentos XML com especificações e uso de namespaces:

- `<x xmlns:edi="http://ecommerce.org/schema">`  
`<!-- the "edi" prefix is bound to http://ecommerce.org/schema`  
`for the "x" element and contents -->`  
`</x>`
- `<book xmlns="urn:loc.gov:books"`  
`xmlns:isbn="urn:ISBN:0-395-36341-6">`  
`<title>Cheaper by the Dozen</title>`  
`<isbn:number>1568491379</isbn:number>`  
`</book>`
- `<schema xmlns:"http://www.w3.org/2001/XMLSchema">`  
`:`  
`<schema>`

## XML – eXtensible Markup Language

### Construções básicas:

- <Elemento> Conteúdo </Elemento>
- <Elemento />
- <Elemento Atributo1="Valor" Atributo2="Valor"/>
- <Elemento Atributo1="Valor" Atributo2="Valor"> Conteúdo </Elemento>
- <Elemento1> <Elemento2> Conteúdo </Elemento2> </Elemento1>
- <Elemento1><Elemento2/></Elemento1>

## Um documento XML

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<ElementoRaiz>
  <Elemento1>
    <Elemento1_1> Conteúdo </Elemento1_1>
    <Elemento1_2> Conteúdo </Elemento1_2>
    <Elemento1_3/>
  </Elemento1>
  <Elemento2>
    <Elemento2_1>
      <Elemento2_1_1> Conteúdo </Elemento2_1_1>
      <Elemento2_1_2/>
    </Elemento2_1>
  </Elemento2>
</ElementoRaiz>
```

## A Versatilidade do XML para Dados Semi-Estruturados

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<Pessoa Nome="João de Souza" DataDeNascimento="12/06/1964" CPF="12312312">
  <Enderecos>
    <Endereco Tipo="Comercial">Rua A 23 - Centro</Endereco>
    <Endereco Tipo="Residencial">Torre Comercial 1923 Sl. 1011</Endereco>
    <MalaDireta Destino="Comercial"/>
  </Enderecos>
  <FormasDeContato>
    <Telefones>
      <Telefone Tipo="Comercial" DDD="61" Ramal="211">222-8545</Telefone>
      <Telefone Tipo="Celular" DDD="61">999-9999</Telefone>
      <Telefone Tipo="Residencial" DDD="61">444-4444</Telefone>
    </Telefones>
    <Email>joão@email.com.br</Email>
    <IQ Numero="127653467654" />
  </FormasDeContato>
</Pessoa>
```

## Doc. XML para Dados Climáticos (Balanço Hídrico)

```
<?xml version="1.0" encoding="ISO8859-1" ?>
<WaterBal xmlns="http://www.agric.gov.br/WaterBal.xsd"
  location="Brotas" latitude="-22.1500 longitude=-47.5800">
  <WeatherData Date="20050823">
    <Temperature> 22.0 </Temperature>
    <AvgRainFall> 201.3 </AvgRainFall>
    <PotET> 115.4 </PotET>
    <RealET> 115.4 </RealET>
    <Stored> 125.0 </Stored>
    <WaterDeficit> 0.0 </WaterDeficit>
    <WaterExcess> 86.0 </WaterExcess>
  </WeatherData>
  :
</WaterBal>
```

## Arquivo texto (não XML) com dados pluviométricos

weather station	latitude	longitude	altitude
674200	185300	874200	1207.0
year			
1994	117	80	0
1995	0	0	0
1996	0	0	0
1997	0	0	0
1998	0	0	0
1999	0	0	0
2000	0	0	0
2001	0	0	0
2002	0	0	0
2003	0	0	0
2004	0	0	0
2005	0	0	0
2006	0	0	0
2007	0	0	0
2008	0	0	0
2009	0	0	0
2010	0	0	0
2011	0	0	0
2012	0	0	0
2013	0	0	0
2014	0	0	0
2015	0	0	0
2016	0	0	0
2017	0	0	0
2018	0	0	0
2019	0	0	0
2020	0	0	0

## 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

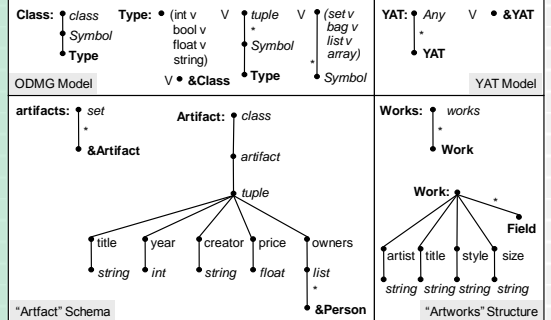
```
<object id="a1" class="artifact">
  <tuple>
    <title> Nymphs </title>
    <year> 1897 </year>
    <creator> Monet </creator>
    <price> 10,000,000 </price>
    <owners refs="p1,p2,p3"/>
  </tuple>
</object>

<object id="p3" class="person">
  <tuple>
    <name> Claudia </name>
    <age> 17 </age>
  </tuple>
</object>
```

```
<work>
  <artist> Monet </artist>
  <name> Nymphs </name>
  <style> Impressionist </style>
  <size> 21 x 61 </size>
  <cplace> Givern </cplace>
</work>

<work>
  <artist> Monet </artist>
  <title> Waterloo Bridge </title>
  <style> Impressionist </style>
  <size> 29.2 x 46.4 </size>
  <history>
    Painted with
    <tech> Oil on canvas </tech>
    in ...
  </history>
</work>
```

## 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.

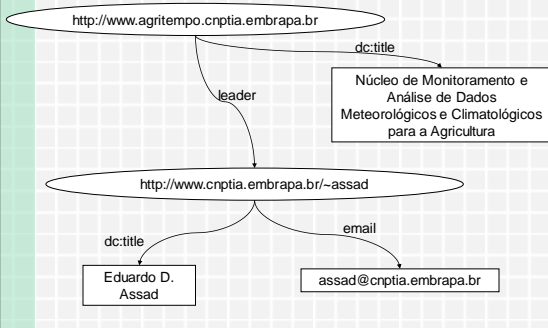
Os formatos para representar ontologias na Web semântica (e.g., DAML+OIL, OWL) são extensões do RDF.

## 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 >
```

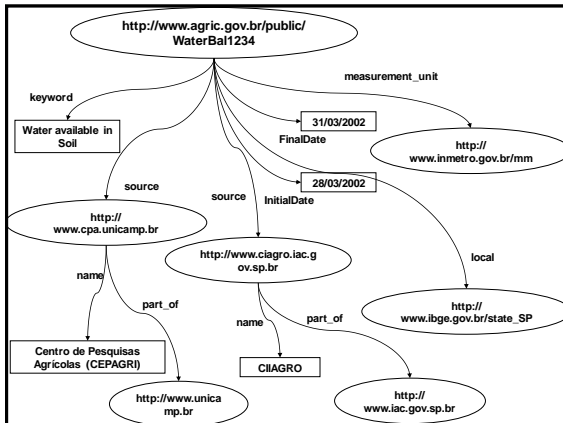
## Estrutura do conhecimento em RDF



## 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/WaterBal1234">
  <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>
```

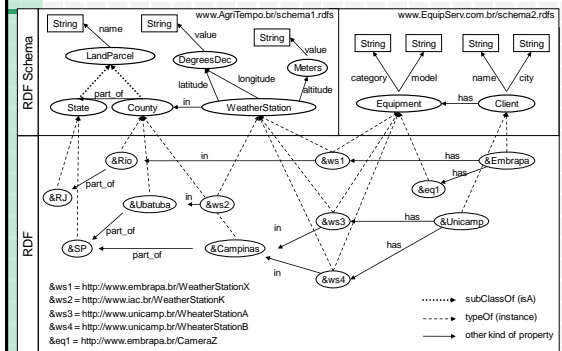
## RDF-Schema

```
<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)))

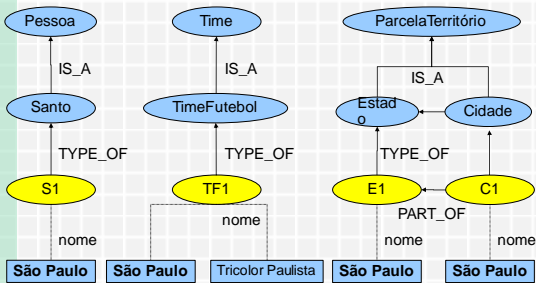
Class(Parent complete
  Person
  restriction (hasChild someValuesFrom(Person)))

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)
- ...

### Query:

- sibling(\_Maria, ?x)
  - ⇒ Yes !!!
  - ⇒ ?b = \_Pedro
  - ⇒ ?b = \_xyz1
  - ⇒ ?b = \_xyz2
  - ⇒ ?b = \_xyz3
  - ⇒ ?b = \_xyz4
  - ...

## 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(

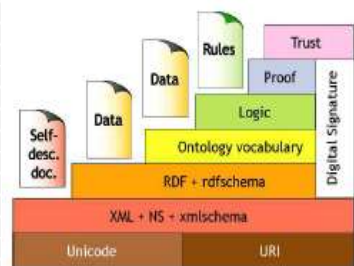
```

