

# LOGICA DEI LINGUAGGI NATURALI E ARTIFICIALI

Towards the Semantic Web

Antonio Lieto - University of Salerno

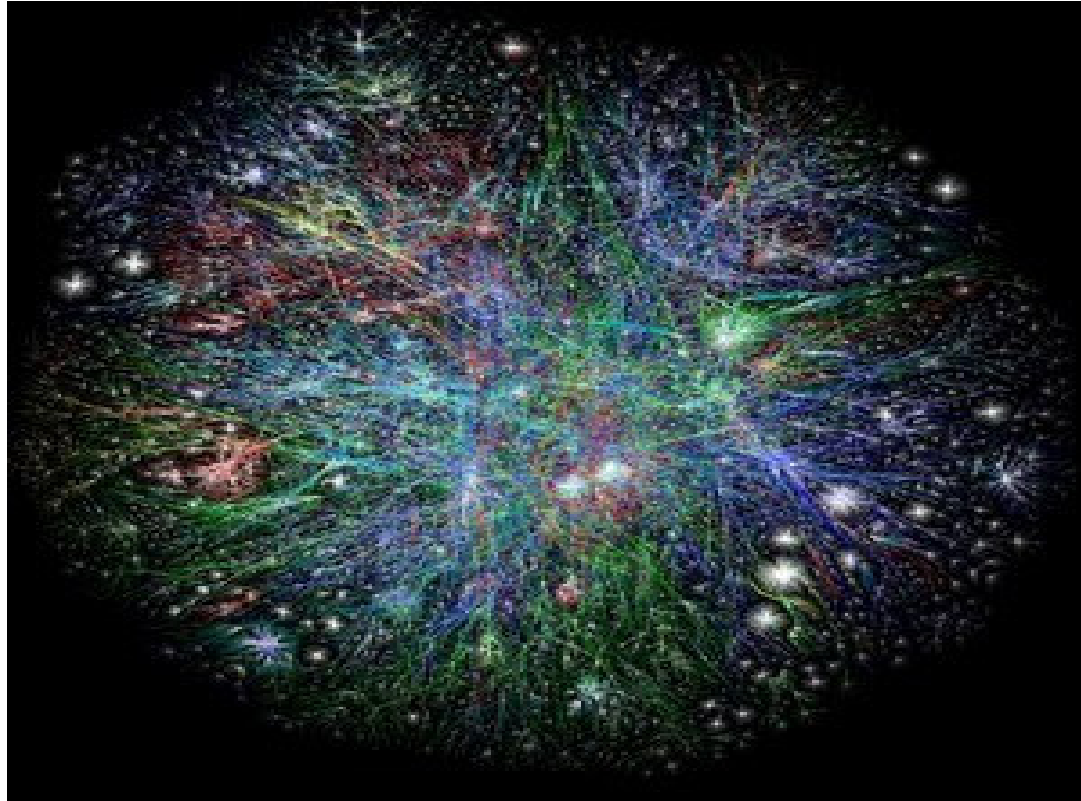
[alieto@unisa.it](mailto:alieto@unisa.it)

<http://antoniolieto.blogspot.com>

# Outline

- Web
- Information Retrieval and Search Engines
- Semantic Web (Definition and Approaches)
- Web Languages Evolution: From the HTML to OWL
- Knowledge Representation Systems and Semantic Web
- Ontologies (Definition and examples)
  
