prolog 2007
:
Introduction to Knowledge Engineering

Monika Lanzenberger

That's me ...

Assistant Professor (Universitätsassistentin)

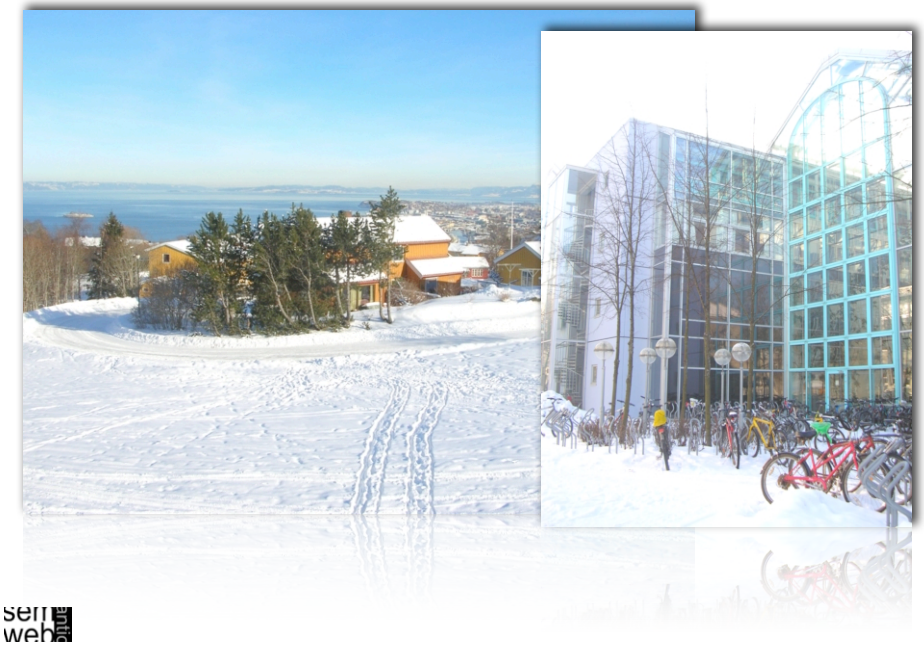
Institute of Software Technology and Interactive Systems,
Information & Software Engineering Group (188/1)
Favoritenstr. 9-11, A-1040 Wien

2001: MSc (Diplomstudium Informatik),
Vienna University of Technology

2003: PhD (Doktoratstudium der Technischen Wissenschaften),
Vienna University of Technology

Post-Doc Fellowships:
February - October 2005, Norwegian University of Science and Technology
November - July 2006, CRP Henri Tudor, Centre for IT Innovation, Luxembourg

That's Trondheim ...



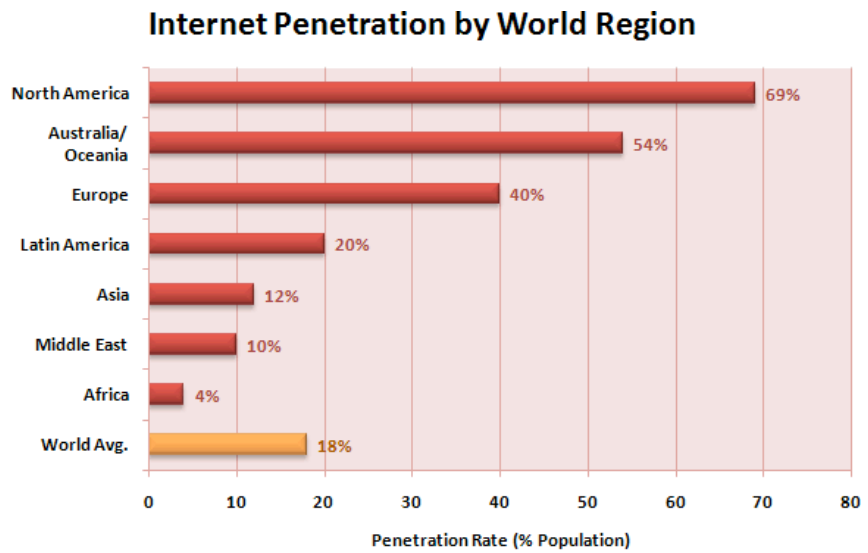
Outline

- World Wide Web
- Web 2.0
- Semantic Web
- Data, Information, and Knowledge
- XML and OWL
- Ontologies in Practice
- Conclusions