Prefix(cpt <http://example.com/concepts#>)
Prefix(ppl <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(ppl:John bks:LeRif ppl:Mary)
)
)
  
```

## Controvérsia sobre pilha de padrões

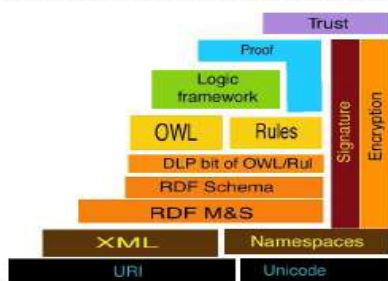
- Peter F. Patel-Schneider. A Revised Architecture for Semantic Web Reasoning. Third Workshop on Principles and Practices of Semantic Web Reasoning. Dagstuhl, Germany, September 2005. LNCS 3703, Springer Verlag, 2005.
- Ian Horrocks, Bijan Parsia, Peter F. Patel-Schneider, and James Hendler. Semantic 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)



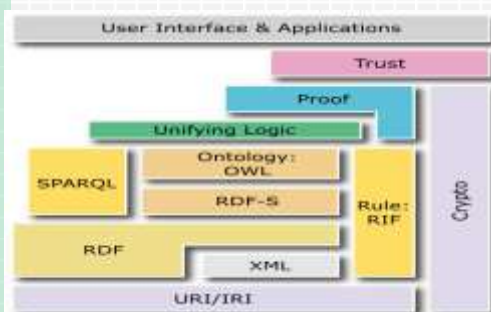
<http://www.w3.org/2000/Talks/1206-xml2k-tbl/>

## 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

## Desafios da Web semântica

- **Construção de ontologias**
  - (Semi-)automatizada e interativa
- **Geração de anotações semânticas de recursos**
  - Muitas aplicações requerem automatização
- **Evolução de ontologias**
  - Checar se alterações não introduzem inconsistências e gerenciar versões
- **Integração de ontologias (*ontology matching*)**
  - A obtenção de consenso é inviável em muitos domínios
- **Dedução/raciocínio**
  - Teorias computáveis e não computáveis; ontologias inconsistentes, ...

## Edição de Ontologias no Protégé

- Editor de ontologias e arcabouço para o desenvolvimento de bases de conhecimento livre e de código aberto baseado em Java.
- Suporta os modelos de *frames* (RDF) e OWL.
- Exporta as ontologias e bases de conhecimento para uma variedade de formatos, incluindo RDF(S), OWL e XML Schema.
- Extensível via *plugins*.

## Construção de ontologias

- **Processo para geração automatizada de ontologias a partir de esboços, dados e documentos** (nova geração de ferramentas)
  - Entendimento do domínio
  - Entendimento dos dados
  - Definição das tarefas
  - **Geração semi-automática**
  - Avaliação de qualidade
  - Refinamento com interação humana

## Definições de "geração de ontologia"

Grobenik and Mladenic, 2006

- **Indução de conceitos** (a partir das instâncias)
- **Indução de relações** (a partir de conceitos e instâncias a eles associadas)
- **Popular ontologia** (dada uma ontologia e instâncias de seus conceitos)
- **Geração de uma ontologia completa** (a partir das instâncias e algum conhecimento/informação)
- **Atualização/extensão de ontologias** (dada uma ontologia e informação como novas instâncias ou padrões de uso da ontologia)

## Document Clustering - Vector Model

### Vector of word occurrences in the documents:

$$d_{jk} = TF(W_k, d_j) * IDF(W_k)$$

#### Where

- $TF(W_k, d_j)$  = number of times  $W_k$  occurs in document  $d_j$
- $IDF(W_k) = \log ( D / DF(W_k) )$
- $D$  = number of documents
- $DF(W_k)$  = number of documents in which word  $W$  occurs

### Document distance metric:

$$\cos(d_i, d_j) = \frac{\sum_k d_{ik} d_{jk}}{\sqrt{\sum_l d_{il}^2} \sqrt{\sum_m d_{jm}^2}}$$



Figure 2.1 An example output of a system for graph-based visualization of document collection. The documents are 1700 descriptions of European research projects in information technology (SFP IST).

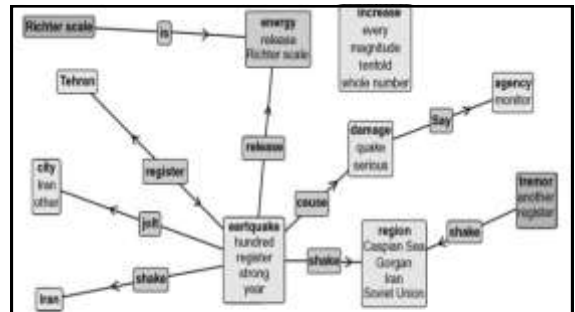


Figure 2.2 Visual representation of an automatically generated summary of a news story about earthquake. The summarization is based on deep parsing used for obtaining semantic graph of the document, followed by machine learning used for deciding which parts of the graph are to be included in the document summary.

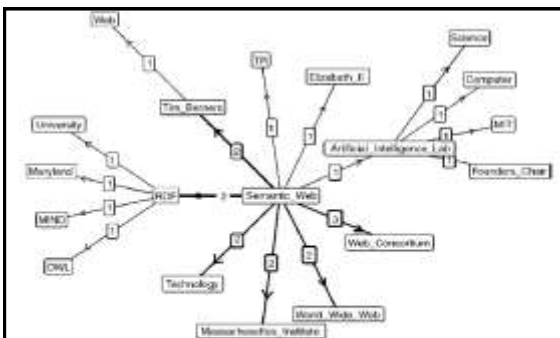


Figure 2.3 Visual representation of relationships (edges in the graph) between the named entities (vertices in the graph) appearing in a collection of news stories. Each edge shows intensity of co-mentioning of the two named entities. The graph is an example focused on the named entity 'Semantic Web' that was extracted from the 11,000 ACM Technology news stories from 2000 to 2004.

## Anotações Semânticas

(Kiryakov et al. 2004)