- Ontologies and Description Logics: OWL DL

# Web



# The actual (syntactic) Web

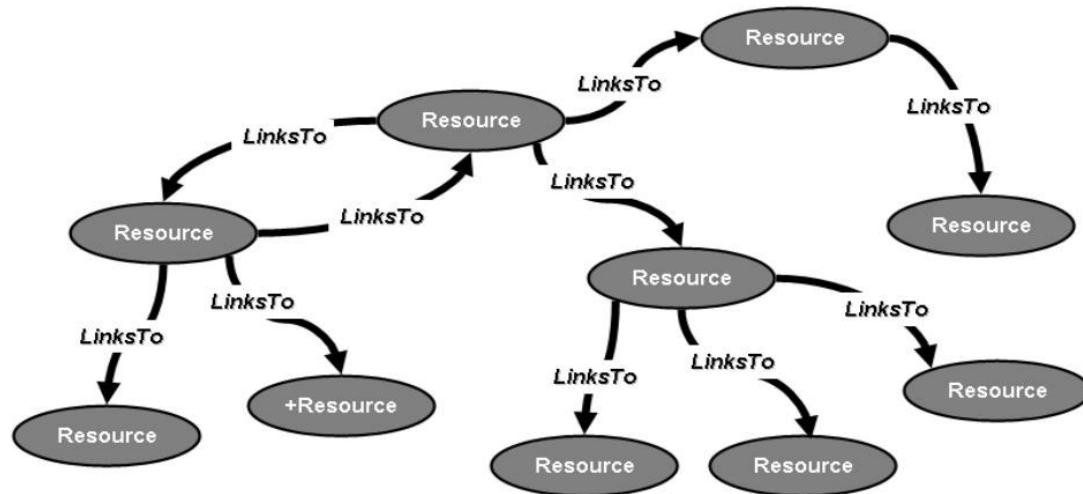


Image adapted from Hendler & Miller (2002)

# The Problem

One of the main “problem” in the World Wide Web is the retrieval of the **relevant** information within a maze of available information and resources.

# Information overload - search process



# Evaluation of an IR System

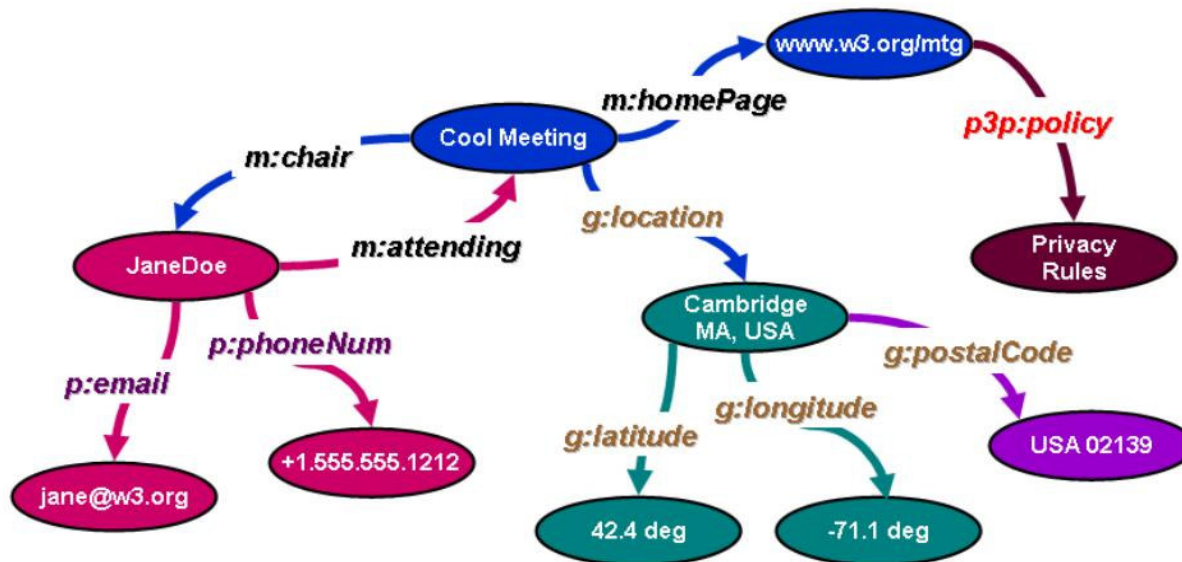
Quantitative measures:

- **PRECISION**= **Relevant Retrieved/ retrieved** ( $R,R / (R,R+NR,R)$ )
- **RECALL**= **Relevant Retrieved/ relevant** ( $R,R / (R,R+R,NR)$ )

	<b>RELEVANT</b>	<b>NON-RELEVANT</b>
<b>RETRIEVED</b>	R,R	NR,R
<b>NOT RETRIEVED</b>	R,NR	NR,NR

# Semantic Web

According to Tim Berners Lee (1999) the Semantic Web is an extension of the current web in which information is given in a **well defined meaning** (e.g. an example of a more structured representation).



