

# Ontologies and data

The application and limits of OWL

Aldo Gangemi Semantic Technology Lab ISTC-CNR, Rome

# Why suq semantics?

The suq is where writing and rationality were born

A metaphor of current knowledge in the age of Internet

The Web of Data ... incredibly cheap information ... the Web as a new mass medium



### Phenomena

Extracting linguistic invariances, collecting semi-structured and deep web data, reusing metadata, open datasets, linguistic resources, ontologies, searching and integrating it all

### Methods

Reengineering, Linking, (Meta) Modeling, Mapping, Formalizing

### **Visions**

Reconciling social/cognitive/lexical with formal semantics

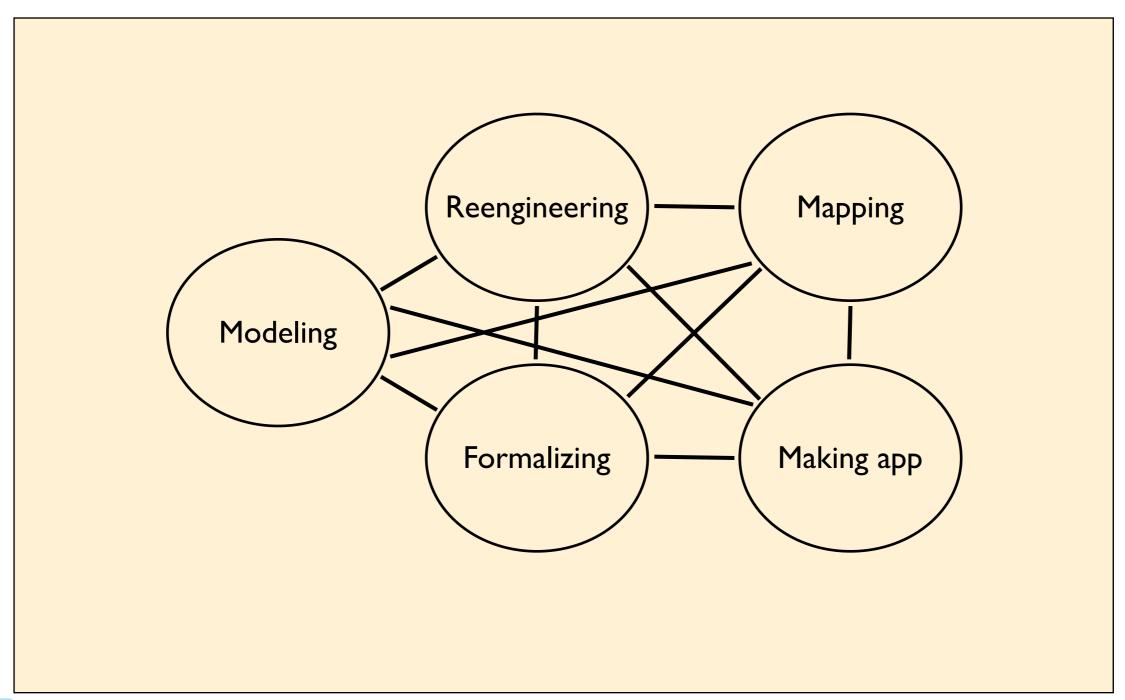
Presenting and exploiting efficiently





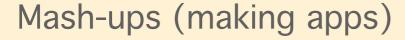
**Lah** Fechnology Lah

## Aspects of dealing with suq semantics

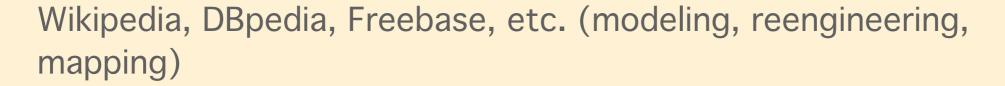




# A lot of data in the web suq



Linked open data (mapping)



Triplify, GRDDL, RDFa, SKOS, SIOC, etc. (reengineering)

Semantic Interoperability (making apps, reengineering, formalizing)



## Annotation on the Web (1)

Web 2.0: Flickr, del.icio.us, etc.

Annotation of Web pages with tags

Semantics depends on intention of annotator



### rewind-button



#### Commenti



#### Vectormonkey [cancellato] dice:

Wierd! I mean, cool, but wierd...
Postato 38 mesi fa. ( link permanente )



mohawk pro dice:

thats all i could hope for!

Postato 37 mesi fa / link permanente )

lips foolish

dance

art.

3 love
3 girl

style
fashion

alternative

chaos

psychosis desire

3 biscuit

3 teacup

ceramicpaisley

models

bald loser

a new

3 home

mohawk
photo

③ S

S limbert

3 look
3 see

surprise

veiw

sean
d

photography

magisch

艺术

③ 藝術

κunstτέχνη

③ 芸術

3 예술

искусствоо

arte

○ 魔术
○ 魔術

+ 1

### Questa foto appartiene anche a:



:Utterly Surreal:. (Pool)

Funny Ha Hal (Pool)

FlickrCentral (Pool)

+ ART DISTRICT (Pool)

"Plastic People" or mannequins, dolls and more (Pool)

+ Kitsch Itch (Pool)

The Secret Life of Toys (Pool)