- Integration of different data types and services
- Simple and freely available standards and protocols
- Client-server architecture or peer to peer
- Compatible upgrades by W3C (World Wide Web Consortium)
- ICANN (Internet Corporation for Assigned Names and Numbers) coordinates the unique identifiers (domain names, IP addresses, etc.)
- Platform independent
- Decentralized resources (hardware and software)

- Encryption and secure connections
- Free tools for authoring, searching, communicating, etc.
- Integration of text, links, pictures, audio and video files, etc.
- Hypertext using HTML - Layout (Hypertext Markup Language)
- Simple elements called tags
  `<A HREF="... "> ... </A>`

- Spam
- Security and privacy problems
- Energy intensive
- Currently over 80 % of the world population cannot access the WWW
- Author or user?
- Web content for humans



Outline

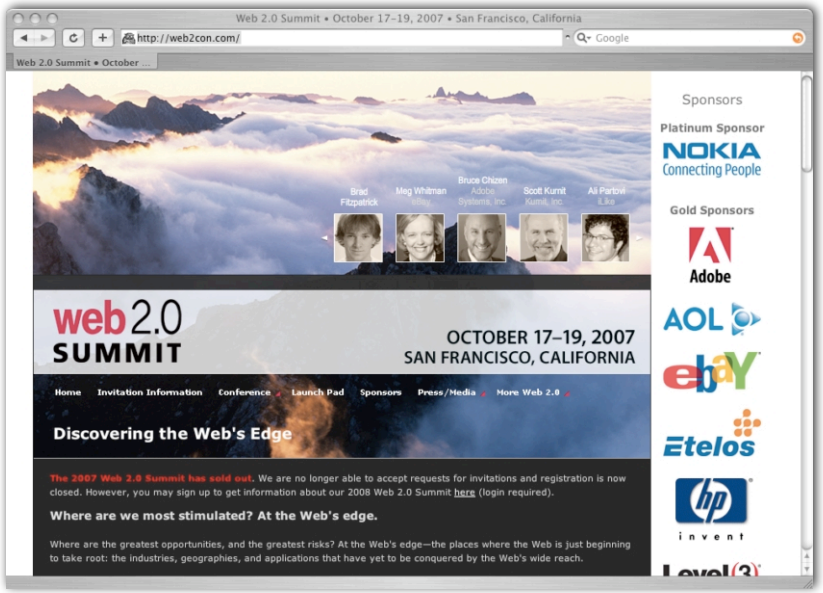
- World Wide Web
- Web 2.0
- Semantic Web
- Data, Information, and Knowledge
- XML and OWL
- Ontologies in Practice
- Conclusions

Web 2.0

- First Web 2.0 conference was held October 5-7, 2004 in San Francisco
- Services like weblogs, social bookmarking, wikis, etc.

[Angermeier/Cremonini]