IBM announced profits in US, planning to build a \$120M plant in Bulgaria... and more and more and more and more text...



## Anotação Semântica de Documentos

- Atribui às entidades que aparecem no documento ligações com suas descrições semânticas na ontologia
- Aplicável a qualquer tipo de texto (documentos HTML, documentos de texto comuns, campos de banco de dados, entre outros)

### Representação das anotações

- Intrusiva
- Não intrusiva

## Confecção das anotações

- **Semi-Automática:** Associa palavras do texto a classes, instâncias e propriedades da ontologia utilizando-se do julgamento humano. Esta associação geralmente é efetuada através de interfaces "arraste-e-solte" (e.g. *OntoMat*)
- **Automática:** Aplica técnicas de processamento de linguagem natural, aprendizado de máquina e extração de informação, entre outras, para associar palavras à ontologia. Essas ferramentas podem utilizar ontologias padrão (e.g. *IEEE Standard Upper Ontology*) ou ontologias de domínios específicos (e.g. *Unified Medical Language System*)
- **Híbrida:** Combinação das técnicas acima

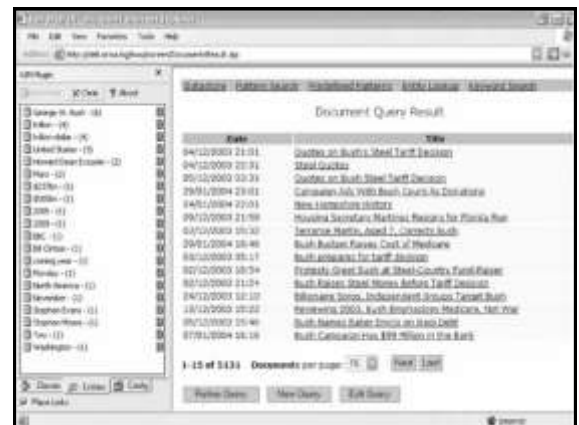
## Geração de anotações semânticas

(Reeve e Han, 2005)

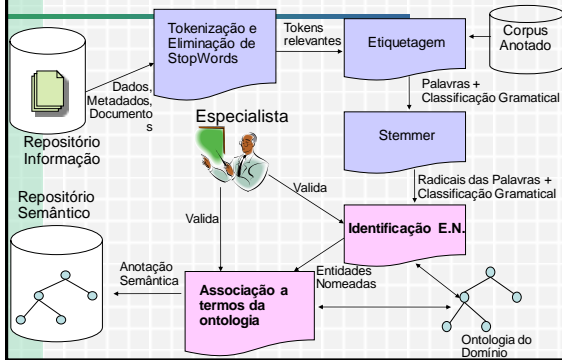


## Ferramentas para anotação

- **Annozilla:** Funciona sobre o Mozilla Firefox
- **Annotea:** projeto da W3C - servidor de anotações em RDF
- **Armadillo:**
- **Muse:**
- **Pankol:**
- **Semantic Word:** anotação interativa de porções de documentos Word
- **OntoMat Annotizer:** anotação interativa de páginas Web com ontologias em OWL
- **Smore:** marcações sobre HTML com ontologias em OWL
- **Melita:** ferramenta de anotação semi-automática que usa algoritmo de IE para ajudar o usuário no processo de anotação
- **MnM:** anotação semântica de páginas Web com metadados descritos em RDF/OCML/DAML+OIL
- **KIM:** permite criar metadados, armazenar e buscar informações e pode ser usada para anotação e procura de documentos



## O processo automatizado de anotação semântica de documentos



## Annotation example

(Davies et al. 2006)

Consider the phrase:  
There are deliverables.

The tokens are:  
[There] [are] [deliverables] [.]

The tagger finds out what kind of a word each of the tokens is (whether it is a noun, an adjective, a verb, etc.)

[There]: existential quantifier

[are]: verb—3rd person singular present

[deliverables]: noun—plural

## Annotation example (cont.)

The morphological analyser gives the roots of all the words.

[There]: root—there

[are]: root—be

[deliverables]: root—deliverable

There are — — —.

triggers the creation of one or more new classes in the ontology.

## Annotation example (cont.)

There are projects.

There are workpackages, tasks and deliverables.

SEKT is a project.

MUSING, 'Knowledge Web', and 'Presto Space' are projects.

Projects have workpackages.

Workpackages can have tasks.

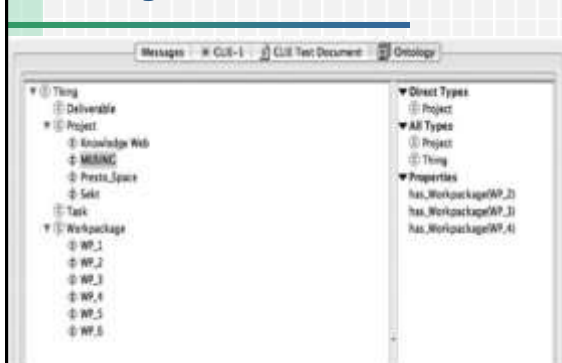
WP1, WP2, WP3, WP4, WP5 and WP6 are workpackages.

SEKT has WP1.

MUSING has WP2, WP3 and WP4.

'Knowledge Web' has WP5 and WP6.

## Ontologia resultante



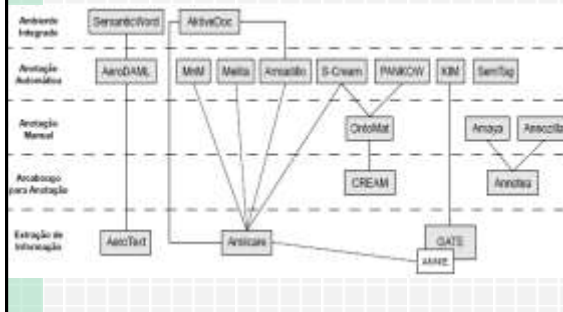
## Quadro Comparativo

(Nunes e Fileto, 2007)

Nome da Ferramenta	Tipo da Ferramenta	Anotações Geradas	Processo Anotação	Evolution Ontologia	Formato Documento
Cent.Mat. Annotator	Semio-automática	CWL	Interativo	Sim	HTML
Semantic Word	Semio-automática	DAML+OIL	IE (Aer/CAML)	-	Word (DOC)
Annotica	Semio-automática	RDF	Interativo	-	HTML
Share	Semio-automática	CWL	OBIE	Sim	HTML, text, email, images
KIM	Automática	RDF+OWL	OBIE (GATE)	Sim	HTML
Midas	Semio-automática	RDF+OIL+OIL	IE (Aer/CAML)	Sim	HTML, text
MAAM	Híbrida	CWL	OBIE (Annotator)	Sim	HTML, text
Criador de Anotação Semântica para o processo de autoria	Semio-automática	CWL	Interativo	Não	Writer (DOC)

## Dependências entre plataformas

(Rui Tramontin Junior, 2007)



## Comparação de desempenho

(Reeve e Han 2005)

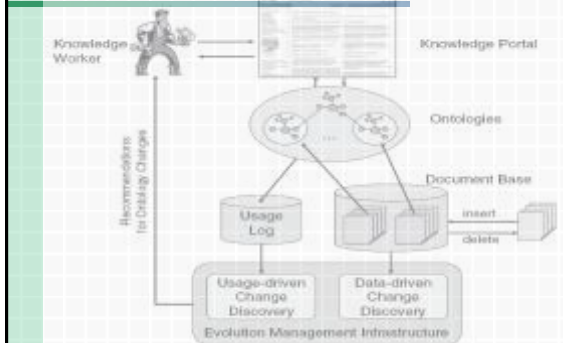
(Reeve e Han 2005)

Plataforma	Categoria	Precisão	Cobertura	Medida F
Armadillo	Descoberta de padrões	91.0	74.0	87.0