Art - eye of the beholder / upload 1,

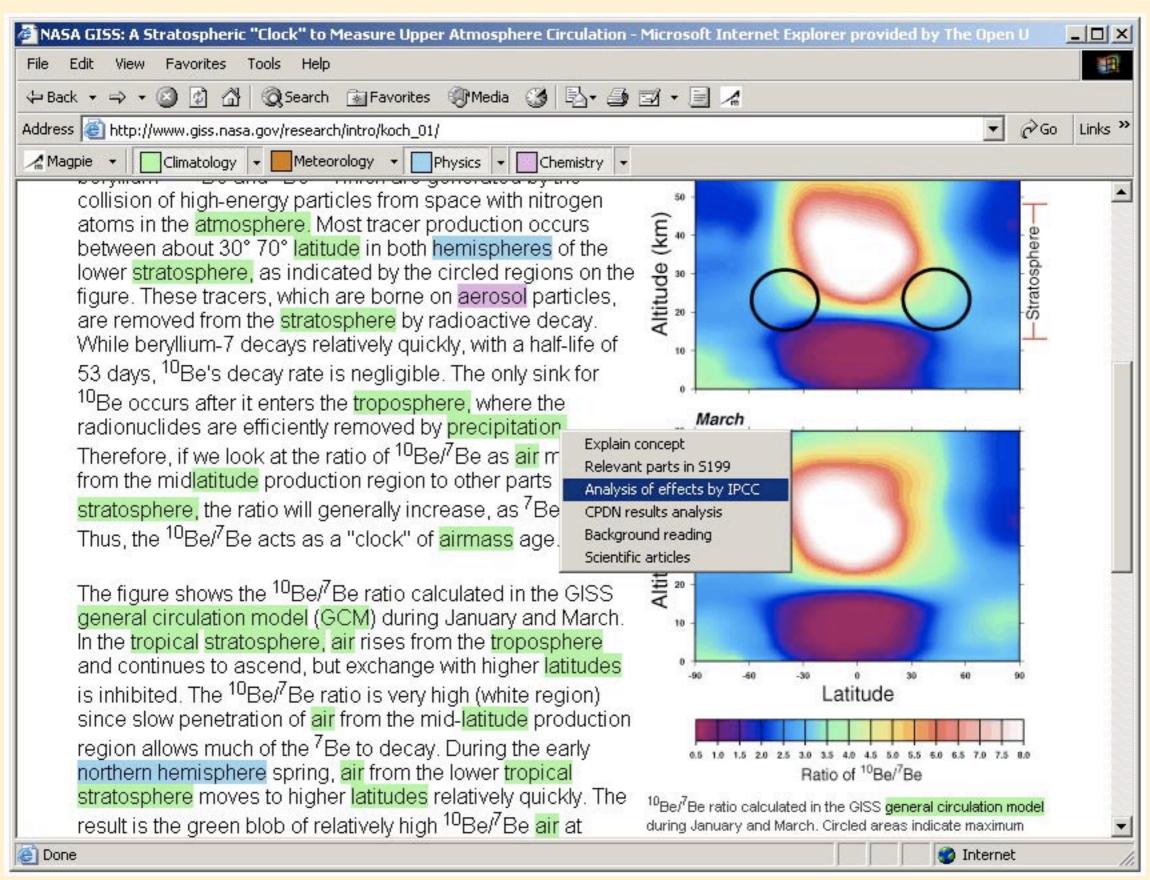
# Annotation on the Web (2)

Semantic Web

Annotation of Web pages with tags interpreted with reference to a logical language (e.g. OWL-RDF)

Semantics depends on intention of annotator, and on the constraints deriving from the reused ontology







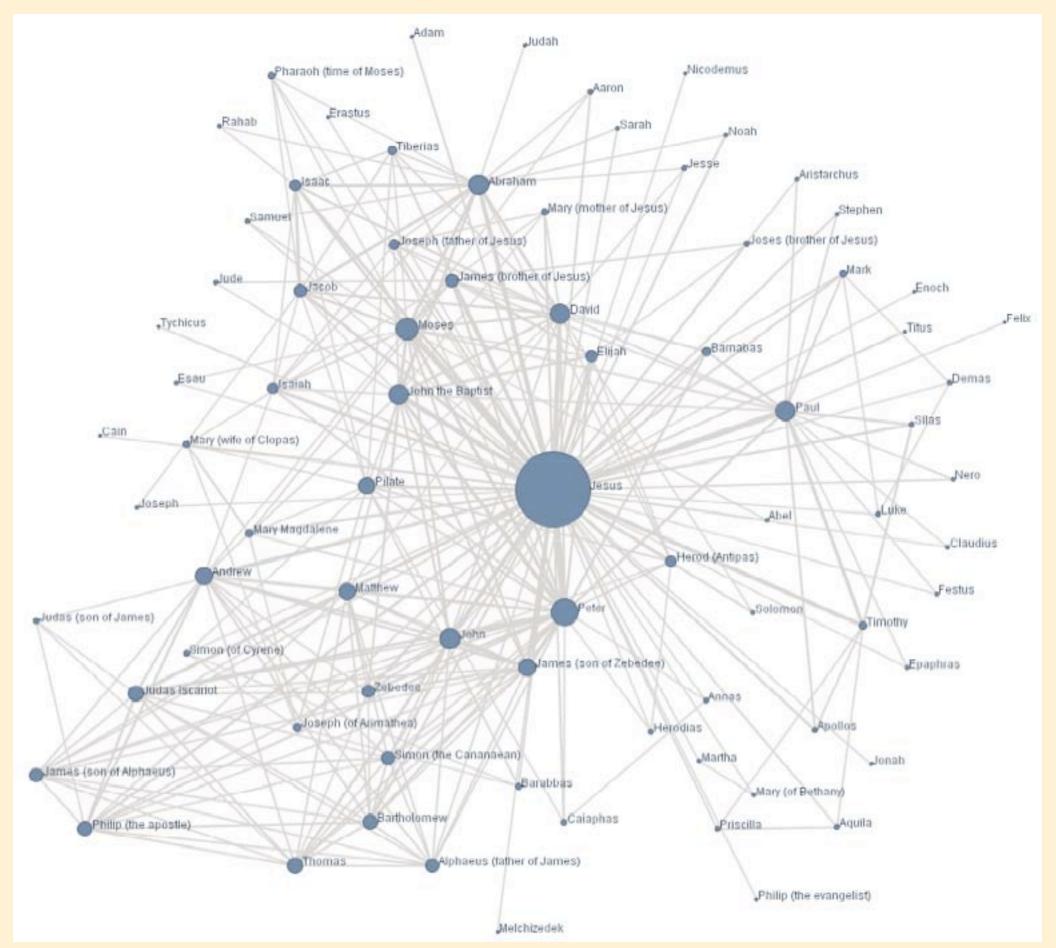
### Social networks

FaceBook, MySpace, Friendster, Orkut, LinkedIn, FOAF, SIOC

Similarity between Web users; various formats

Semantics depends on the model of the application, and on how distance between the nodes is computed