Web 2.0 Conference 2007

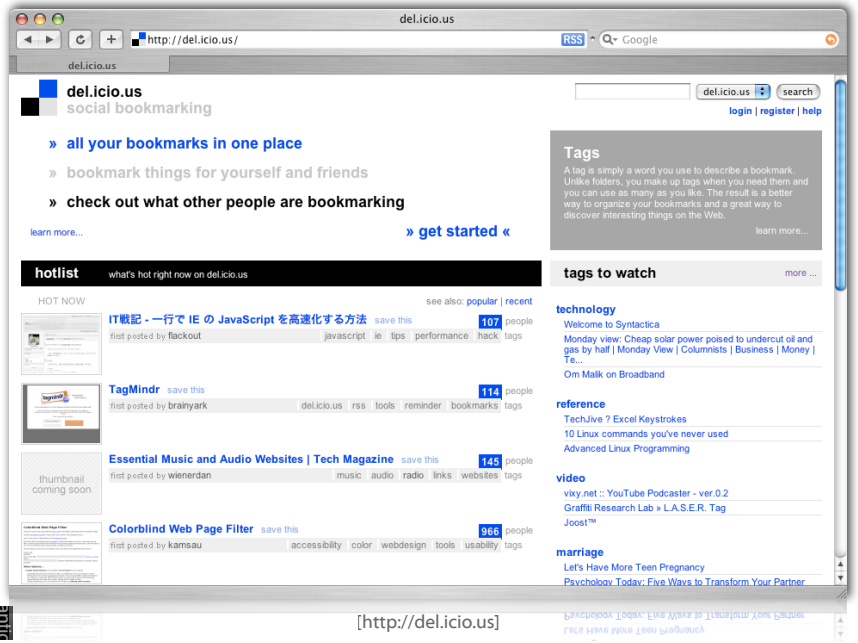


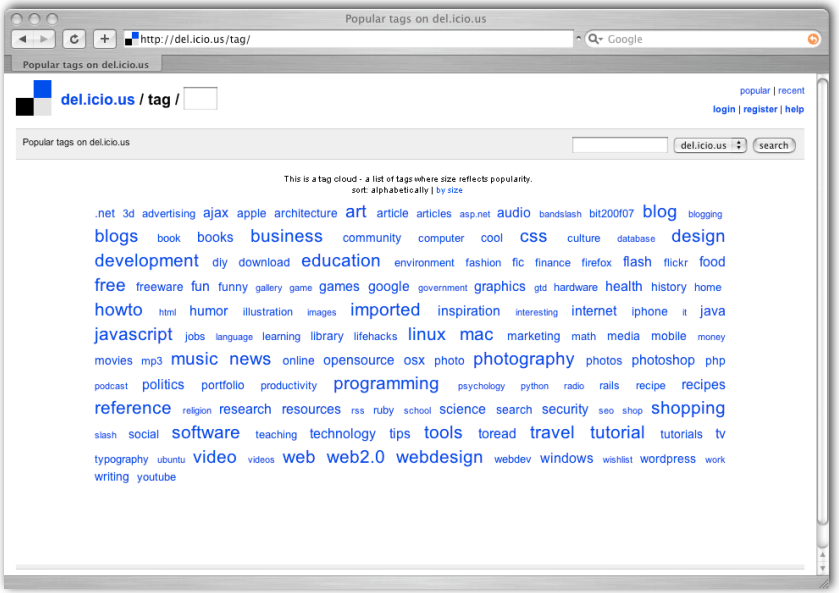
[http://web2con.com]

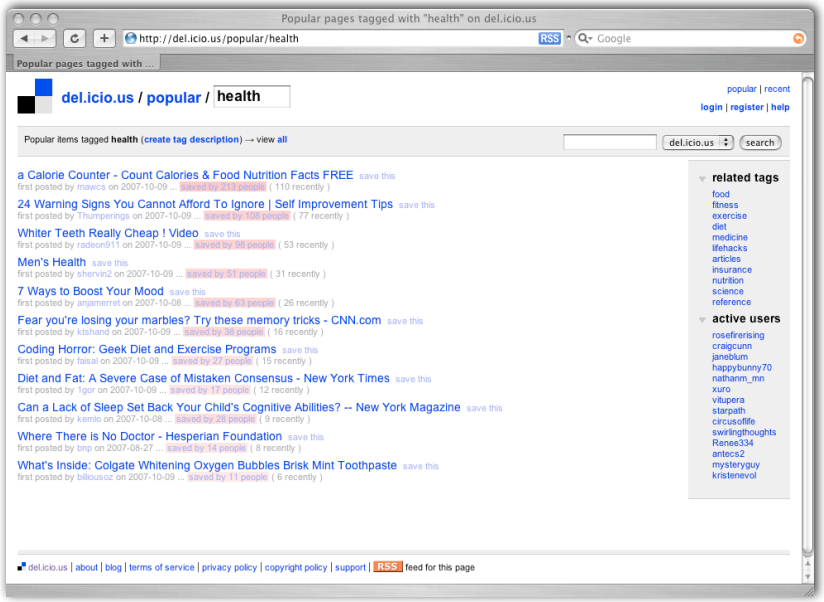
Social Bookmarking

- Keep links to your favorite articles, blogs, music, reviews, recipes, and more, and access them from any computer on the web.
- Share favorites with friends, family, coworkers, and the del.icio.us community.
- Discover new things. Everything on del.icio.us is someone's favorite -- they've already done the work of finding it. So del.icio.us is full of bookmarks about technology, entertainment, useful information, and more.

[http://del.icio.us]







- Tags
  - One-word descriptors;
  - Assign them to your bookmarks;
  - Similar to keywords but non-hierarchical;
  - Assign as many tags to a bookmark and rename or delete them later;
  - Tagging is easier and more flexible than using categories or folders.
- Feedback from a user
  - "Feedback is immediate. As soon as you assign a tag to an item, you see the cluster of items carrying the same tag. If that's not what you expected, you're given incentive to change the tag or add another ... you can adapt to the group norm, keep your tag in a bid to influence the group norm, or both."
- Folksonomy
  - The emerging category structure that results from the combination of a large number of users tagging resources for their own use.

Social Bookmarking

From Folksonomy to Taxonomy ...