KIM	Regras	86.0	82.0	84.0
MnM	Indução	95.0	90.0	n/d
MUSE	Regras	93.5	92.3	92.9
PANKOW	Descoberta de padrões	65.0	28.2	24.9
SemTag	Regras	82.0	n/d	n/d

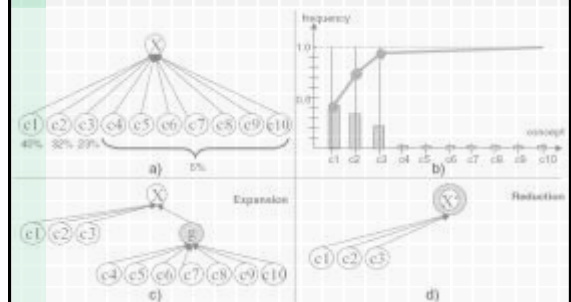
## Evolução de ontologias

(Bloehdorn et al. 2006)

(Bloehdorn et al. 2006)



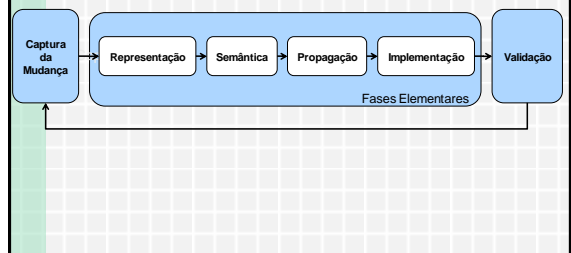
## Exemplo: Uso não uniforme de conceitos



## Alternativas para remoção do conceito C que tem sub-conceitos



## Processo de evolução de ontologias



## Mapeamentos entre ontologias (Ontology Matching)

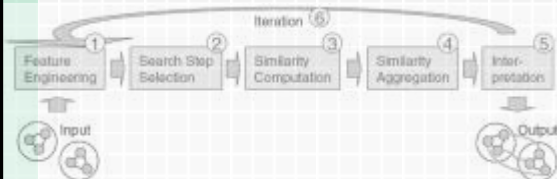


## Fusão de ontologias



## Processo semi-automático de alinhamento

(Huang et al., 2006)



## Referência

Jerome Euzenat, INRIA Rhone-Alpes, France,  
Pavel Shvaiko, University of Trento, Italy.

### Ontology Matching

URL: <http://book.ontologymatching.org>  
Springer-Verlag, Heidelberg (DE), 2007  
343 p., 67 illus., Hardcover  
ISBN: 978-3-540-49611-3

## E as buscas semânticas?

### ● Recuperação da informação de documentos

### ● Medidas da eficácia das soluções

$$\text{precisão} = \frac{\text{nro. de docs relevantes recuperados}}{\text{nro. de docs recuperados}}$$

$$\text{cobertura} = \frac{\text{nro. de docs relevantes recuperados}}{\text{nro. de docs relevantes}}$$

## Classificação de abordagens para buscas semânticas

(Mangold, 2007)

- Arquitetura
- Acoplamento entre ontologias e documentos
- Transparência
- Contexto do usuário
- Modificação de consultas
- Estrutura das ontologias
- Tecnologia para representar as ontologias
- Desempenho, escalabilidade e distribuição
- Adaptabilidade
- Ranking

## Arquiteturas para buscas semânticas

- **Máquina de busca *stand alone***: armazena metadados sobre os documentos em índices com os quais avalia as consultas
- **Meta-máquina de busca**: distribui as consultas para outras máquinas de busca e combina os resultados

## Acoplamento ontologias-documentos

- **Acoplamento forte**: os metadados dos documentos referem-se explicitamente aos conceitos de uma ontologia e vice-versa (requer anotação semântica)
- **Acoplamento fraco**: os documentos são independentes de quaisquer ontologias disponíveis (pode restringir a capacidade de busca)

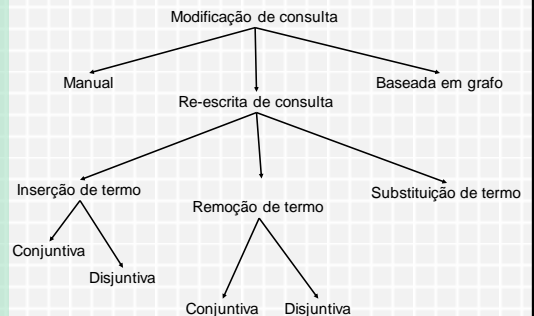
## Contexto do usuário

- **Learning**: extraído na interações do usuário com o sistema (histórico das consultas e refinamentos de consultas)
- **Hard-coded**: as consultas são organizadas em categorias que definem a informação solicitada (e.g., "location of", "resources for" ...)

## Modificação de consultas

- **Manual**: o sistema retorna um trecho apropriado da ontologia junto com a resposta, permitindo ao usuário navegar na ontologia e reformular a consulta pela adição e remoção de termos
- **Re-escrita automática**: a consulta é otimizada automaticamente pelo sistema
- **Baseada em grafo**: considera conceitos e documentos como nodos de um grafo, efetuando propagação da consulta a partir dos nodos correspondentes a termos de busca

## Modificação de consultas





## Estrutura da ontologia

- **Propriedades anônimas:** as relações entre conceitos não são nomeadas, indicando apenas que eles compartilham o mesmo contexto
- **Propriedades padronizadas:** sinônimia, hipernímia, meronímia, negação, etc.
- **Propriedades específicas de domínio:** além das propriedades padrão, considera propriedades específicas de um domínio (e.g., "tipo de câmera" em sistema para a área de fotografia)

## Tecnologia da ontologia

- F-Logic
- RDF
- DAML+OIL
- Frames
- OWL
- :

## Comparação de abordagens

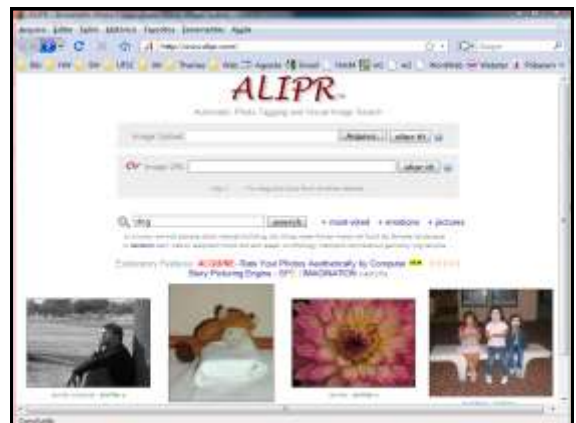
Prototype/preparator	SWOL	Inquiral	ZAP	Minimal grounding approach	ZSA	Librarian agent
By:	Hollan and Rescher (2000)	Stevan et al. (2001)	Doko et al. (2005)	Aschke et al. (2004)	Burton-Jones et al. (2001)	Ingwersen (2000)
Focus:	WWW	WWW	WWW	WWW	WWW	WWW
Architecture:	Stand-alone	Meta	Meta	Stand-alone	Meta	Stand-alone
Coupling:	Tight	-	Loose	Tight	Loose	Tight
Transparency:	Interactive	Transparent	Hybrid	Transparent	Hybrid	Interactive