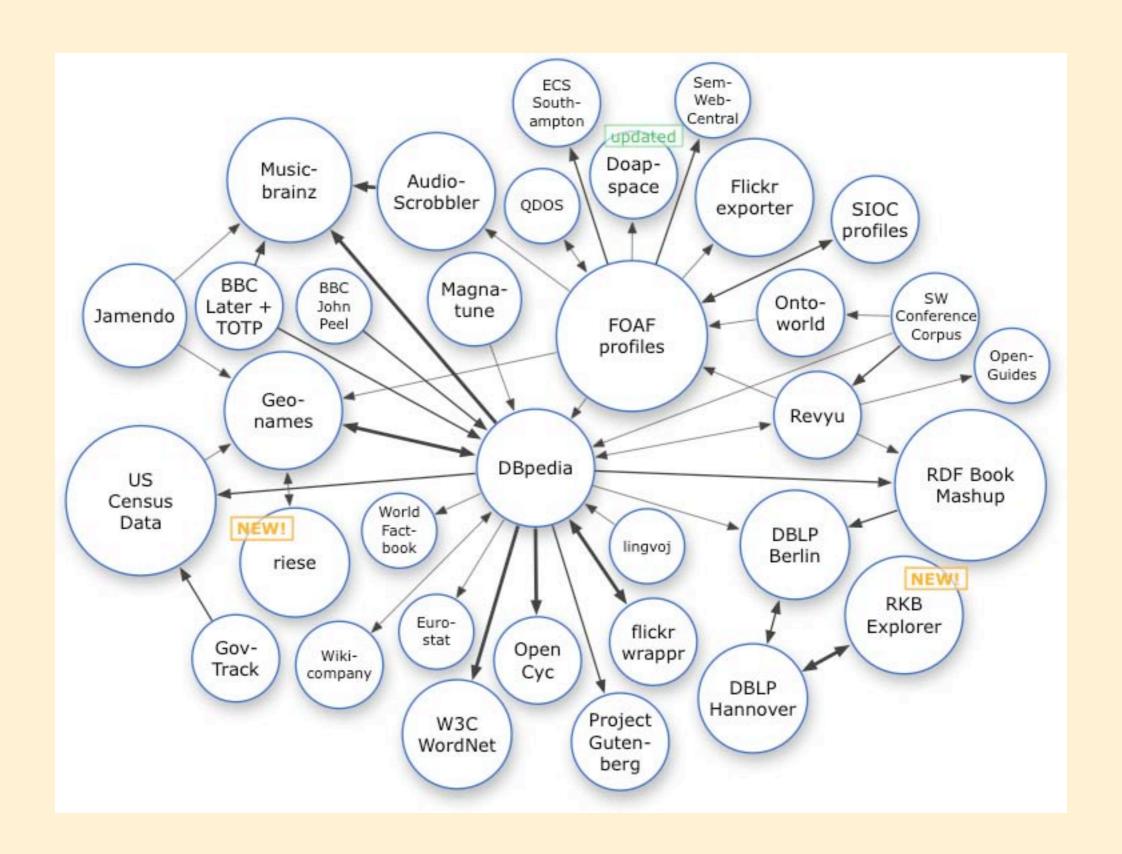
# RDF data ("triples")

### Linked Data

Dataset extracted from heterogeneous sources, and triplified.
Associations (linkings) are sometimes added in direct, naïve ways

Semantics depends on intended semiotics of data and relations assumed for linking







# Knowledge retrieval

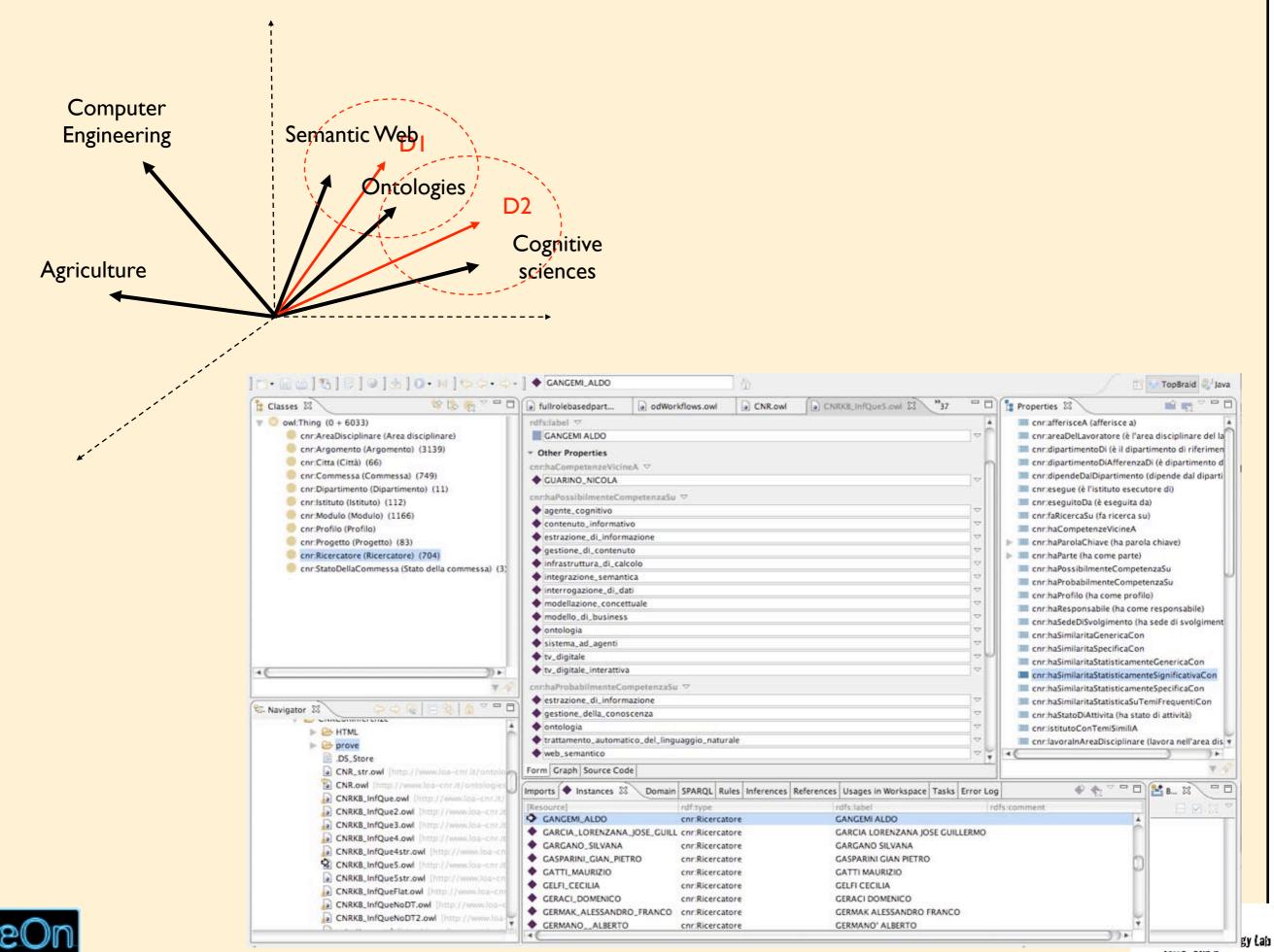
NLP, LSA, IE, etc.

Automatic association of non-structured information

Sources have a linguistic semantics

Depending on intended semiotics for extraction/learning (what features, how related, etc.)





### Data and behavior

Virtual worlds: where are the data?

Semantics of dynamic systems: Second Life, Wii, ...?

Separation of data and presentation

Also behavioral data

Also dynamics of non-behavioral data



# Dealing with suq semantics: current state

Much enthusiasm, a lot of nice ideas

Much confusion and mutual misunderstanding between 'and "neats"

Pushing formal semantics beyond its limits (e.g. the "owl:sameAs" dispute)



Mixing up strings, classes, terms, concepts, topics, tags, etc.