Websites like Yahoo! organize and present links in a fixed hierarchy (taxonomy)

Taxonomy directed folksonomies: Text analysis of the website using a lexical database like WordNet

- English nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept with 206941 word-sense pairs.
- Synsets are interlinked by means of conceptual-semantic and lexical relations resulting in a network of meaningfully related words and concepts.
- WordNet is also freely and publicly available (http://wordnet.princeton.edu/)
- Tags could be suggested by using the lexical database

Using folksonomies to create simple ontologies

- Examine which tags are used in combination to estimate the relations among these tags
- Map the tags to a lexical database
- Learn additional relations based on the tag clouds and the lexical database

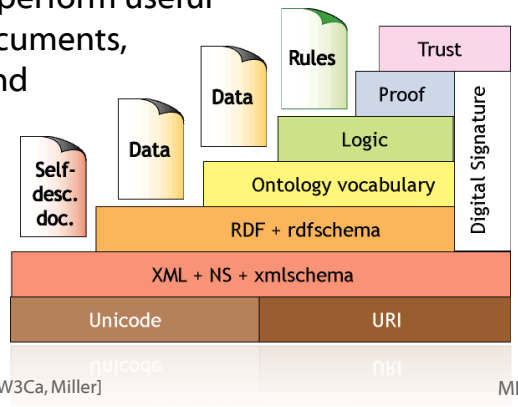
Outline

- World Wide Web
- Web 2.0
- Semantic Web
- Data, Information, and Knowledge
- XML and OWL
- Ontologies in Practice
- Conclusions

Semantic Web

The Semantic Web is a vision for the future of the Web [...] information is given explicit meaning, [...] machines automatically process and integrate information available on the Web.

If machines are expected to perform useful reasoning tasks on these documents, the language must go beyond the basic semantics of RDF Schema.



[W3Ca, Miller]

Requirements for Ontology Languages

Ontology languages allow users to write explicit, formal conceptualizations of domain models.

The main requirements are:

- a well-defined syntax
- efficient reasoning support
- a formal semantics
- sufficient expressive power
- convenience of expression

[Antoniou and van Harmelen, 2004]

- Class membership

  If x is an instance of a class C,

  and C is a subclass of D,

  then we can infer that x is an instance of D.

- Equivalence of classes

  If class A is equivalent to class B,

  and class B is equivalent to class C,

  then A is equivalent to C, too.

- Consistency

  Consider x being an instance of classes A and B,

  but A and B are disjoint.
  --> Indication of an error in the ontology.

- Classification

  Certain property-value pairs are a sufficient condition for membership in a class A; if an individual x satisfies such conditions, we can conclude that x must be an instance of A.

Reasoning support is important for...
... checking the consistency of the ontology and the knowledge.
... checking for unintended relationships between classes.
... automatically classifying instances in classes.

Checks like the preceding ones are valuable for...
... designing large ontologies, where multiple authors are involved.
... integrating and sharing ontologies from various sources.

The Friend of a Friend (FOAF) project:

- is creating a Web of machine-readable pages describing people, the links between them and the things they create and do
- applies simple technology that makes it easier to share and use information about people and their activities (eg. photos, calendars, weblogs)
- FOAF uses W3C's RDF technology to integrate information from your home page with that of your friends, and the friends of your friends, and their friends...
- FOAF-a-matic is a simple Javascript application that allows you to create a FOAF ("Friend-of-A-Friend") description of yourself
- Make file publically accessible (foaf.rdf) and link it on your website

Semant Web Application FOAF Code


```
<rdf:RDF
      ...
<foaf:Person rdf:ID="me">
   <foaf:name>Monika Lanzenberger</foaf:name>
   <foaf:title>Dr.</foaf:title>
   <foaf:givenname>Monika</foaf:givenname>
   <foaf:family_name>Lanzenberger</foaf:family_name>
   <foaf:mbox rdf:resource="mailto: ... "/>
   <foaf:phone rdf:resource="tel:+43-1-58801-18871"/>
   <foaf:workplaceHomepage rdf:resource="http://www.ifs.tuwien.ac.at/
   ~mlanzenberger"/>
   <foaf:knows>
      <foaf:Person>
         <foaf:name>Elke Michlmayr</foaf:name>
         <foaf:mbox rdf:resource="mailto: ..."/>
      </foaf:Person>
   </foaf:knows>
</foaf:Person>
</rdf:RDF>
```

[http://www.foaf-project.org]