# Semantic Web (2)

There are two main approaches to the Semantic Web:

**Formal Approach:** Top - Down (based on Ontology): (have some limits).

**Informal Approach:** Folksonomy (folks + taxonomies): based on *tagging* and bottom -up (have limits too).

A **tag** is a “string” (usually a word) with which an user label a source.



# Web Languages Evolution: HTML

HTML is a mark-up language based on the *presentation* (e.g. “text formatting”) rather than *content* (Horrocks et al. 2003).

Even the retrieval is “formatting” based and not “content” based.

HTML is based on easy to understand (for humans) **tags**.

# Web Languages Evolution (2)

## Examples of HTML well known tags:

**<IMG SRC="image.jpg" WIDTH="100" HEIGHT="50">; search and visualize images hosted elsewhere.**

**<A HREF="url">text name link</a> ;; hypertextual link**

**<A HREF="mailto:alieto@unisa.it">alieto@unisa.it</a> ;; link to my e-mail**

**<TABLE>....</TABLE>; create table**

.....


# Example: Unisa Code

Università degli Studi di Salerno - Mozilla Firefox

File Modifica Visualizza Cronologia Segnalibri Strumenti Aiuto

http://www3.unisa.it/

Antonio Lieto Università degli Studi di Salerno

 UNIVERSITÀ DEGLI STUDI DI SALERNO  
CAMPUS VIVENDI


CERCA

LINK RAPIDI Seleziona il link

Contatti Rubrica Webmail

Ateneo Didattica Ricerca Vivere il Campus

Studenti  
Personale  
Enti e Imprese  
Territorio




**inevidenza**

- Dopo la guerra fredda  
Democrazia, autoritarismo e ... >>>
- Il CUT presenta **Giudizio Universale**  
Una nuova rappresentazione del ... >>>
- Il Dante di Bruno Lucrezi  
>>>

**Caratterizzazione degli odori**  
Laboratorio di Ingegneria Sanitaria Ambientale (SEED), Dipartimento di Ingegneria Civile >>>

**Master DAOSan**  
Formazione: leva strategica per l'innovazione ed il cambiamento in Sanità >>>

 servizio dell'Università degli Studi di Salerno

[leggi tutte le news](#)

UniSa - Università degli Studi di Salerno  
P.IVA 00851300657

Completato

# Example: Unisa Code (header)

```
<head> <meta http-equiv="Content-Type" content="text/html; charset=UTF-8" />
```

```
....
```

```
<meta name="description" content="Portale web dell'Università degli Studi di Salerno" />
```

```
  <meta name="keywords" content="università, salerno, studi, unisa, didattica" />
```

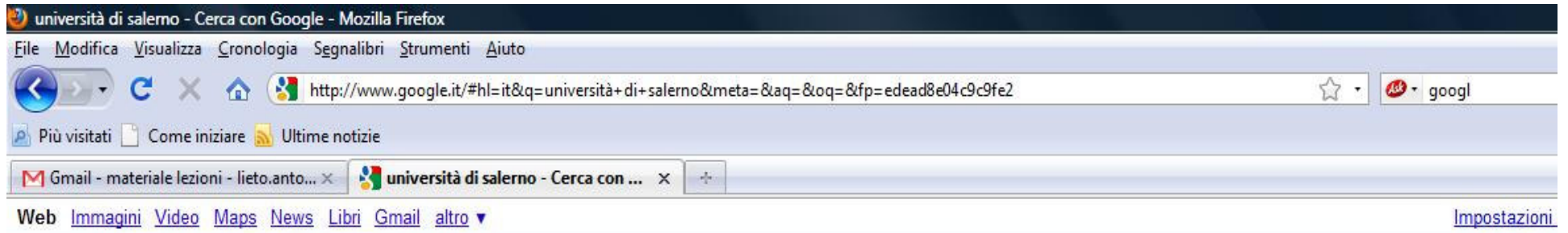
```
  <meta name="author" content="C.S.I. - Ufficio Applicazioni" />
```

```
<title>Università degli Studi di Salerno</title>
```

```
....
```

```
</head>
```

# Search and code



università di salerno

Cerca

[Ricerca avanzata](#)

Cerca:  nel Web  pagine in Italiano  pagine provenienti da: Italia

Web [Mostra opzioni...](#)

Risultati 1 - 10 su circa 1.090.000 per università di salerno. (0,15 secondi)

## [Università degli Studi di Salerno](#)

Portale web dell'Università degli Studi di Salerno.