Suppose we need to design a desire ontology: where to start from?





### Ontology-related data

Informal vs. formal

Text corpora

Folksonomies (tag sets, directories, topic trees, subject indexes, infoboxes)

Lexica (dictionaries, wordnets, terminologies, nomenclatures)

Knowledge organization systems (thesauri, classification schemes)

Frames, semantic networks

DB schemas

Linked Open Data datasets

(Computational) ontologies

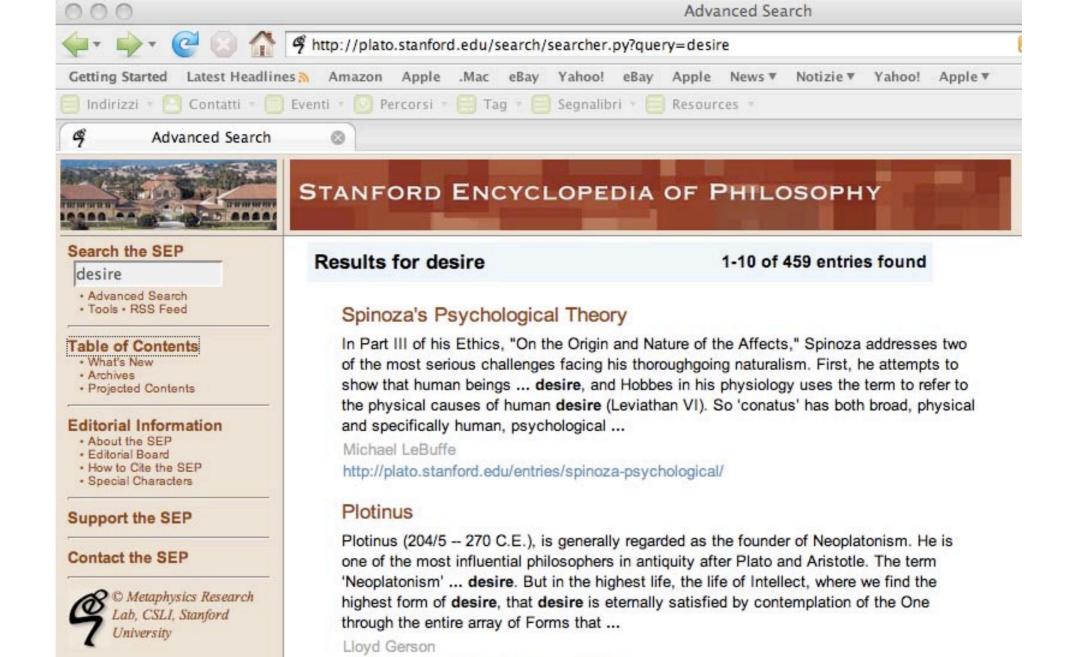


### Text ...

Plena mujer, manzana carnal, luna caliente, espeso aroma de algas, lodo y luz machacados, qué oscura claridad se abre entre tus columnas? Qué antigua noche el hombre toca con sus sentidos? Ay, amar es un viaje con agua y con estrellas, con aire ahogado y bruscas tempestades de harina: amar es un combate de relámpagos y dos cuerpos por una sola miel derrotados. Beso a beso recorro tu pequeño infinito, tus márgenes, tus ríos, tus pueblos diminutos, y el fuego genital transformado en delicia corre por los delgados caminos de la sangre hasta precipitarse como un clavel nocturno, hasta ser y no ser sino un rayo en la sombra.

[Pablo Neruda, Cien sonetos de amor]





### Encyclopedia ...

#### Moral Motivation

In our everyday lives, we confront a host of moral issues. Once we have deliberated and formed judgments about what is right or wrong, good or bad, these judgments tend to have a marked hold on us. ... desire or disposition of the individual herself. Second, moral properties not only motivate entirely on their own: they provide overriding motivation. Of course, their motivational power depends on an ...

