

Semantic Web

Ing. Federico Chesani

Corso di Fondamenti di Intelligenza Artificiale M

Outline

1. Introduction

- a) The map of the Web (accordingly to Tim Berners-Lee)
- b) The current Web and its limits
- c) The Semantic Web idea

2. Semantic Information (a bird's eye view)

- a) Semantic Models
- b) Ontologies
- c) Few examples

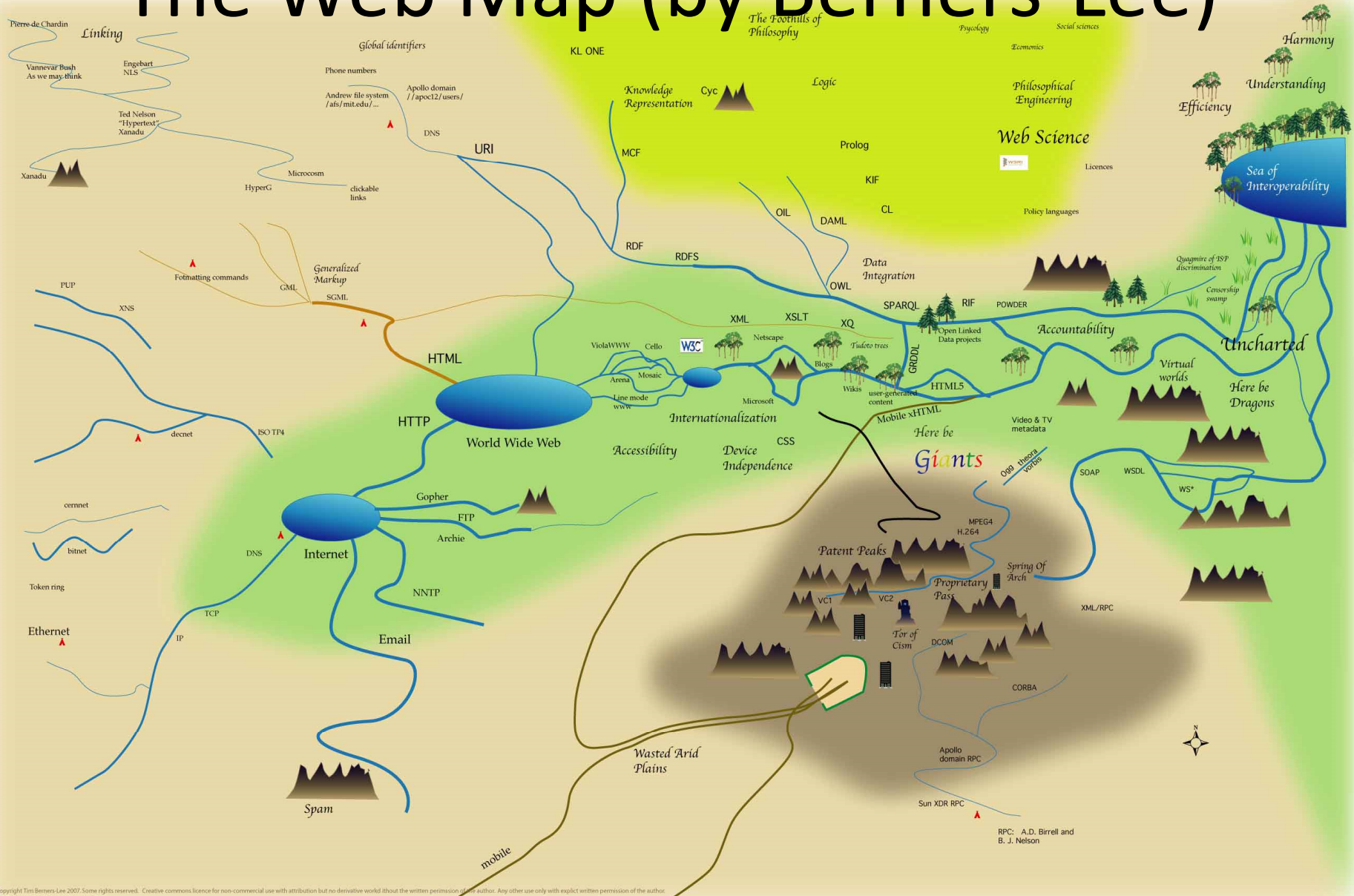
3. Semantic Web Tools

- a) Unique identifiers - URI
- b) XML
- c) RDF and SPARQL
- d) OWL

4. Semantic Web: where are we?

- a) Problems against the success of SW proposal
- b) Critics against SW
- c) Few considerations
- d) Few links to start with