User control:	None	Hard-coded	Hard-coded and learning	None	None	Learning
Query evaluation:	Manually	Comp. supp.	Developer-part, graph-based, Database-part, manually	Graph-based	All sets of query rewriting	Comp. reasoning, Comp. supp., Substitution
Ontology structure:	Hypertext, Anonymous	-	Anonymous	Domain specific	Hypertext, Synonym, Negation	Anonymous
Ontology technology:	Proprietary	-	RDF	Unclear	ZINC concepts + Friend set	Proprietary

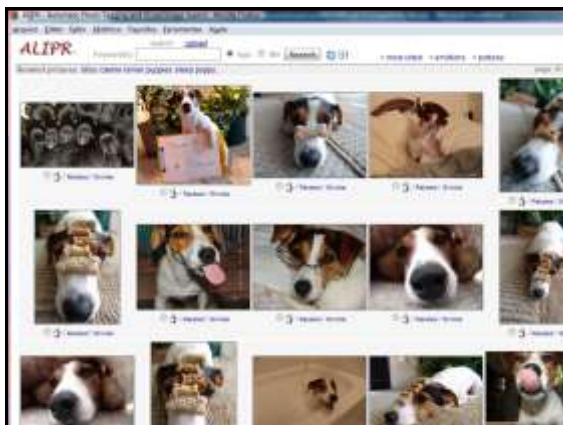
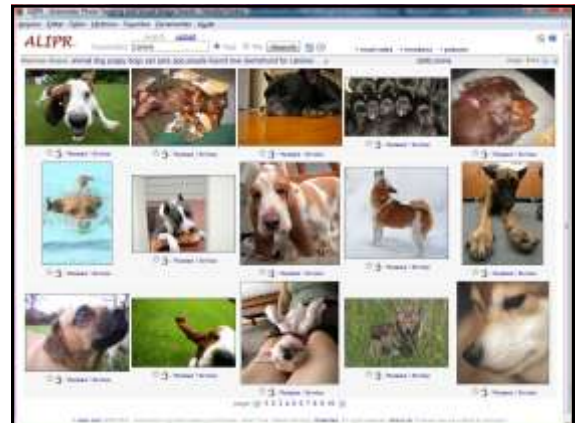
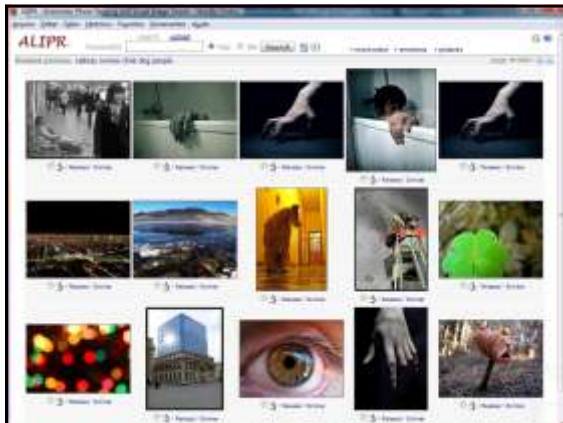
## Comparação de abordagens (cont.)

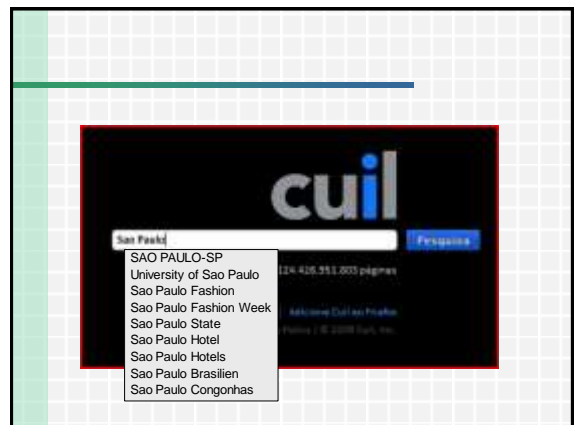
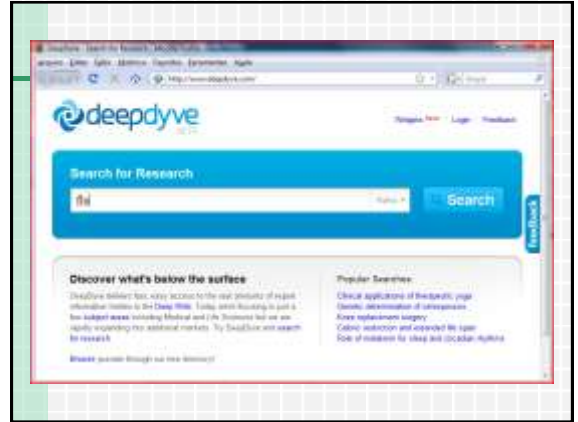
Prototype/preparator	SCORE	TRUST	Andie data-oriented	Designator	Semantically more effective
By:	Shen et al. (2001)	Jamnik et al. (2004)	Klein et al. (2000)	Wolstetter et al. (2001)	-
Focus:	Information systems	Information systems	Is for audio data retrieval	Is for image retrieval	-
Architecture:	Stand-alone	Hybrid	Stand-alone	Stand-alone	Open
Coupling:	Tight	Hybrid	Tight	Tight	Hybrid
Transparency:	Interactive	Unclear	Transparent	Interactive	Hybrid
User control:	Unclear	Hard-coded	None	Hard-coded	Hard-coded and learning
Query modification:	Manually	Comp. supp.	Dev. supp., Substitution	Substitution	Open
Ontology structure:	Hypertext, Domain specific	Hypertext, Synonym, Negation	Hypertext, Synonym, Negation	Hypertext, Synonym, Domain specific	Facilitation, use of AI
Ontology technology:	Unclear	Proprietary	Proprietary	Proprietary	Standard(s)

## Buscas semânticas na Web hoje

- <http://www.alipr.com> (imagens)
- <http://www.cognition.com>
- <http://www.deepdive.com>
- <http://www.bing.com>
- <http://www.cuil.com>
- <http://www.freebase.com>
- <http://www.google.com>
- <http://www.kosmix.com>
- <http://www.hakia.com>
- <http://www.powerset.com>
- <http://www.sensebot.net>
- <http://developer.yahoo.com/searchmonkey>
- <http://swoogle.umbc.edu> (ontologias)



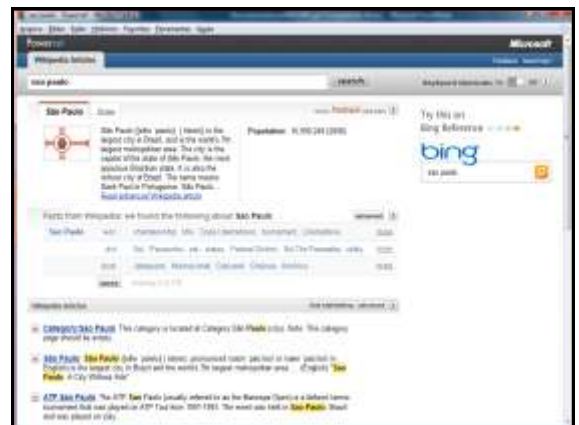


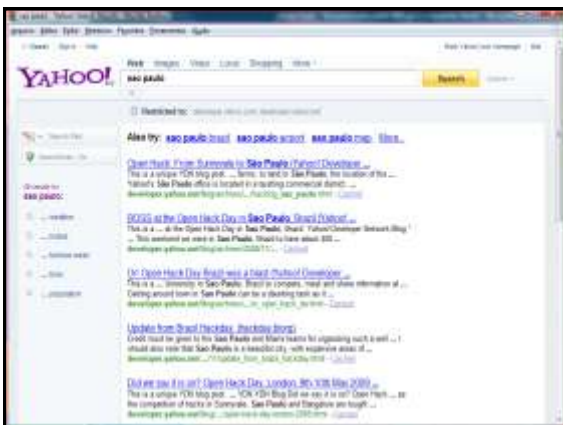






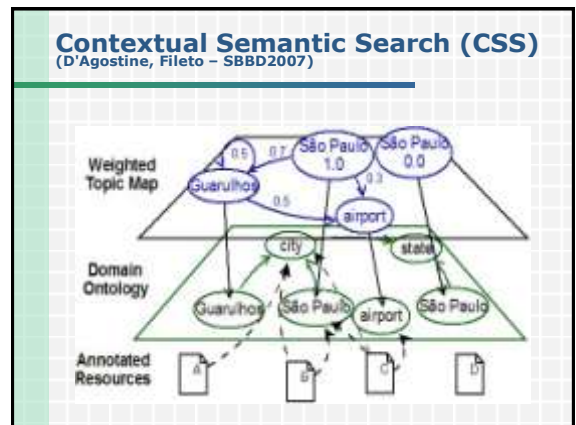
Sao Paulo Airport  
Sao Paulo Plane Crash  
Sao Paulo Zoo



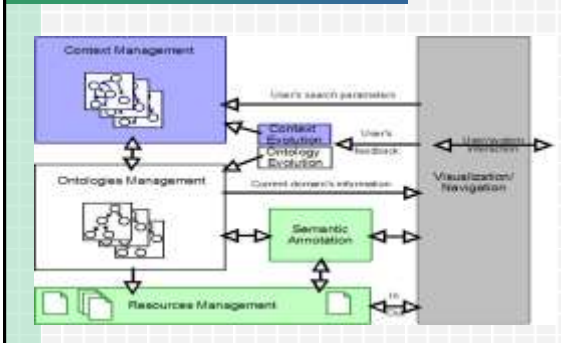




- ### Algumas questões em aberto
- Representação de contextos de usuários e mapeamento desses a ontologias
  - Análise dos efeitos de modificações de consultas
  - Meta-buscas semânticas
  - Análise da aceitação dos usuários
  - Adaptabilidade a diferentes ontologias
  - Ranking dos resultados
  - Integração com sistemas de gerenciamento de documentos/conteúdo
  - Tratamento de dados multimídia
  - Interfaces homem-máquina
  - Desempenho e escalabilidade



## General Architecture for CSS



## The Weighted Topic Graph

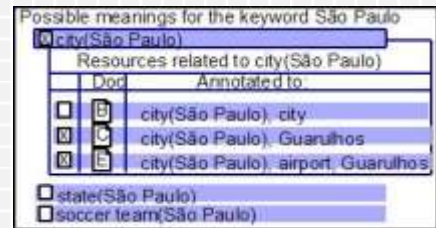
(D'Agostine, Fileto, Dantas, Gauthier - ICEIS 2008)

- A weighted topic graph is a graph  $TM(T, A)$ 
  - $T$  is a set of topics (vertices)
  - $A$  is a set of associations (edges)
- Let  $\mathcal{O}$  be an ontology
  - Each topic  $t$  in  $T$  corresponds to a ontology term  $o$  in  $\mathcal{O}$
- Each topic  $t$  has a weight  $[0,1]$ . The sum from the weights from all topics with same name equals 1
- Each association  $a$  has a weight  $[0,1]$ . The sum from all the weights from all association departing a single topics equals 1

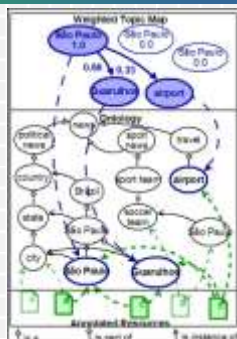
## The Weights in the Topic Graph

- **Topic weights:** used to disambiguate
  - E.g., interest in *São Paulo* city instead of *São Paulo* state
- **Association weights:** used to semantically expand searches
  - E.g., Actual interest in Guarulhos and airport while searching for *São Paulo* city

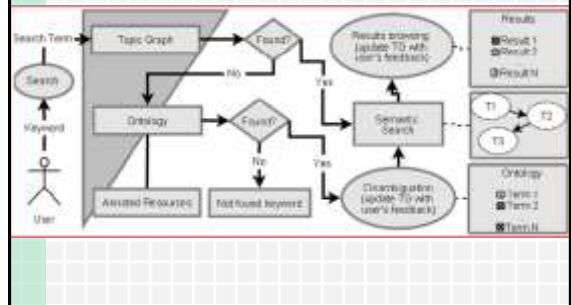
## Capturing the User's Context



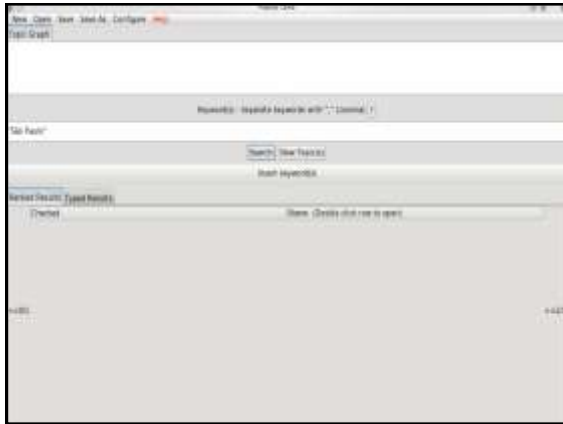
## Generated Topic Graph with the Ontological User's Context



## The Contextual Semantic Search Process





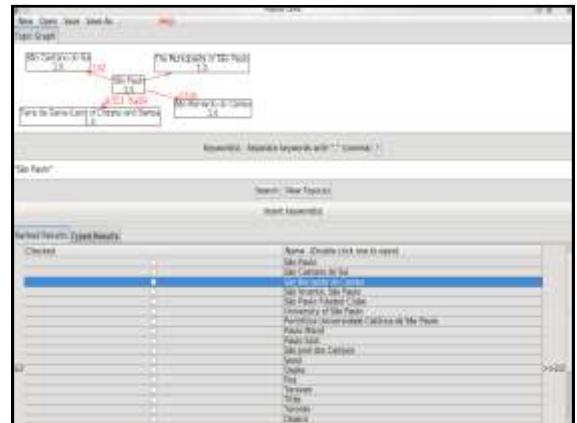
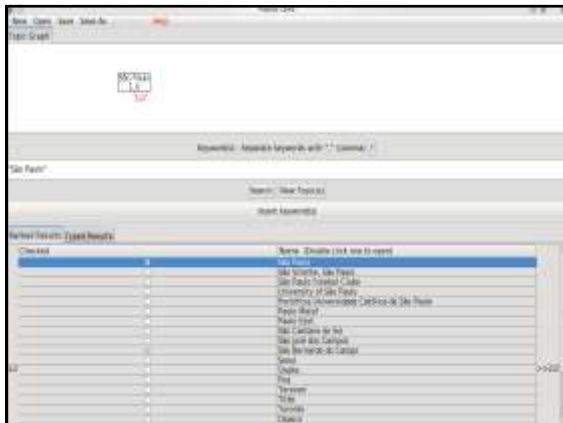