[www.unisa.it/](http://www.unisa.it/) - [Copia cache](#) - [Simili](#)

[Facoltà](#)

[Webmail](#)

[Didattica](#)

[Ateneo](#)

[Studenti](#)

[Concorsi, bandi e gare](#)

[Segreterie studenti](#)

[Personale](#)

[Altri risultati in unisa.it »](#)

## [Università degli Studi di Salerno: facoltà di ateneo](#)

logo università degli studi di salerno. Università degli Studi di Salerno ... Ufficio

Amministrativo di Facoltà: tel. 089/96 9787 fax: 089/96 9608 ...

[www.unisa.it/include/view\\_struct.php?do=view...](http://www.unisa.it/include/view_struct.php?do=view...) - [Copia cache](#) - [Simili](#)

## [Università degli Studi di Salerno: Facoltà di Ingegneria](#)

Portale web dell'Università degli Studi di Salerno.

[www.ingegneria.unisa.it/](http://www.ingegneria.unisa.it/) - [Copia cache](#) - [Simili](#)

## [Università di Salerno](#)

Gruppo Collegato di Salerno - Sezione di Napoli. Responsabile: M. Fusco Girard - HomeDip.

# Web Languages Evolution: XML

**XML** (eXtensible Mark-up language) add some support to HTML for capturing the meaning of the web contents.

It allows to create personalized tags (e.g. DTD)

Enables to add metadata to define the content of a web source.

# Ex. File XML

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<catalogue>
```

```
<book="1">
```

```
<author>Cesare Pavese</author>
```

```
<title>La casa in collina</title>
```

```
</book>
```

```
<book="2">
```

```
<author>Francesco Petrarca</author>
```

```
<title>Il Canzoniere</title>
```

```
</book>
```

```
</catalogue>
```

# DTD and XML Schema

The Document Type Definition defines the allowed elements (and their structure) in a XML Document

It is not necessary a DTD to form well formed XML Documents, but it is necessary to form validated ones.

XML Schema – XSD (XML Schema Definition).

# XML Limits

It doesn't allow the interchange of information between different resources (it's only possible to annotate a single source).

A new language (from XML) has been developed:  
RDF.

# RDF

Resource Description Framework (RDF) allows to describe resources (usually web sources) by using a triple of:

- Resource (identified by an URI)
- Property
- Values

# An RDF File

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<Book rdf:about="http://bookandbook.it/PublisherName"
```

```
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns="http://mybooks.it/book#">
```

```
<title> Il Canzoniere</title>; e.g. Property (title) – Value (Il Canzoniere)
```

```
<author rdf:resource="http://author.it/Petrarca">
```

```
<Year> 1970</Year>
```

```
<page> 230 </230>
```

```
</book>
```

# RDF Limits

e.g: an RDF file doesn't allow to define a level of hierarchy between the represented resources.

- RDFS

# RDFS

RDFS stands for RDF Schema

It is an XML structure based on RDF

Allows to define hierarchy ( $\Rightarrow$  inheritance)  
and structured relations among resources.

# RDFs

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-
ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xml:base="http://www.pubblicazioni.it/standard#">

  <rdfs:Class rdf:ID="Rivista">
    <rdf:subClassOf
rdf:resource="#PubblicazioneCartacea"></rdf:subClassOf>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Libro">
    <rdf:subClassOf rdf:resource="#PubblicazioneCartacea"></rdf:subClassOf>
  </rdfs:Class>

  <rdfs:Class rdf:ID="Sito">
    <rdf:subClassOf rdf:resource="#PubblicazioneOnline"></rdf:subClassOf>
  </rdfs:Class>

  <rdf:Property rdf:ID="scrittoDa">
    <rdf:domain rdf:resource="#Libro"></rdf:domain>
    <rdf:range rdf:resource="#Autore"></rdf:range>
  </rdf:Property>