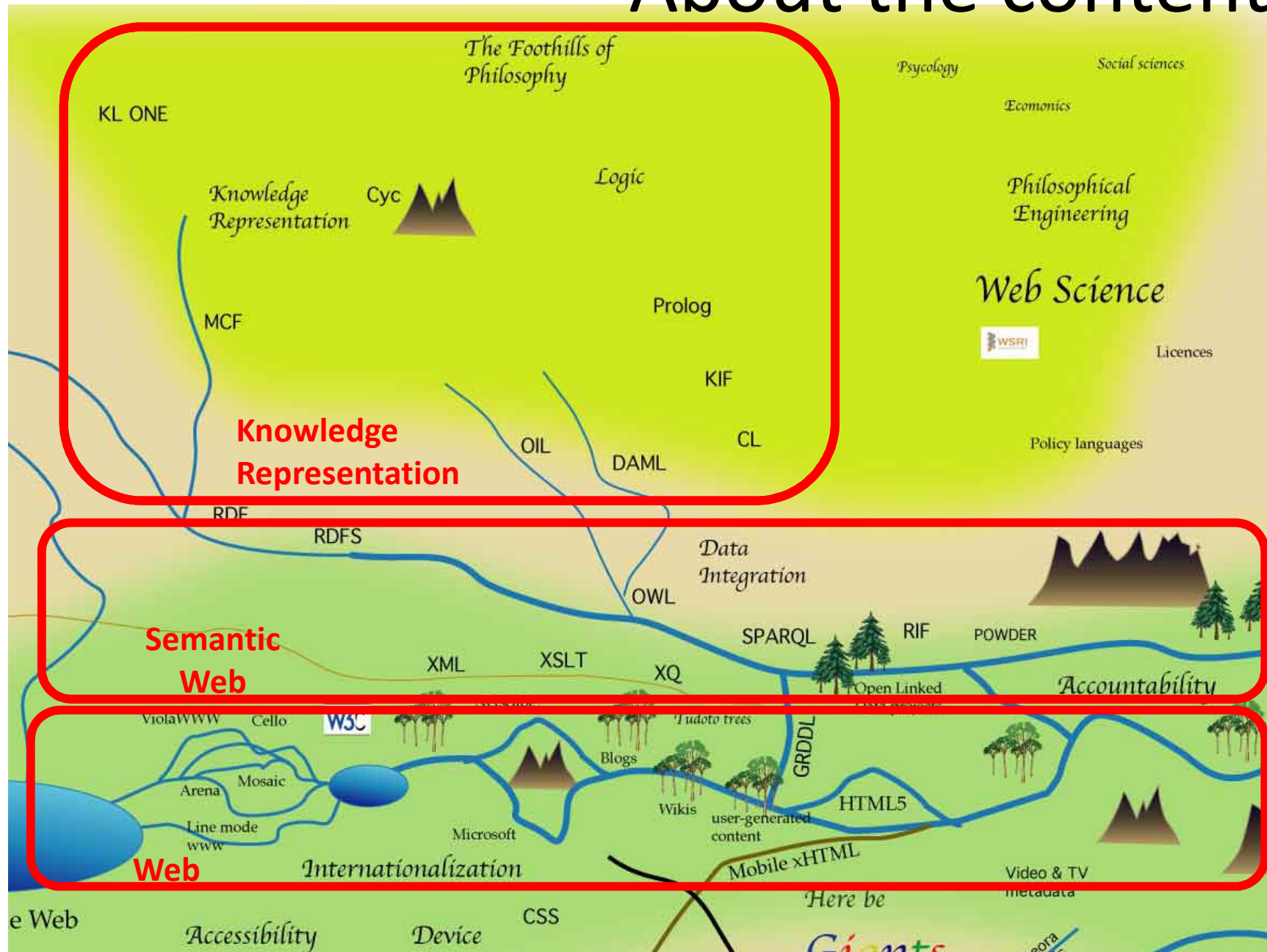
The Web Map (by Berners-Lee)



Copyright Tim Berners-Lee 2007. Some rights reserved. Creative Commons licence for non-commercial use with attribution but no derivative work about the written permission of the author. Any other use only with explicit written permission of the author. <http://www.w3.org/2007/09/map/main>

©Tim Berners-Lee, <http://www.w3.org/2007/09/map/main.jpg>

About the content



The Web 1.0 ...