## Desambiguação

Choose the intended meaning(s) for the keyword "São Paulo"

- São Paulo Futebol Clube
- SoccerClub
- SportsTeam
- Organisation
- Resource
- São Paulo
- Area
- PopulatedPlace
- Place
- Resource

OK



## Fundamento dos algoritmos de busca sobre contexto ontológico

Heurísticas de estigmergia (*Ant Colony Optimization - ACO*)

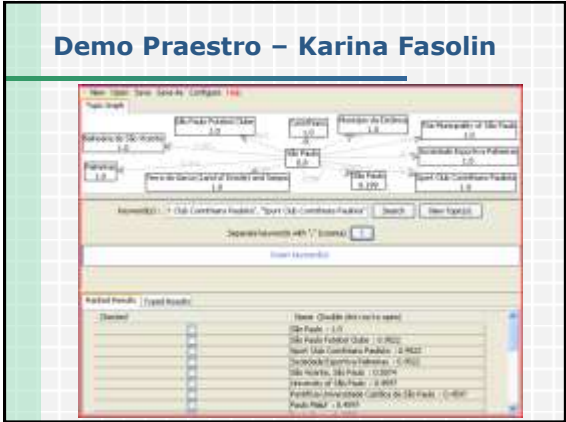
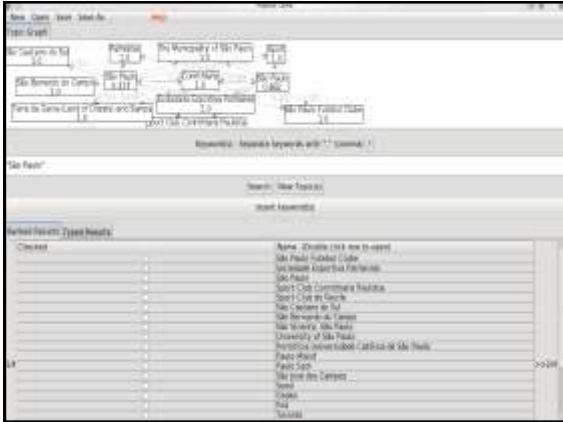
## Algoritmo de busca semântica baseada em ACO

Entrada: palavras-chave[ ];

Dados: G(T,A), Onto, repositório,  
 limite\_de\_profundidade,  
 limite\_mínimo\_de\_peso,  
 $\lambda$  // fator de atenuação de peso;

início Algoritmo Busca Contextual 1 Estigmérica  
 resultados = busca\_G\_ou\_Onto(palavras-chave[]);  
 atenuaG(indice) ; //  $\lambda \in [0, 1]$   
 manutenção(coleta\_feedback(resultados));

fim



## Demo Praestro – Karina Fasolin

### Alguns trabalhos relacionados

Name	Article	Comments
Graupmann et al. 2005	The SphereSearch engine for unified ranked retrieval of heterogeneous XML and web documents	Searches through XML around given element, also to both contexts. Made tests with stress testing tool. benchmark for this kind of novel system.
Park e Cheyer 2006	Just For Me: Topic Maps and Ontologies	Manages knowledge in three layers: Topics, Knowledge Structures and Documents. Not implemented (at least at the time of publication).
Michlmayr, et al. 2007	Adaptive User Profiles for Enterprise Information Access	Represents context as an individual graph. The graph is constructed based on tags used to semantically annotate the content which the user considers to be relevant. However, the tags are only that, only labels. They are not tied-down to a formal definition.

### Alguns trabalhos relacionados (cont.)

Name	Article	Comments
Aleman-Meza et al. 2003	Context-Aware Semantic Association Ranking	Also defines context as regions. But the regions are static, defined in the ontology which is used by the system. So the contexts are defined the same for every single user.
Mani and Sundaram 2007	Modeling user context with applications to media retrieval	Creates a graph representing the user's context. Each node corresponds to an instance in the ontology. The edges represent relations between the instances. Built to search through multimedia documents. However, it does not associate the vocabulary the user uses with the instances.
Challam et al. 2007	Contextual Search Using Ontology-Based User Profiles	Monitors the user's activity capturing content from open Internet Explorer, MS-Office and MSN. The captured information is stored and used to build a user's contextual profile bases on an ontology. The context is tightly coupled with the ontology.
Vallet et al. 2006	Personalized Information Retrieval in Context	Stores context as a graph constructed based on the ontology used by the system. Similar to Mani and Sundaram, including its limitations (mapping the vocabulary), however it built for text documents.
Sieg et al. 2007	Ontological User Profiles for Personalized Web Search	Similar to Challam et al., but it gets the context information from the ontology terms used to annotate the retrieved relevant context. Also limited by the structure of the ontology.

### Alguns trabalhos relacionados (cont.)

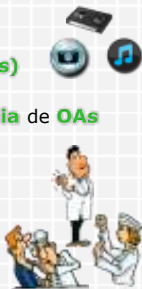
Name	Deals with objective knowledge (ontology)?	Deals with subjective knowledge (user's context)?	The knowledge management is transparent?	Idea / Implemented
Graupmann et al. 2005	X			IMPL.
Park e Cheyer 2006	X	X	X	IDEA
Michlmayr, et al. 2007		X	X	IMPL.
Aleman-Meza et al. 2003	X		X	IMPL.
Mani e Sundaram 2007	X	X	X	IMPL.
Challam et al. 2007	X	X?	X	IMPL.
Vallet et al. 2006	X	X?	X	IMPL.
Sieg et al. 2007	X	X?	X	IMPL. (incomplete)
<b>Praestro</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>IMPL. (prototype)</b>

### UnA-SUS – Universidade Aberta do SUS

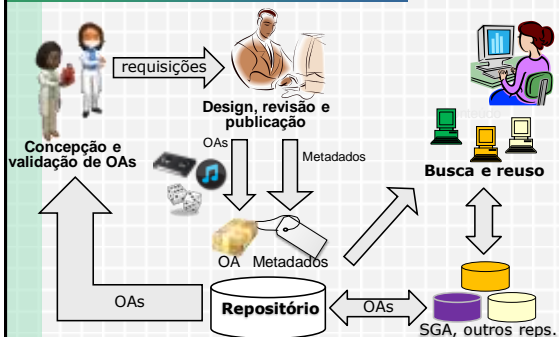
- Programa do Ministério da Saúde para atender necessidades de **formação e educação** permanente dos **trabalhadores do SUS**