Semant Web Application FOAF Code


```
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:admin="http://webns.net/mvcb/">


<foaf:PersonalProfileDocument rdf:about="">
  <foaf:maker rdf:resource="#me"/>
  <foaf:primaryTopic rdf:resource="#me"/>
  <admin:generatorAgent rdf:resource="http://www.ldodds.com/
foaf/foaf-a-matic"/>
  <admin:errorReportsTo
rdf:resource="mailto:leigh@ldodds.com"/>
</foaf:PersonalProfileDocument>
```

[http://www.foaf-project.org]

Outline

- World Wide Web
- Web 2.0
- Semantic Web
- Data, Information, and Knowledge
- XML and OWL
- Ontologies in Practice
- Conclusions

Definitions

- Data
  "input signals to sensory and cognitive processes"
- Information
  "data with an associated meaning"
- Knowledge
  "the whole body of data and information together with cognitive machinery that people are able to exploit to decide how to act, to carry out tasks and to create new information"

[Schreiber et al., 2000]

Examples

- Some pieces of data

```
a:= 15
b:= "Huber"
```

- A piece of information

```
<Region rdf:ID="CentralCoastRegion" />
```

- A piece of knowledge

```
<owl:Class rdf:ID="ConsumableThing" />

<owl:Class rdf:ID="NonConsumableThing">
  <owl:complementOf rdf:resource="#ConsumableThing" />
</owl:Class>
```


Outline

- World Wide Web
- Web 2.0
- Semantic Web
- Data, Information, and Knowledge
- XML and OWL
- Ontologies in Practice
- Conclusions

XML

- eXtensible Markup Language
- Foundation for the creation of purpose specific, XML-based markup languages
- Individually defined tags such as `<company_name>`
- Using Unicode (multi-lingual character-encoding system)
- Free, open standard recommended by W3C
- Two levels of correctness: Well-formed and Valid
- Namespace provide uniquely named elements and attributes (`xmlns="http://www.w3.org/1999/xhtml"`)

XML


```
<?xml version = "1.0"?>
    <buecher schlagwort = "semantic web" />
     <buch>
         <titel>Visualizing the Semantic Web</titel>
         <isbn>1-85233-576-9</isbn>
         <ed>
             <vorname>Vladimir</vorname>
             <nachname>Geroimenko</nachname>
         </ed>
         <seiten>202</seiten>
         <photo filename = "vis_semweb.jpg" />
     </buch>
 </buecher>
```

From XML to OWL

33

- XML provides a surface syntax for structured documents, but imposes no semantic constraints on the meaning of these documents.
- XML Schema is a language for restricting the structure of XML documents and also extends XML with data types.
- RDF is a data model for objects ("resources") and relations between them, provides a simple semantics for this data model, and these data models can be represented in an XML syntax.
- RDF Schema is a vocabulary for describing properties and classes of RDF resources, with a semantics for generalization-hierarchies of such properties and classes.
- OWL adds more vocabulary for describing properties and classes: among others, relations between classes (e.g. disjointness), cardinality (e.g. "exactly one"), equality, richer typing of properties, characteristics of properties (e.g. symmetry), and enumerated classes.

OWL Lite Constructs: Simple Classes and Individuals

34

Simple Named Classes:

- `Class`
- `rdfs:subClassOf`

```
<owl:Class rdf:ID="Winery"/>
<owl:Class rdf:ID="Region"/>
<owl:Class rdf:ID="ConsumableThing"/>
```

`Individual`

Defining Properties:

- `rdf:Property`
  - subproperties:
    - `owl:ObjectProperty` (Instance - Instance)
    - `owl:DatatypeProperty` (Instance - `rdfs:Literal` / XML Schema datatypes)
- `rdfs:subPropertyOf`
- `rdfs:domain`
- `rdfs:range`

Properties of Individuals

OWL Lite Constructs: Simple Classes and Individuals

35

Simple Named Classes:

- `Class`
- `rdfs:subClassOf`

```
<owl:Class rdf:ID="Winery"/>
<owl:Class rdf:ID="Region"/>
<owl:Class rdf:ID="ConsumableThing"/>
```

```
<owl:Class rdf:ID="PotableLiquid">
  <rdfs:subClassOf rdf:resource="#ConsumableThing" />
  ...
</owl:Class>
```

`Individual`

Defining Properties:

- `rdf:Property`
  - subproperties:
    - `owl:ObjectProperty` (Instance - Instance)
    - `owl:DatatypeProperty` (Instance - `rdfs:Literal` / XML Schema datatypes)
- `rdfs:subPropertyOf`
- `rdfs:domain`
- `rdfs:range`