</rdf>
```

# RDFS limits

RDFS **too weak** to describe resources in sufficient detail

- No **localised range and domain** constraints
  - Can't say that the range of hasChild is person when applied to persons and elephant when applied to elephants
- No **existence/cardinality** constraints
  - Can't say that all *instances* of person have a mother that is also a person, or that persons have exactly 2 parents
- No **transitive, inverse or symmetrical** properties
  - Can't say that isPartOf is a transitive property, that hasPart is the inverse of isPartOf or that touches is symmetrical
- ...

Difficult to provide **reasoning support**

- No “native” reasoners for non-standard semantics

From Bechhofer et al. Tutorial on OWL (2003)

# OWL

OWL (Ontology Web Language) => (DAML + OIL) is a W3C (World Wide Web Consortium) standard for building KRS (ontologies) in Semantic Web.

It adds vocabulary for describing:

- – relations between classes (e.g. disjointness)
- – cardinality (e.g. “exactly one”)
- – equality
- – richer typing of properties
- – characteristics of properties (e.g. symmetry).

# OWL

The Web Ontology Language (OWL) is a language for defining and instantiating *Web ontologies*.

It is based on a description logic model that makes possible to define and describe concepts and to make reasoning about them.

In OWL an ontology is a knowledge representation system.

OWL has three sub-languages: OWL Lite, OWL-DL e OWL-Full.

# OWL sub-languages

**OWL-Lite** is the simplest sub-language and it is used for simple class hierarchy and simple constraints.

**OWL-DL** is based on **Description Logics**. It maintained the decidability or computational completeness.

OWL DL support Description Logic and for that reason has good computational properties for reasoning systems 😊

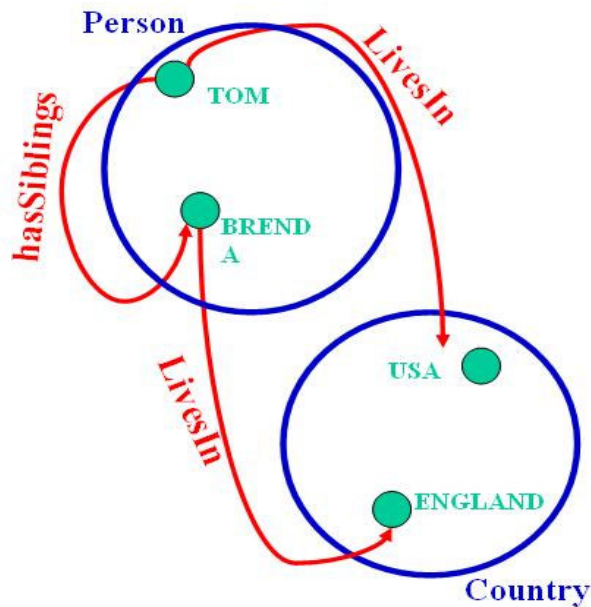
**OWL-Full** is the most expressive sub-language but doesn't guarantee decidability and computational completeness.

# OWL Ontologies

With **OWL** it's possible to create ontologies and describe the knowledge we have in a specific domain (**classes**, **relationships** between classes and **individuals** belonging to classes).

This formalized knowledge is processable automatically by a machine through a reasoner that implements inferential and deductive processes.

# OWL Elements



Individuals: objects in the domain of interest, are named *Instances* in Protegè . (In figure are green)