- Ações focadas em **e-learning**:
  - formulação de conteúdo
  - bibliotecas digitais
  - cursos a distância

## Metas da UnA-SUS

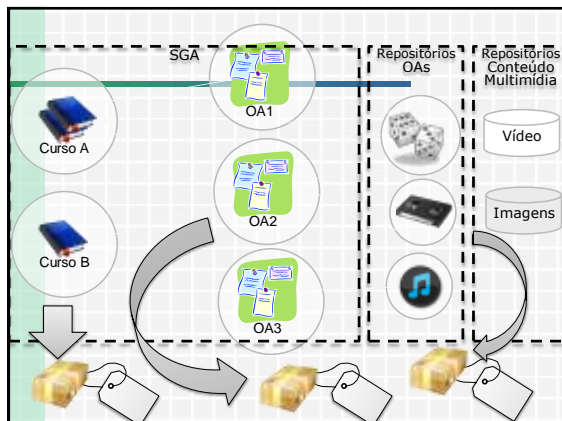
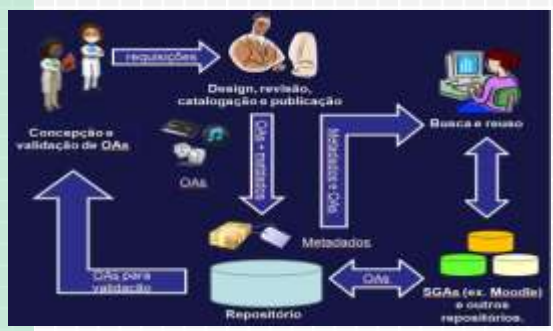
- Desenvolver um **Ambiente Virtual de Ensino e Aprendizagem**
  - plataforma **Web** para **gestão** de **Objetos de Aprendizagem (OAs)**
  - composição **modular** e **multimídia** de **OAs**
- Montar e oferecer cursos para formação continuada de profissionais da saúde
  - Especialização em Saúde da Família (ESF)**



## O processo de produção, catalogação e reuso de OAs

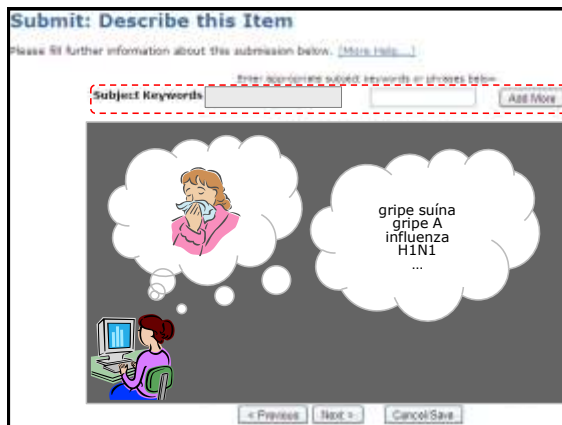


## O processo de produção, catalogação e reuso de OAs

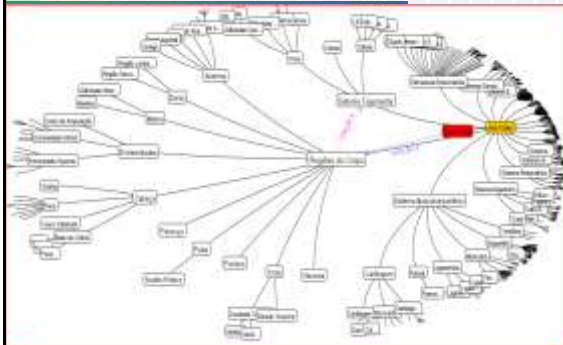


## Reuso é essencial

- OAs são caros para produzir !!!
- Demandas emergenciais na saúde pública exigem agilidade e pronta resposta com cursos, para a qualificação de profissionais



## Navegação hiperbólica no DeCS



## Catlogação de Objetos de Aprendizagem (OAs) usando Vocabulários Controlados

- Seleção de termos do DeCS, CID-10, SNOMED, ...
  - via fornecimento de palavras-chave que são pesquisadas na base de conhecimento
  - via navegação em uma visão do conhecimento em forma de árvore

## Seleção via contexto ontológico

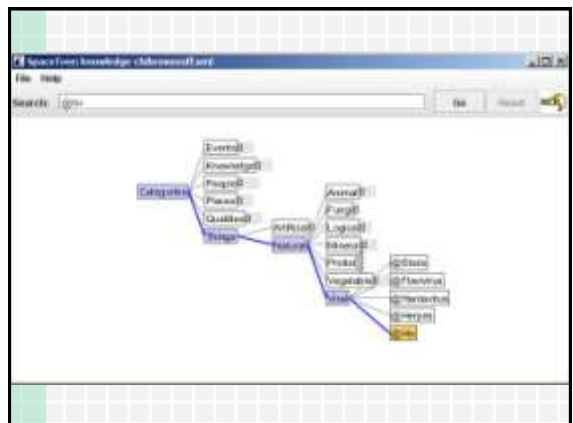


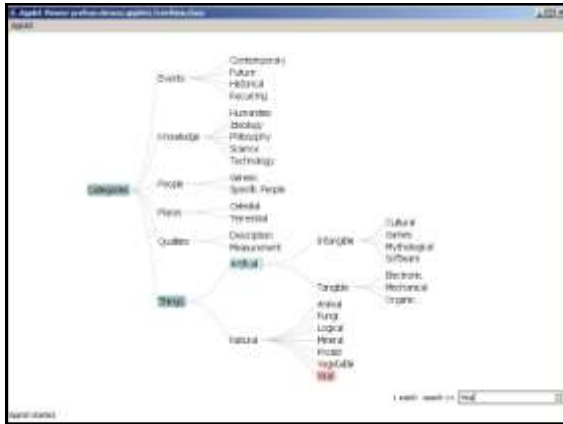
## Protótipo (catlogação com entrada de palavra-chave)

[DSpace:Catlogação de OA.swf](#)

## Protótipo (catlogação com navegação na base de conhecimento)

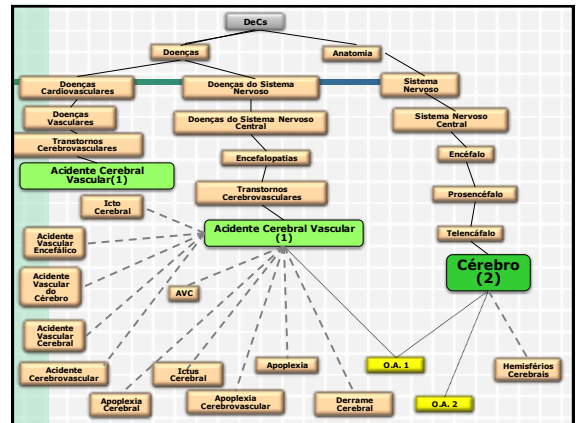
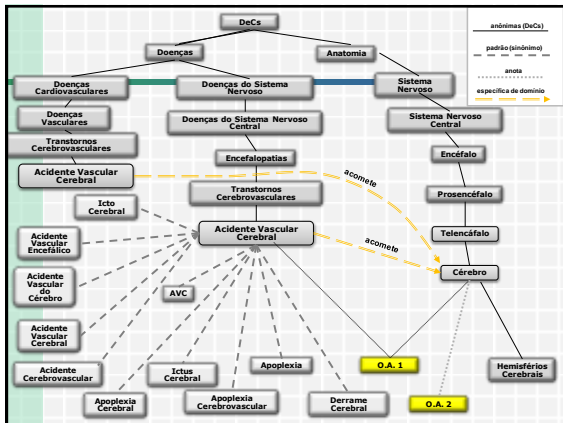
[DSpace:Recuperação de OA.swf](#)





## Recuperação de Objetos de Aprendizagem (OAs) Baseada em Conhecimento

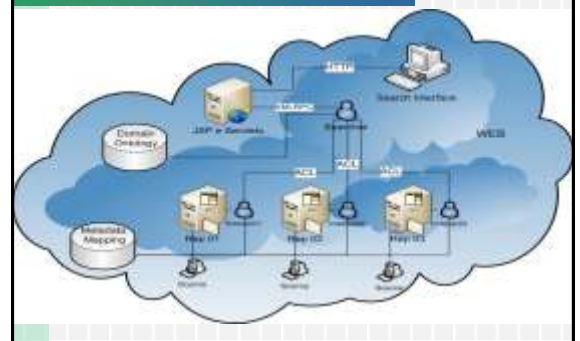
- Vocabulários e relações semânticas:
  - Oriundos do DeCS
  - Definidos pelos catalogadores de OAs
  - Gerados pelo cruzamento de informações com o CID-10 (Classificação Internacional de Doenças)
  - Ex: sinônimos, é um(a), parte de, causa, efeito, sintoma, etc.



## Busca de OAs na área de saúde usando conhecimento de domínio

- ANATOMIA
  - Regiões do Corpo
    - Abdome
      - Cavidade Abdominal
        - Peritônio
          - Saco de Douglas
          - Mesentério + Omento
          - Cavidade Peritoneal
          - Estomas Peritoneais
          - Espaço Retroperitoneal

## Semantic Learning Objects (VIAN, J. ; SILVEIRA, R. A. ; FILETO, R. , WCCE 2009)



## Conclusões

- A Web semântica tem potencial de contribuir na obtenção de melhores resultados para buscas
- Suporte automatizado é fundamental para a construção, evolução e integração de ontologias, anotações semânticas e contextos de usuários
- Nossos primeiros resultados focam na captura e utilização de contextos de usuários mapeados a ontologias de domínios específicos para desambiguar e estender as buscas

## Trabalhos futuros

- Processos parcialmente automatizados para geração e atualização de ontologias e anotações semânticas
- Experimentos para comprovar a eficácia das técnicas e ferramentas utilizadas
- Validação de soluções de problemas de busca semântica em diversos domínios de aplicação
- Determinação de níveis adequados de acoplamento entre contextos e ontologias e dessas com os recursos a serem recuperados

## Notícias e análises

- <http://news.cnet.com/new-search-engine-cuil-takes-aim-at-google/>
- [http://www.readwriteweb.com/archives/is\\_google\\_a\\_semantic\\_search\\_engine.php](http://www.readwriteweb.com/archives/is_google_a_semantic_search_engine.php)
- [http://www.readwriteweb.com/archives/semantic\\_search\\_the\\_myth\\_and\\_reality.php](http://www.readwriteweb.com/archives/semantic_search_the_myth_and_reality.php)
- <http://www.pandia.com/sew/1262-top-5-semantic-search-engines.html>
- <http://mindset.research.yahoo.com/>
- [http://news.cnet.com/8301-13953\\_3-9982015-80.html](http://news.cnet.com/8301-13953_3-9982015-80.html)
- <http://www.searchenginejournal.com/askcom-focuses-on-semantic-search/8252/>