Properties of Individuals

OWL Lite Constructs: Simple Classes and Individuals

36

Simple Named Classes:

- `Class`
- `rdfs:subClassOf`

`Individual`

Defining Properties:

- `rdf:Property`
  - subproperties:
    - `owl:ObjectProperty` (Instance - Instance)
    - `owl:DatatypeProperty` (Instance - `rdfs:Literal` / XML Schema datatypes)
- `rdfs:subPropertyOf`
- `rdfs:domain`
- `rdfs:range`

```
<owl:ObjectProperty rdf:ID="madeFromGrape">
  <rdfs:domain rdf:resource="#Wine"/>
  <rdfs:range rdf:resource="#WineGrape"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:ID="course">
  <rdfs:domain rdf:resource="#Meal" />
  <rdfs:range rdf:resource="#MealCourse" />
</owl:ObjectProperty>
```

Properties of Individuals

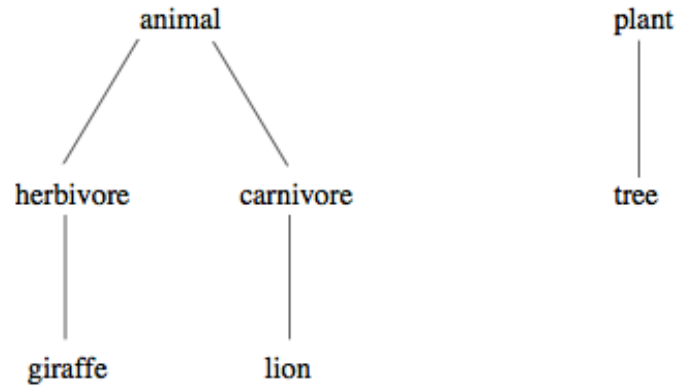
Outline

37

- World Wide Web
- Web 2.0
- Semantic Web
- Data, Information, and Knowledge
- XML and OWL
- Ontologies in Practice
- Conclusions

An African Wildlife Ontology – Class Hierarchy

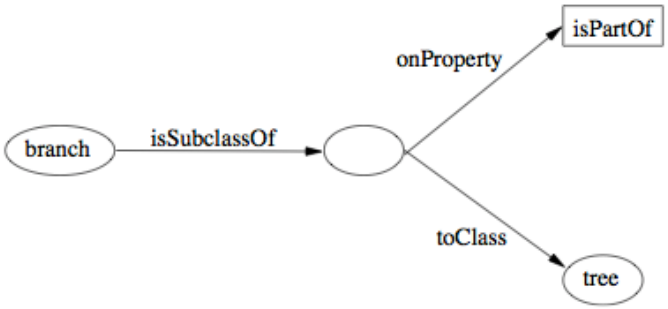
38



[Antoniou and van Harmelen, 2004]

An African Wildlife Ontology – Schematic Representation

39



[Antoniou and van Harmelen, 2004]

An African Wildlife Ontology – Properties

40

```
<owl:TransitiveProperty rdf:ID="is-part-of"/>

<owl:ObjectProperty rdf:ID="eats">
    <rdfs:domain rdf:resource="#animal"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:ID="eaten-by">
    <owl:inverseOf rdf:resource="#eats"/>
</owl:ObjectProperty>
```

[Antoniou and van Harmelen, 2004]

An African Wildlife Ontology – Plants and Trees


```
<owl:Class rdf:ID="plant">
  <rdfs:comment>Plants are disjoint from animals.
  </rdfs:comment>
  <owl:disjointWith="#animal"/>
</owl:Class>

<owl:Class rdf:ID="tree">
  <rdfs:comment>Trees are a type of plant.</rdfs:comment>
  <rdfs:subClassOf rdf:resource="#plant"/>
</owl:Class>
```


An African Wildlife Ontology – Branches


```
<owl:Class rdf:ID="branch">
    <rdfs:comment>Branches are parts of trees.</rdfs:comment>
    <rdfs:subClassOf>
        <owl:Restriction>
          <owl:onProperty rdf:resource="#is-part-of"/>
          <owl:allValuesFrom rdf:resource="#tree"/>
       </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>
```


An African Wildlife Ontology – Leaves


```
<owl:Class rdf:ID="leaf">
    <rdfs:comment>Leaves are parts of branches. </rdfs:comment>
    <rdfs:subClassOf>
       <owl:Restriction>
         <owl:onProperty rdf:resource="#is-part-of"/>
         <owl:allValuesFrom rdf:resource="#branch"/>
       </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>
```