Properties: binary relations on individuals (In figure are red)

Classes: sets that contain individuals and organized into superclass-subclass hierarchy that is known as taxonomy. (In figure are blue)

# What is an ontology? (1)

Ontology in Philosophy: the “science of being in qua being” (Aristotle, *Metaphysics*, IV, 1).

An ontology is a document or file that formally defines the relations among terms. The most typical kind of ontology for the Web has a taxonomy and a set of inference rules. (T.B.Lee 1991).

Article: G.Vetere, *Nova 24 Il Sole 24*. Che razza di ontologi sono gli informatici?

# What is an ontology? (2)

An ontology is an explicit specification of a conceptualization. A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose.(T.R. Gruber 1993)

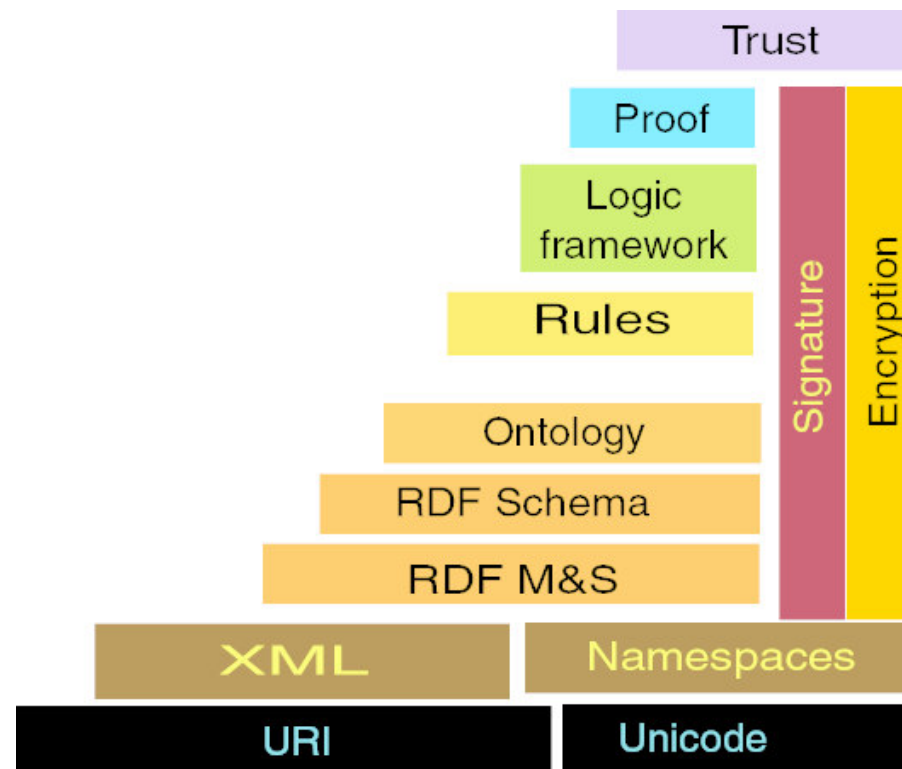
In other terms: it's a formal knowledge representation of a specific domain (or world).

# Ontologies and Semantic Web

To obtain a **structured representation** of the information through the ontologies is one of the main objectives in order to realize the so called Semantic Web (T.B. Lee et al., 2001).

According to T.B. Lee's vision, Semantic web should enable the machines to “understand” the semantics of the web resources and, therefore, to have a more “intelligent” behaviour in their activities of search.

# Semantic Web Infrastructure

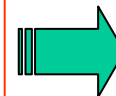
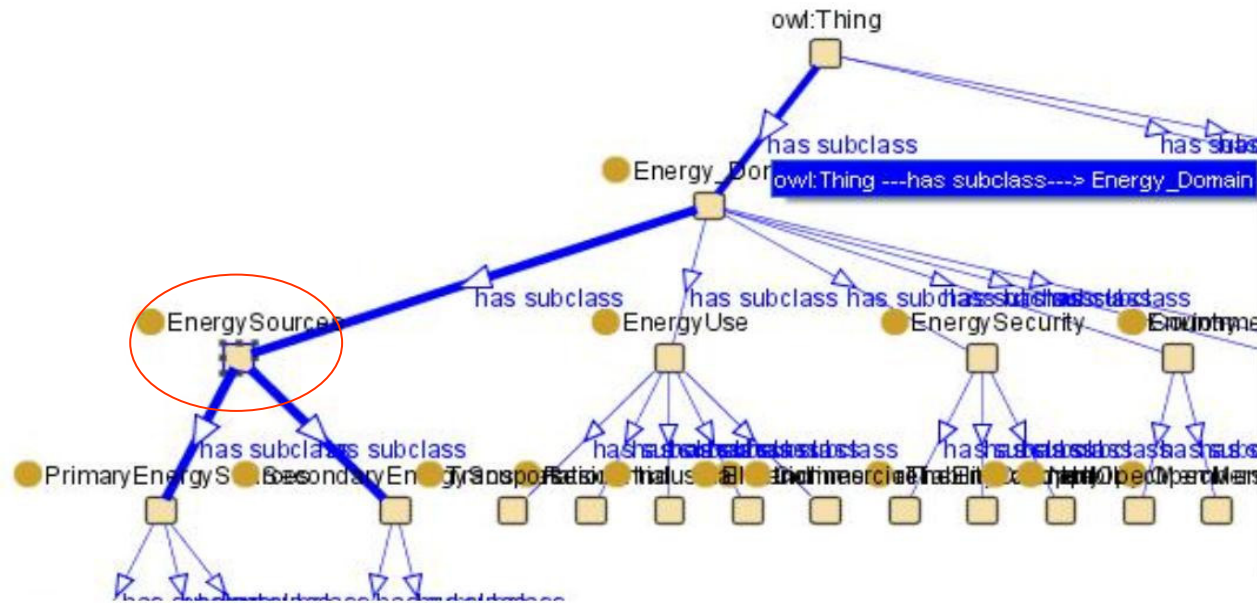
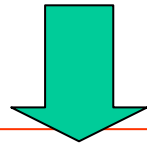


Source [www.w3.org](http://www.w3.org)

# Ontologies and Semantic Web (2)

SEARCH ENGINE

energy sources



Decision Maker

# OWL Ontology code: a toy example...(Dept.)

ie.pdf (Oggetto application/pdf) - Mozilla Firefox  
Visualizza Cronologia Segnalibri Strumenti Aiuto  
http://www.di.uniba.it/~nadja/sysag/  
Come iniziare Ultime notizie  
OWL Web Ontology Language Gui... reasoning.pdf (Oggetto applicatio... Risultato della ricerca immagini di ... ontologie.pdf (Oggetto applicati...  
Trova

OWLClasses Properties Forms Individuals Metadata

**SUBCLASS RELATIONSHIP**  
For Project: dipartimento

Asserted Hierarchy

- owl:Thing
  - Dipartimento
    - corsi\_in\_aula
    - corsi\_in\_lab
  - Staff
    - Professore
    - Studiante
    - Technico

**CLASS EDITOR**  
For Class: corsi\_in\_aula (instance of owl:Class)

Name SamsAs DifferentFrom

corsi\_in\_aula

rdfs:comment