- [http://news.bbc.co.uk/2/hi/programmes/click\\_online/default.stm](http://news.bbc.co.uk/2/hi/programmes/click_online/default.stm)
- [http://news.bbc.co.uk/2/hi/programmes/click\\_online/8144765.stm](http://news.bbc.co.uk/2/hi/programmes/click_online/8144765.stm)

## Web semântica e buscas semânticas

- <http://www.w3.org/2001/sw/>
- <http://semanticweb.org/>
- [http://en.wikipedia.org/wiki/Semantic\\_Web](http://en.wikipedia.org/wiki/Semantic_Web)
- [http://en.wikipedia.org/wiki/Semantic\\_search](http://en.wikipedia.org/wiki/Semantic_search)

## Ferramentas de desenvolvimento

- <http://protege.stanford.edu/>
- <http://jena.sourceforge.net/>
- <http://www.w3.org/2001/Annotea/>
- <http://www.ontotext.com/kim/>
- <http://esw.w3.org/topic/SemanticWebTools>
- <http://www.daml.org/2003/05/swmu-tools-tutorial/Overview.html>

## Referências em Web semântica

- T. Berners-Lee, J. Hendler e O. Lassila. **The Semantic Web**. Scientific American, May 2001.
- Thomas B. Passing **Explore's Guide to the Semantic Web**, Manning Publications, Greenwich, CT, 2005.
- Davies, J., Studer, R., Warren, P. (Eds.) **Semantic Web Technologies: trends and research in ontology-based Systems**, John Wiley & Sons, 2006.
- Breitman, K.K., Casanova, M.A., Truszkowski, W. **Semantic Web: Concepts, Technologies and Applications**. Series: NASA Monographs in Systems and Software Engineering, Springer 2007.
- Grigoris Antoniou and Frank van Harmelen **A Semantic Web Primer**, 2nd edition, The MIT Press, Cambridge, MA, USA, 2008.
- Kashyap, V., Bussler, C., Moran, M. **The Semantic Web - Semantics for Data and Services on the Web**, Series: Data-Centric Systems and Applications. Springer, 2008.
- Hitzler, P., Krotzsch, M., Rudolph, S. **Foundations of Semantic Web Technologies**. Chapman & Hall/CRC 2009.

## Referências em buscas semânticas

- Guha, R., McCool, R., and Miller, E. 2003. **Semantic search**. In *Proc. of the 12th international Conference on World Wide Web (WWW)*, Budapest, Hungary. ACM, New York, NY, 2003, 700-709.
- Reeve, L. and Han, H. 2005. **Survey of semantic annotation platforms**. In *Proceedings of the ACM Symposium on Applied Computing*, Santa Fe, New Mexico. ACM, New York, NY, 2005, 1634-1638.
- Eetu Makela. **Survey of Semantic Search Research**
- Mangold, C. **A Survey and classification of semantic search approaches**. *Journal of Metadata Semantics and Ontology*, 2(1), 2007.
- Stephan Bloehdorn, Philipp Cimiano, Alistair Duke, Peter Haase, Jörg Heizmann, Ian Thurlow, Johanna Völker. **Ontology-Based Question Answering for Digital Libraries**. *ECDL 2007*: 14-25.
- Hai Dong; Hussain, F.K.; Chang, E. **A survey in semantic search technologies**. In *2nd IEEE International Conference on Digital Ecosystems and Technologies*. 2008, 403-408

## Algumas referências grupo UFSC

- D'Agostini, C. S. ; Fileto, R. **Capturing Users' Preferences and Intentions in a Semantic Search System**. In: *21st International Conference on Software Engineering & Knowledge Engineering (SEKE)*, Boston, 2009. p. 587-591.
- D'Agostini, C. S. ; Fileto, R. ; Dantas, M. A. R. ; Gauthier, F. A. O. **Contextual Semantic Search - Capturing, using the User's Context to Direct Semantic Search**. In: *10th International Conference on Enterprise Information Systems (ICEIS)*, Barcelona, Spain, 2008. v. SAIC. p. 154-159.
- D'Agostini, C. S. ; Fileto, R. **Capturing and managing user context for improving information retrieval**. In: *Workshop de Teses e Dissertações do Simpósio Brasileiro de Bancos de Dados (WTDBD/SBBD)*, Campinas: Unicamp, Brazil, 2008.
- Vian, J. ; Silveira, R. A. ; Fileto, R. **Proposal of a Multi-agent System for Indexing and Recovery applied to Learning Objects**. In: *9th IFIP World Conference on Computers in Education (WCCE)*, Bento Goncalves, Brazil, 2009.

## Alguns projetos na UFSC

- <http://www.lisa.ufsc.br/projetos>
- <http://www.unasus.ufsc.br>
- <http://www.literaturabrasileira.ufsc.br>

## Perguntas?