An African Wildlife Ontology – Carnivores


```
<owl:Class rdf:ID="carnivore">
    <rdfs:comment>Carnivores are exactly those animals
    that eat also animals.</rdfs:comment>
    <owl:intersectionOf rdf:parsetype="Collection">
        <owl:Class rdf:about="#animal"/>
        <owl:Restriction>
          <owl:onProperty rdf:resource="#eats"/>
          <owl:someValuesFrom rdf:resource="#animal"/>
       </owl:Restriction>
    </owl:intersectionOf>
</owl:Class>
```

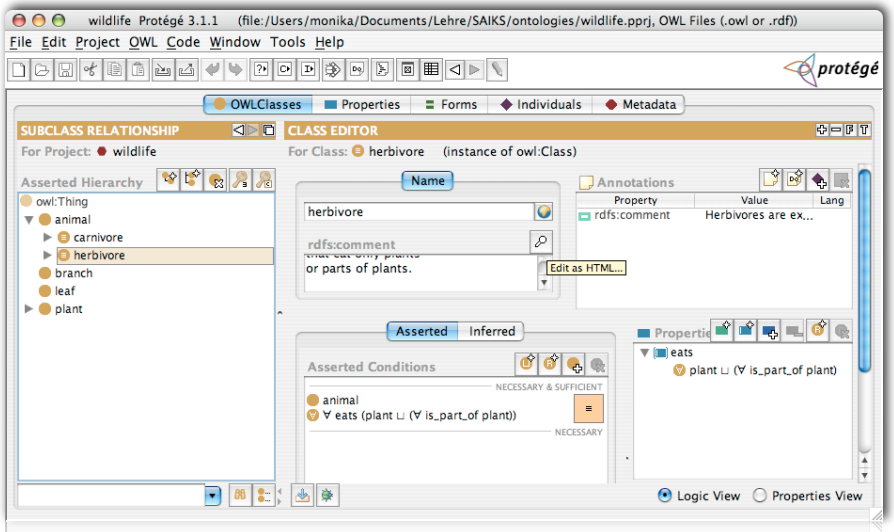
An African Wildlife Ontology – Herbivores


```
<owl:Class rdf:ID="herbivore">

  <rdfs:comment>Herbivores are exactly those animals that
  eat only plants or parts of plants.</rdfs:comment>

  ...                                                    ?

</owl:Class>
```


An African Wildlife Ontology – Herbivores


```
<owl:intersectionOf rdf:parseType="Collection">                    !
  <owl:Class rdf:about="#animal"/>
  <owl:Restriction>
    <owl:onProperty rdf:resource="#eats"/>
    <owl:allValuesFrom>
      <owl:Class>
        <owl:unionOf rdf:parseType="Collection">
          <owl:Class rdf:about="#plant"/>
          <owl:Restriction>
            <owl:onProperty rdf:resource="#is_part_of"/>
            <owl:allValuesFrom rdf:resource="#plant"/>
          </owl:Restriction>
        </owl:unionOf>
      </owl:Class>
    </owl:allValuesFrom>
  </owl:Restriction>
</owl:intersectionOf>
```


An African Wildlife Ontology – Herbivores



An African Wildlife Ontology – Giraffes


```
<owl:Class rdf:ID="giraffe">
    <rdfs:comment>Giraffes are herbivores, and they
    eat only leaves.</rdfs:comment>
    <rdfs:subClassOf rdf:type="#herbivore"/>
    <rdfs:subClassOf>
        <owl:Restriction>
          <owl:onProperty rdf:resource="#eats"/>
          <owl:allValuesFrom rdf:resource="#leaf"/>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>
```

An African Wildlife Ontology – Lions


```
<owl:Class rdf:ID="lion">
    <rdfs:comment>Lions are animals that eat
    herbivores.</rdfs:comment>
    <rdfs:subClassOf rdf:type="#animal"/>
    <rdfs:subClassOf>
        <owl:Restriction>
          <owl:onProperty rdf:resource="#eats"/>
          <owl:someValuesFrom rdf:resource="#herbivore"/>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>
```