Connie S. Rosati

http://plato.stanford.edu/entries/moral-motivation/

http://plato.stanford.edu/entries/plotinus/

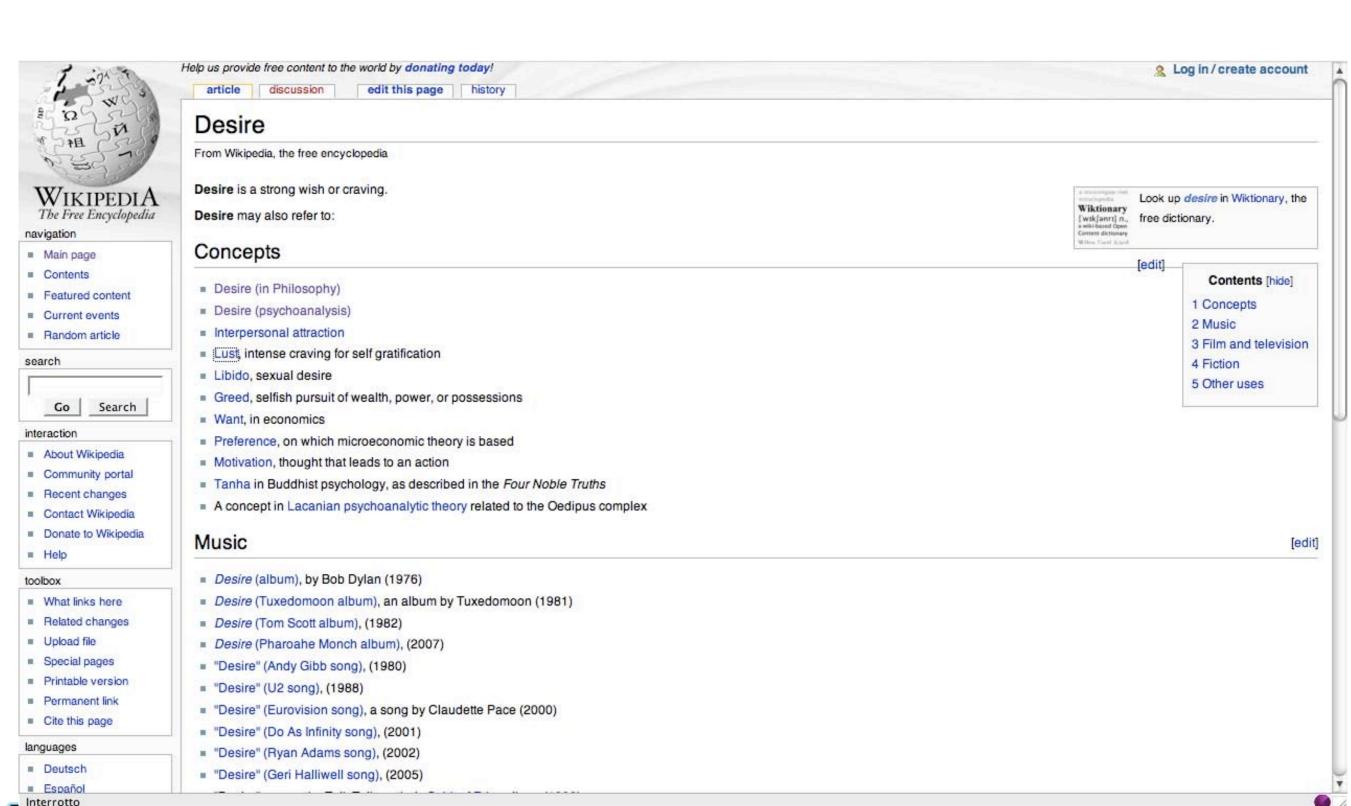
#### Well-Being

Well-being is most commonly used in philosophy to describe what is non-instrumentally or ultimately good for a person. The question of what well-being consists in is of independent interest, but it ... desire theories, or objective list theories. According to the view known as welfarism, well-being is the only value. Also important in ethics is the question of how a



STLab The Semantic Technology Lab IST'C-CNR Rome

### Wikipedia ...





### Linguistic dictionaries and thesauri

### Oxford American Dictionary

 $desire \hspace{0.1cm} |\hspace{0.1cm} \text{d} \hspace{0.1cm} \text{$ 

#### noun

a strong feeling of wanting to have something or wishing for something to happen: [with infinitive ] a desire to work in the dirt with your bare hands.

• strong sexual feeling or appetite: they were clinging together in fierce mutual desire.

verb [trans.]

strongly wish for or want (something): he never achieved the status he so desired | [as adj. ] ( **desired**) it failed to create the desired effect.

- want (someone) sexually: there had been a time, years ago, when he had desired her.
- archaic express a wish to (someone); request or entreat.

ORIGIN Middle English: from Old French desir (noun), desirer (verb), from Latin desiderare (see desiderate).

#### **Thesaurus**

### desire

#### noun

1 a desire to see the world wish, want, aspiration, fancy, inclination, impulse; yearning, longing, craving, hankering, hunger; eagerness, enthusiasm, determination; informal yen, itch, jones.

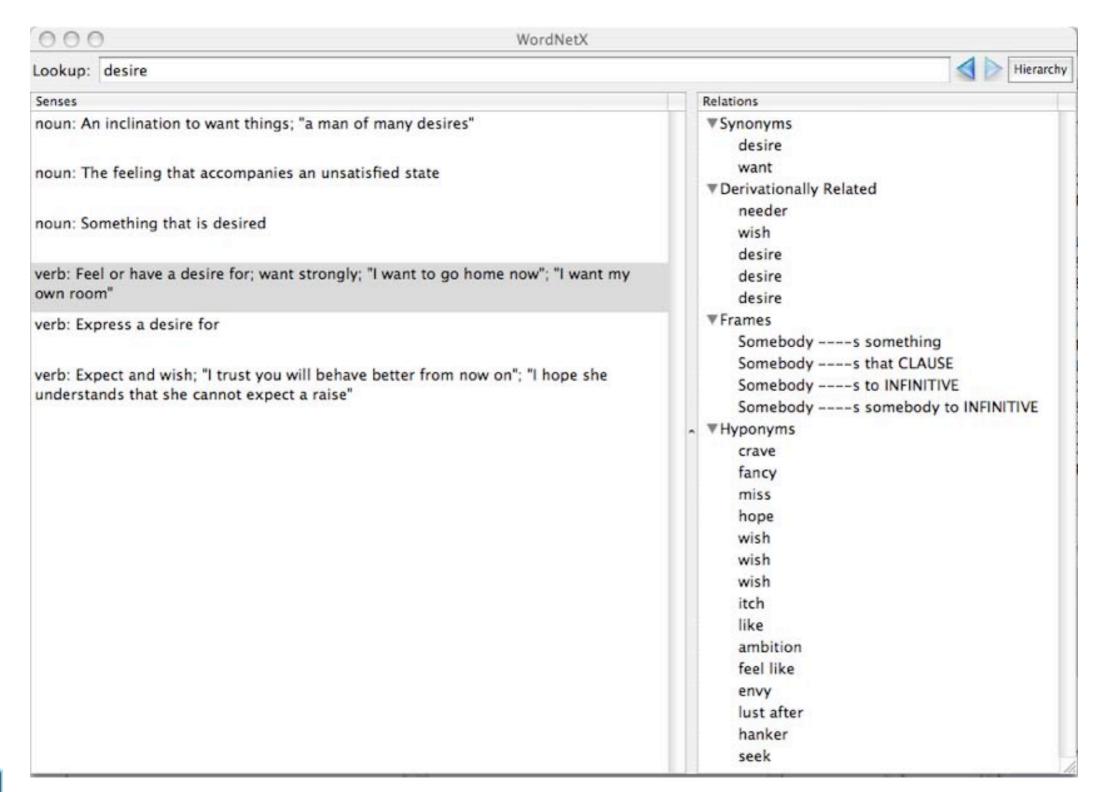
2 his eyes glittered with desire lust, sexual attraction, passion, sensuality, sexuality; lasciviousness, lechery, salaciousness, libidinousness; informal the hots, raunchiness, horniness. verb

1 they desired peace want, wish for, long for, yearn for, crave, hanker after, be desperate for, be bent on, covet, aspire to; fancy; informal have a yen for, have a jones for, yen for, hanker after/for.

2 she desired him be attracted to, lust after, burn for, be infatuated by; informal fancy, have the hots for, have a crush on, be mad about, be crazy about.

TLab

### WordNets ...





### FrameNets ...

Frame Report (recent data)

| Top of Frame Index | Top of Lexical Unit Index |

### **Desiring**

#### Definition:

An Experiencer desires that an Event occur. (Note that commonly a resultant state of the Event will stand in for the Event.) In some cases, the Experiencer is an active participant in the Event, and in such cases the Event itself is often not mentioned, but rather some Focal participant which is subordinately involved in the Event.

Generally, the use of a word in this frame implies that the specific Event has not yet happened, but that the Experiencer believes that they would be happier if it did. Sometimes the Time\_of\_Event, Purpose\_of\_Event, or the Location\_of\_Event are mentioned without the explicit mention of the Event.

only WANTED one piece of candy.

The company was EAGER for him to leave as soon as possible

Susan really WISHES that you 'd listen to her

FEs:

Core:

Semantic Type
State\_of\_affairs

The ch

The change that the Experiencer would like to see.

Experiencer [Exp]
Semantic Type Sentient

The Experiencer is the person (or sentient being) who wishes for the Event to occur.