```
<owl:Class rdf:ID="Dipartimento">  
  <owl:disjointWith>  
    <owl:Class rdf:ID="Staff"/>  
  </owl:disjointWith>  
</owl:Class>  
<owl:Class rdf:ID="corsi_in_aula">  
  <rdfs:subClassOf rdf:resource="#Dipartimento"/>  
  <owl:disjointWith>  
    <owl:Class rdf:ID="corsi_in_lab"/>  
<owl:FunctionalProperty rdf:about="#professore">  
  <rdfs:domain>  
    <owl:Class>  
      <owl:unionOf rdf:parseType="Collection">  
        <owl:Class rdf:about="#corsi_in_aula"/>  
      </owl:unionOf>  
    </owl:Class>  
  </rdfs:domain>  
</owl:FunctionalProperty>
```

Annotations

Property	Value
----------	-------

**Datatype property**

Properties

- nome\_corso (multiple xsd:string)
- professore (single Professore)

SUFFICIENT NECESSARY

**Object property**

Disjoints

- corsi\_in\_lab

From <http://www.di.uniba.it/~nadja/sysag/>

# OWL Classes and Properties

OWL classes are defined using the element:

*owl:Class*

Predefined classes:

*owl:Thing* and *owl:Nothing*

Properties:

*owl:ObjectProperty*, relations between instances of two classes.

*owl:DatatypeProperty*, relations between instances of classes and RDF literals and XML Schema datatypes.

# Datatypes Properties Ex.

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