An African Wildlife Ontology – Tasty Plants

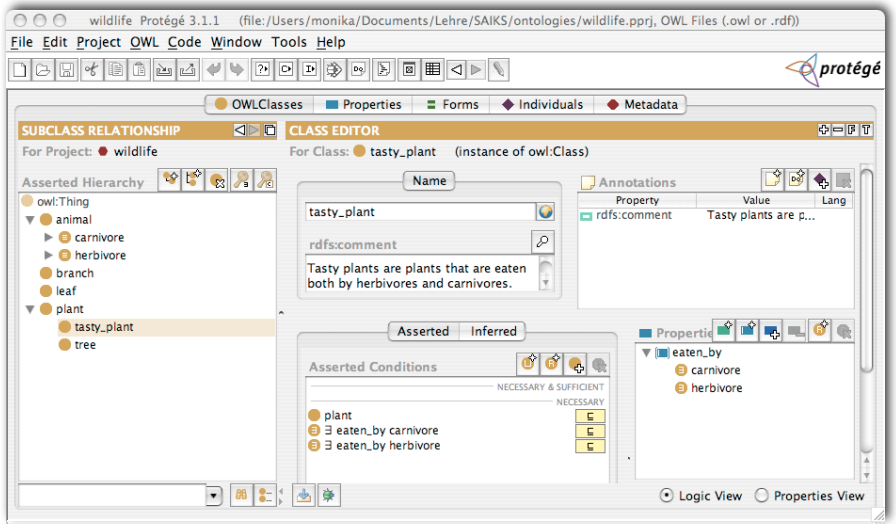

```
<owl:Class rdf:ID="tasty-plant">
  <rdfs:comment>Plants eaten both by herbivores and
  carnivores</rdfs:comment>

    ...                                          ?

</owl:Class>
```


An African Wildlife Ontology – Tasty Plants


```
<rdfs:subClassOf rdf:resource="#plant"/>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty rdf:resource="#eaten_by"/>
    <owl:someValuesFrom>
      <owl:Class rdf:about="#herbivore"/>
    </owl:someValuesFrom>
  </owl:Restriction>
</rdfs:subClassOf>
<rdfs:subClassOf>
  <owl:Restriction>
    <owl:onProperty rdf:resource="#eaten_by"/>
    <owl:someValuesFrom>
      <owl:Class rdf:about="#carnivore"/>
    </owl:someValuesFrom>
  </owl:Restriction>
</rdfs:subClassOf>
```

!

An African Wildlife Ontology – Tasty Plants



An African Wildlife Ontology

What problem would emerge if we replace
`owl:someValuesFrom` by
`owl:allValuesFrom`
in the definition of carnivores?

[Antoniou and van Harmelen, 2004]

Outline

- World Wide Web
- Web 2.0
- Semantic Web
- Ontologies in Practice
- XML and OWL
- Data, Information, and Knowledge
- Conclusions

Please enter your conclusions ...

- Verwenden Sie Web 2.0 Tools?
- Hatten Sie schon mal mit Semantic Web zu tun?
- Könnte Ihnen eine Ontologie, die Wissen über das Studium und die Technische Universität zur Verfügung stellt, nützlich sein?
- Wenn Sie ein Semantic Web Tool entwickeln würden, welches fänden Sie nützlich und interessant?

References & Resources

[Angermeier/Cremonini] M. Angermeier, Web 2.0 Mindmap, 2005; extension by L. Cremonini, http://en.wikipedia.org/wiki/Image:Web_2.0_Map.svg (checked online 10. Oct. 2007).

[Antoniou and van Harmelen, 2004] G. Antoniou and F. van Harmelen, A Semantic Web Primer, MIT Press, Massachusetts, 2004.

[InternetWorldStats] Miniwatts Marketing Group, Internet World Stats, Usage and Population Statistics, http://www.internetworldstats.com/stats.htm (checked online 10. Oct. 2007).

[Miller] Eric Miller, W3C Layer Cake, http://www.w3.org/2001/09/06-ecdl/slide17-0.html (checked online 29. Sep. 2006).

[Protégé 3.1.1] Stanford Medical Informatics, Protégé-OWL, http://protege.stanford.edu/download/download.html (checked online 8. Nov. 2006).

[Schreiber et al., 2000] G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. van de Felde, and B. Wielinga. Knowledge Engineering and Management: The CommonKADS Methodology, MIT Press, 2000.

[W3Ca] OWL Web Ontology Language Overview, http://www.w3.org/TR/2004/REC-owl-features-20040210/, W3C Recommendation 10 February 2004, (checked online 11. Oct. 2007).

[W3Cb] OWL Web Ontology Language Guide, http://www.w3.org/TR/2004/REC-owl-guide-20040210/, W3C Recommendation 10 February 2004, (checked online 11. Oct. 2007).