Focal\_participant [Foc] This is the entity

This is the entity that the Experiencer wishes to be affected by some Event.

Location\_of\_Event [PEv]

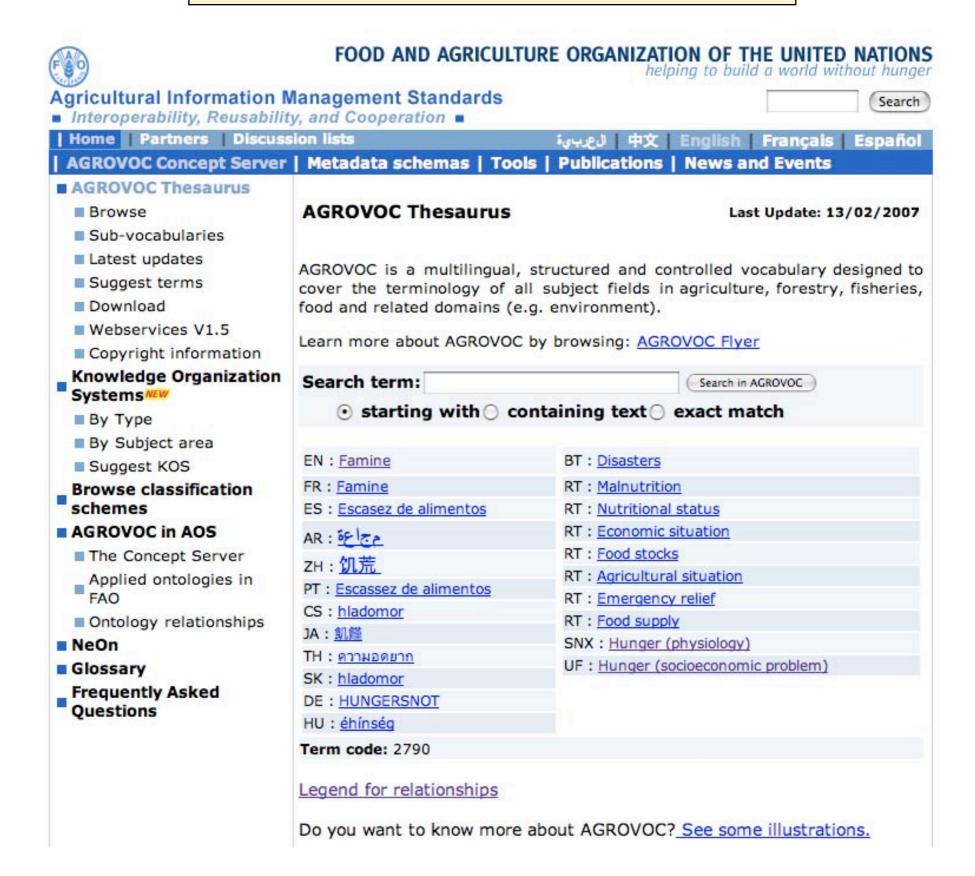
The Location\_of\_Event is the place involved in the desired Event.

I WANT that box on top of the other one.

The prince WISHES you here before matins.



### Thesauri ...





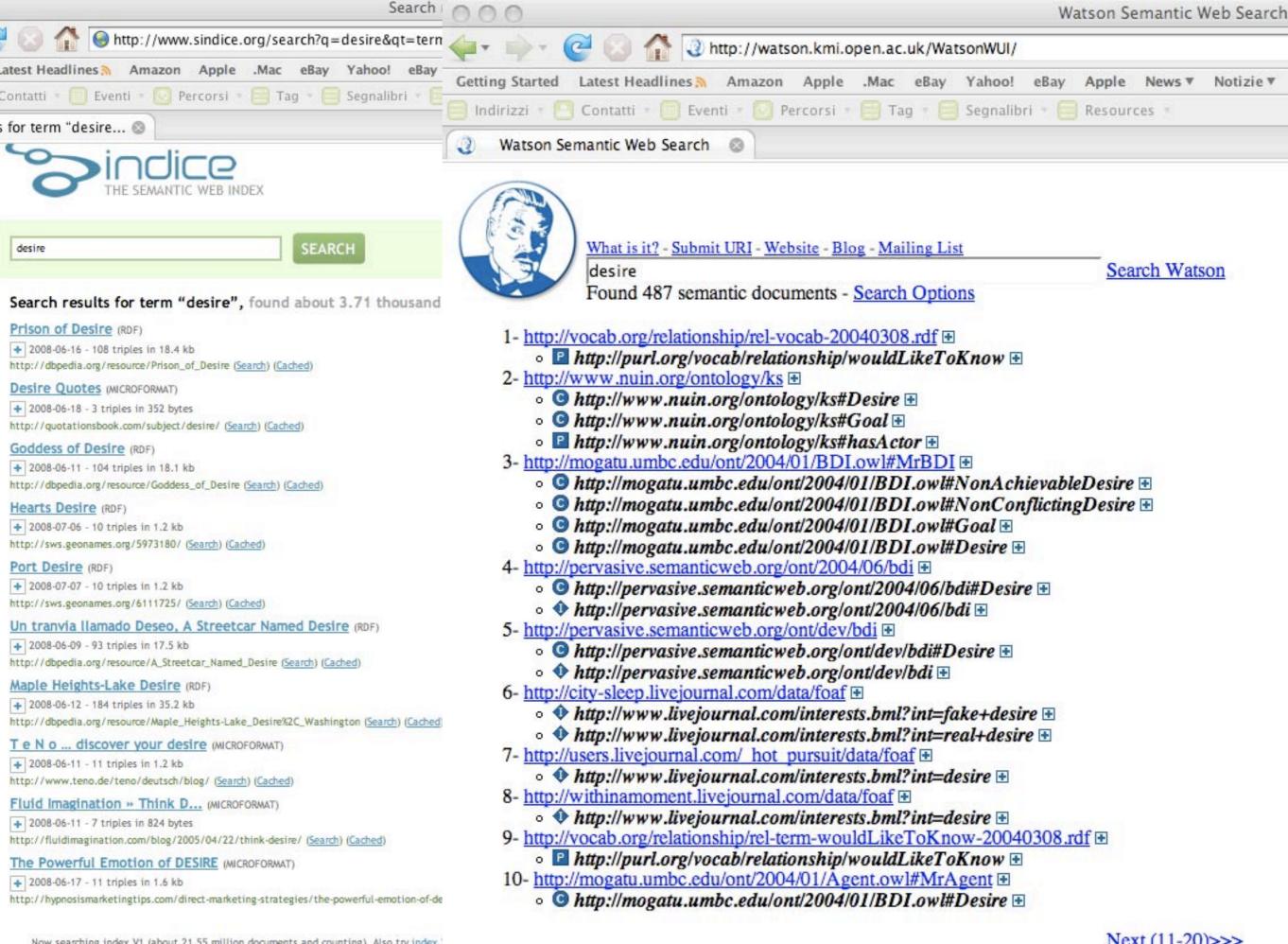
# Standard languages help

Transform all in RDF, or even OWL

Cf. Triplify initiative

Then search for RDF or make integrating apps





### Integrated knowledge search

#### YAGO-query:



#### Search Results

#### Narrow Results



Freebase

#### Items 1 - 30 of 60+

#### Desire



ona Composition

"Desire" is the lead single from U2's 1988 album, Rattle and Hum. It was also their first #1 single in the UK. It reached #3 on the Billboard Hot 100 in the US, and reached #1 on both the Mainstream and Modern Rock Tracks charts, the first song to reach the top of both of these charts. "Desire" debuted live on the first night of the Lovetown Tour on 21 September 1988, and appeared at almost every concert on that tour. It segued into a cover of Bob Dylan's "All Along the Watchtower", and the...