```
<owl:DatatypeProperty rdf:ID="yearValue">
```

```
  <rdfs:domain rdf:resource="#WineYear" />
```

```
  <rdfs:range rdf:resource="&xsd;positiveInteger"/>
```

```
</owl:DatatypeProperty>
```

From <http://www.w3.org/TR/owl-guide/>

# Examples of OWL Properties

- Cardinality: (owl:minCardinality, owl:maxCardinality);
- Simmetric:(owl:SymmetricProperty);
- Transitive: (owl:TransitiveProperty);
- Functional: (max: 1 value for each object: ex: weight etc. (owl:FunctionalProperty);
- .....

# Description Logics

Description Logics are a class of formalisms that allows the representation of a terminological knowledge of a domain (e.g. ontologies are vocabularies).

DL' KRS are different from Semantic Networks and Frames: both do not provide a formal well defined semantics.

DL's allows, for the domain modelling, the use of:

- classes
- objects
- relations (rules)

On those notions can be inserted constraints and restrictions.

# Reasoning Capabilities in DL Systems

- Subsumption Algorithms: determine superconcept – subconcept relations.
- Instantiation Algorithms: determines instances relations (ex. Individual C is an instance of the Class A).
- Consistency Algorithms: determine if a knowledge base is non contradictory.

# OWL Ontologies and Description Logics

Logical Restriction are used to restrict the individuals that belong to a class. Two main categories.

1. ( $\exists$ ) **Existential restriction**: for a set of individuals it specifies the existence of a (at least one) relationship along given property to an individual that is a member of a specific class.
2. ( $\forall$ ) **Universal Restriction**: constrain the relationships along a given property to individuals that are members of a specific class.

# A simple example (energy domain)

Verbal Proposition : Some Fossil Fuels cause some environmental consequences or some Risks

First Order Predicate Logic:  $\exists x (Fx \rightarrow Ce \vee Cr)$

Protegé Construction:  $\exists$  Fossil Fuels cause some (Environmental Consequences or Risks)

# The example in Protégé...

Ontoenergy Protégé 3.3.1 (file:\C:\Documents%20and%20Settings\Antonio%20Lieta\Documenti\ontoenergy.pprj, OWL / RDF Files)

Edit Project OWL Code Tools Window Help

Metadata (Ontology1195308540.owl) OWLClasses Properties Individuals Forms Queries Jambalaya

**CLASS EXPLORER**

Project: ontoenergy

Sorted Hierarchy

- owl:Thing
  - Energy\_Domain
    - Country
      - NonOpecMembers
      - OpecMembers
    - EnergySecurity
    - EnergySources
      - PrimaryEnergySources
        - NonRenewable
          - FossilFuel
        - Nuclear
        - Renewable
      - SecondaryEnergySources
        - SecondaryNonRenewable
        - SecondaryRenewable
    - EnergyUse
    - EnvironmentalConsequences
    - Infrastructures
  - Risks
  - Solutions
    - LongTermActions
    - ShortTermActions

**CLASS EDITOR**

For Class: FossilFuel (instance)

Property	Value
rdfs:comment	

NonRenewable

cause some (EnvironmentalConsequences or Risks)

# Ontology Engineering and Modelling

Poesio (2005) states that there are, at least, two different research traditions in the domain modelling literature.

One school of thought supports the thesis of the need of more rigorous logical and philosophical foundations for domain modeling formalisms.

The second school of thought, instead (that Poesio defines as “cognitive”), argue that the best way to identify epistemological primitives is to study concept formation and learning in humans.

# Ontology Engineering and Modelling (2)

The first approach's aim is both to establish a “Tarskian Semantics” for the formalism used in the domain ontologies (leading to description logics) and to have cleaner domain ontologies.

“clean ontology” = “ontology with a clear semantics”.

# Ontology Engineering and Modelling (3)

The second approach is more “empiristic”.

It argues that the best approach to the construction of domain ontologies is by the use of machine learning techniques to automatically extract ontologies from language corpora: language based point of view, no conjectures made *a priori* by the ontologists .

# Ontology Types

- **Top-level ontologies**

General concepts. Domain independent.

- **Domain ontologies**

Vocabulary of a generic domain. Ex: biology, physics.

- **Task ontologies**

Vocabulary for a specific task or activity. Ex: selling.

- **Application ontologies**

Specialization of domain or task ontologies

# LOGICA DEI LINGUAGGI NATURALI E ARTIFICIALI

Towards the Semantic Web

Antonio Lieto - University of Salerno

[alieto@unisa.it](mailto:alieto@unisa.it)

<http://antoniolieto.blogspot.com>