#### Desire



Musical Albu

Desire is singer-songwriter Bob Dylan's 17th studio album, released by Columbia Records in 1976. It is one of Dylan's most collaborative efforts, featuring the same caravan of musicians from the acclaimed Rolling Thunder Revue tours the previous year (later documented on The Bootleg Series Vol. 5). Most of the album was co-written by Jacques Levy, and is composed of lengthy story-songs, two of which quickly generated controversy: the over 11-minute long "Joey", which is seen as glorifying the...

#### Desire



TV Program

Desire is an American telenovela which debuted at 8 p.m. ET/7 p.m. CT on September 5, 2006, on the American network MyNetworkTV, and ended on December 5. It was produced by Twentieth Television. The program starred Sofia Milos, Michelle Belegrin, Nate Haden, Kelly Albanese, Zack Silva, Jessie Ward, Tanisha Harper. Haden and Silva played two brothers on the run from the Gamarras, a New Jersey crime family. They run from Bayonne to Los Angeles and become restaurateurs. Along the way, the pair...

#### Desire

Comic Book Character, Fictional Charact

Desire is one of the Endless, a fictional character from Neil Gaiman's comic book series, The Sandman. Desire is the third youngest of the Endless and the twin of Despair. It is a strikingly beautiful figure whose gender is mutable, becoming male, female, both, or neither as the situation warrants. It is often referred to as "sister-brother" by its siblings, particularly Dream. Desire blends in effortlessly with whatever environment it finds itself in. It lives in the heart of a massive flesh...

#### esire



"Desire" is a song recorded by Swedish singer Darin. It was released as the third single (and first digital only) from Darin's third studio album Break the News in Sweden on May 7 2007. In Germany and selected other European countries, it was released as the album's second single on December 7 2007. Darin performed the song on the finale of the German Popstars on December 6 2007. "Desire" debuted in Germany at number fifty-three in the last week of 2007.

# everything based on a centralized ontology ... of mixed quality

STLab

The Semantic Technology Lab
ISTC-CNR Rome

### Microformats

```
* hCal
       <span class="vevent">
       <abbr class="dtstart" title="2006-12-05">December 5-</abbr>
       <abbr class="dtend" title="2006-12-07">7th</abbr>
       </b> At <b><span class="summary">;XML 2006</span></b>
       (<span class="location">Boston, MA USA</span>)
      for a presentation on "Social Semantic Mashups".</span>
  * XFN for Social Networks
<a href="http://www.w3.org/People/Connolly/" rel="colleague">Dan Connolly</a>
<a href="http://seanmcgrath.blogspot.com/" rel="colleague">Sean McGrath</a>
<a href="http://www.jclark.com/" rel="colleague">James Clark</a>
  * hCard Example
       <div class="vcard">
       <span class="fn n">
       <span class="given-name">Harry </span>
       <span class="additional-name">Reeves</span>
       <span class="family-name"> Halpin
       <span class="tel">+44-131-650-4421</span>
       <
       <span class="street-address">2 Buccleuch Place</span>
```



### Microformats issues and GRDDL

The main problem with microformats is that they put your data into HTML,

but you have no standard way to get the data out.

Another problem is that they cannot be validated easily.

You can mix hCard and hCal and there's no way to guarantee you will interpret it correctly. Domain-specific: You can not make a microformat for just anything!

With GRDDL, microformat data can be viewed as Semantic Web data

Gleaning Resource Descriptions from Dialects of Languages

A markup for declaring that an XML document includes gleanable data (PROFILE) linking to an algorithm (typically XSLT) for gleaning the RDF data from the document (TRANSFORMATION)

- \* A large number of web sites are using microformats:
- \* Eventful
- \* LinkedIn
- \* Yedda
- \* Yahoo! Local
- \* Yahoo! Tech Reviews
- \* Dreamweaver plug-in to help authors

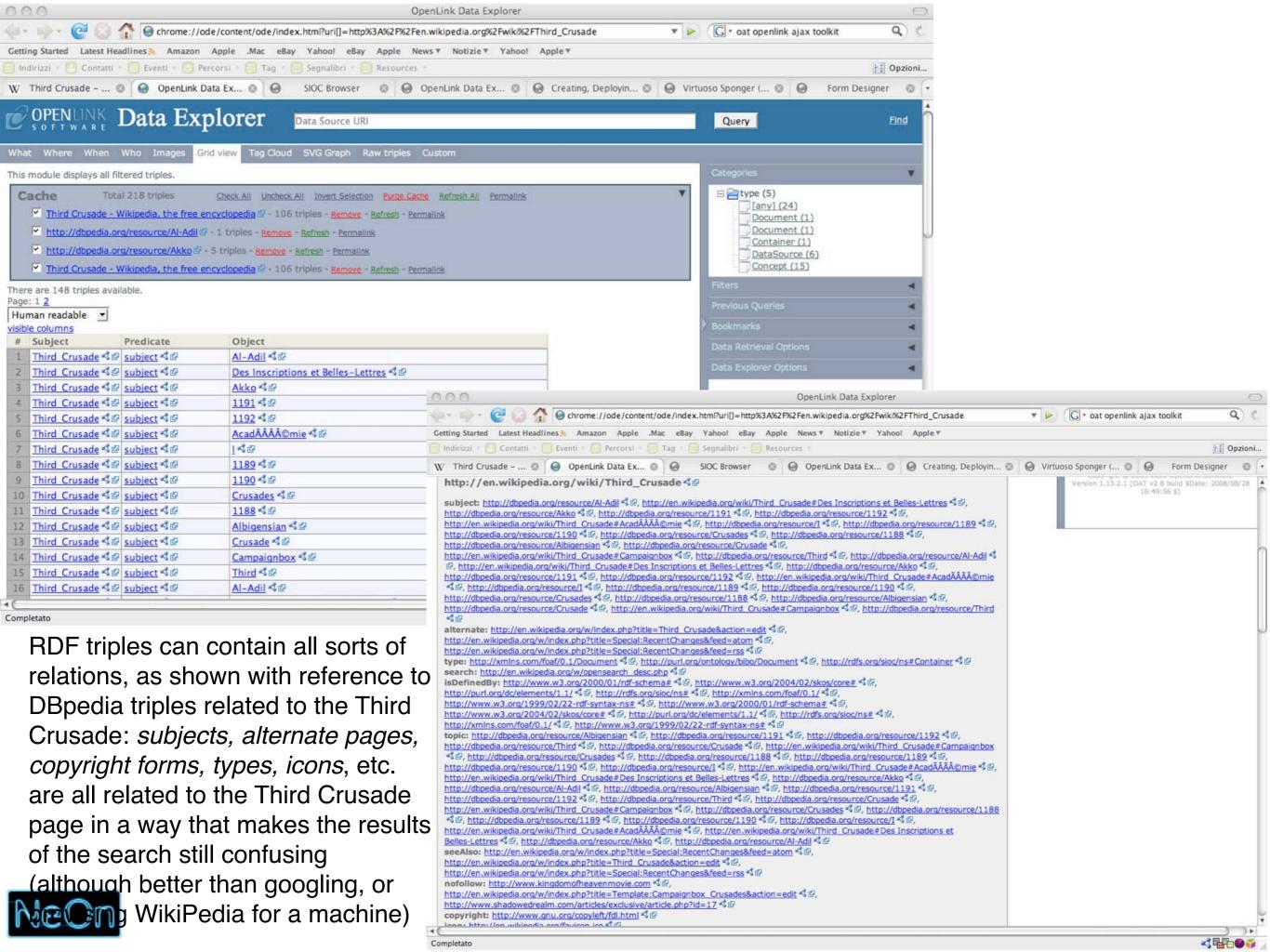
credits Gandon-Halpin-Adida



# Linking Open Data

LOD Presentation by K. Idehen







Semantic Technology Lab ISTC-CNR Rome

Now we have all those data expressed in a language that allows semantic interoperability ...



### What we can do with OWL

```
... (maybe) we can check the consistency, classify, and query all this knowledge
this is great, but ...
... when I locally reuse parts of such a big bunch of knowledge, inferences sometimes produce
strange results:
    a web page same as an email address (e.g. http://.../Aldo owl:sameAs mailto://aldo@...)
    a person same as a wikipedia article (e.g. Aldo owl:sameAs http://en.wikipedia.org/Aldo)
    Italy is a continent (e.g. (Italy rdf:type (Country) rdfs:subClassOf Continent))
... and problems are hardly fixable on a large scale
Logical consistency is not the main problem
    e.g. owl:sameAs can be wrongly used and still we have consistency
Why OWL is not enough?
```



### When to use owl:Individual, Class, ObjectProperty, DatatypeProperty?

OWL gives us logical language constructs, but does not give us any guidelines on how to use them in order to solve our tasks. E.g. modeling something as a class or an object property is mostly arbitrary

cf. Semantic Web Interest Group (<a href="mailto:semantic-web@w3.org">semantic-web@w3.org</a>) post on May 27th, 2008 by Zille Huma:

"I have been wondering for sometime now that why isn't it a popular trend to store standard activities of a domain in the ontology and not only the concepts, e.g., for the tourism domain, ontologies normally contain concepts like Tourist, Resort, etc. but I have not so far come across an ontology that also contains the standard activities like searchResort, bookHotel, etc. Why is it so? What support is provided in the ontology languages to model the standard activities of the domain as well?"

- (1) a functionality for searching resorts is implemented in our web service
   owl:Individual(searchResort) rdf:type(Functionality)
- (2) searching resorts is a type of functionality required for this kind of services
   owl:Class(searchResort) rdfs:subClassOf(Functionality)
- (3) who has been searching for what resorts in our web service?
   owl:ObjectProperty(searchResort) rdfs:range(Resort)
- (4) how many users have been using our resort searching functionality?
   owl:DatatypeProperty(searchResort) rdfs:range(xsd:boolean)



### solutions?

... OWL is not enough for building a good ontology and we cannot ask all web users either to learn logic, or to study ontology design

Reusable solutions are described as "Ontology Design Patterns", which help reducing arbitrariness without asking for sophisticated skills ...

... provided that good tools will be built :)

# Ontology Design Patterns (OP)

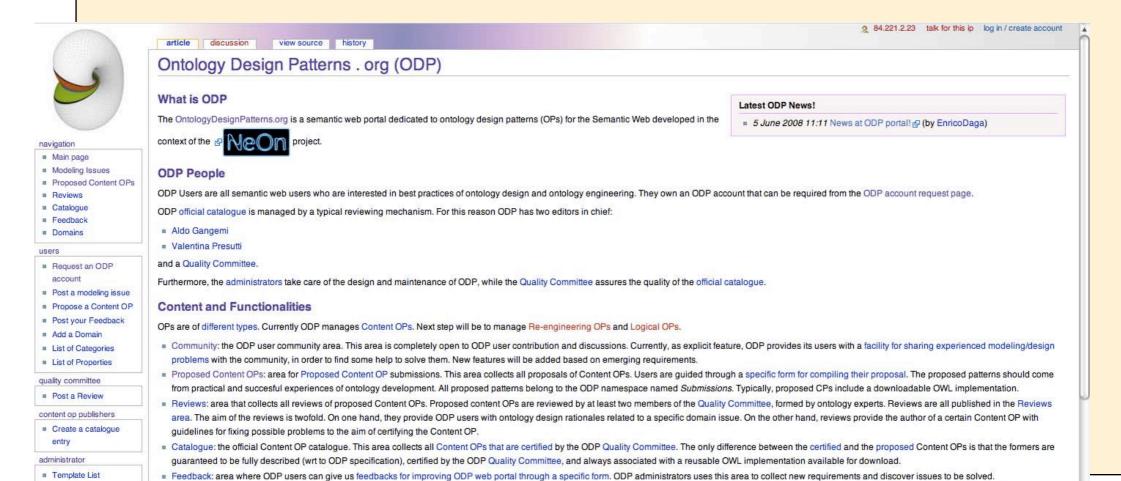
### Definition

■ Form list

Create a category

Create a property
 Create a template

An Ontology Design Pattern is a modeling solution to solve a recurrent ontology design problem



Domain: the list of domains. This page lists all domains that are defined in ODP, and provides users with a facility to create new ones. Each Content OP or Modeling Issue is associated to a domain, this is why this page is

important. Before to propose your Content OP or to post a modeling issue be sure your domain is already on the list or create a page for it.