- Information represented by means of:
 - Natural language
 - Images, multimedia, graphic rendering/aspect
- Human Users easily exploit all this means for:
 - Deducting facts from partial information
 - Creating mental associations (between the facts and, e.g., the images)
 - They use **different communication channels** at the same time (contemporary use of many primitive senses)

The Web 1.0 ...

- The content is published on the web with the principal aim of being “human-readable”
 - Standard HTML is focused on *how* to represent the content
 - There is no notion of *what* is represented
 - Few tags (e.g. <title>) provide an implicit semantics but ...
 - ... their content is not structured
 - ... their use is not really standardized

The Web 1.0 ...

Sembrava tutto fatto, invece approvato un solo articolo. All'origine del blocco il no di Mastella all'art.91

Finanziaria, il voto slitta a domani Al Senato va in scena l'assurdo

Dini ripete: "Mani libere fino all'ultimo". Prodi ottimista
di CLAUDIA FUSANI



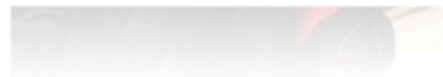
Lamberto Dini

ROMA - L' "assurdo" è andato in scena oggi nell'aula di palazzo Madama sfidando la logica, il buon senso, la corretta gestione della politica e lasciando il passo alle monovre di palazzo. Doveva essere il giorno dell'atteso e faticoso voto finale alla Finanziaria. Alle nove e mezzo del mattino, quando il presidente del Senato Franco Marini riprende i lavori mancano sette articoli - dal 91 al 97 - e circa ottanta votazioni. Poche ore ed è finita. Alle otto di sera, invece, l'articolo 91 è stato accantonato - rinviato - il 92 approvato, il 93 in piena discussione. E anche molto vivace. Si e no sono state fatte una decina di

votazioni. Un impasse assurdo, kafkiano, che racconta una delle pagine più incredibili di questa

votazioni. Un impasse assurdo, kafkiano, che racconta una delle pagine più incredibili di questa

Lamberto Dini



<h1>

<!-- inizio TITOLO -->

Finanziaria, il voto slitta a domani

Al Senato va in scena l'assurdo

<!-- fine TITOLO -->

</h1>

We can identify the title by means of its representation (<h1>,) ...

... what if tomorrow the designer changes the format of the web pages?

The Web 1.0 ...

- Web pages contain also links to other pages, but ...
 - No information on the link itself ...
 - ... what does a link represent?
 - ... what does the linked page/resource represent?
 - E.g.: in my home page there are links to other home pages ...
 - Which ones link to colleagues?
 - Which ones link to friends?

The Web 1.0 ...

Actual Web = Layout + Routing

The problem: it is not possible to
automatically reason about the data

The Web 1.0 ...

- We can see the Web as an immense *database*, every day queried by millions of users
 - Users access it through *search engines* and *keywords ...*
 - ... successful search depends on many parameters
 - the “quality” of the indexing and search algorithm
 - the number of total pages that have been indexed
 - the (meta-)content of the pages
 - E.g.: google, US election in 2005, and the keyword “stupid”

The Web 1.0 ...

- The web is *global*
 - Any page can link to anything
 - Approximatively, anyone can publish anything on the web, about any topic
 - *Distribution* of the information
 - *Inconsistency* of the information
 - *Incompleteness* of the information
 - Some recent attempts to limit such freedom (with mixed results)

Web 2.0 (ten years ago...)

- Term referring to O'Reilly Media Web 2.0 Conference, 2004 (but not coined there).
- A new way of using the web (rather than technical advances)
- Roughly (but really roughly) speaking:
 - Possibility of user of *adding/sharing content* (without being web editors)
 - Strong, unpredictable (???) *social participation* (blogs, wikis, social networking, participation, youtube, folksonomies)
 - Possibility of *net-distributing applications* (hosted services, web services, cloud computing, web-office)

Web 3.0 ? Web 4.0 ?

Ummh... Oooh...
Well... I am not
really sure...



The "Press Any Key" Dilemma

How the current web is changing?

Which is the impact of the social media?

Semantic Web

Goal: “*use*” and “*reason upon*” all the available data on the internet *automatically*

How? By *extending* the current web with *knowledge* about the content (*semantic information*)

Semantic Web

*“The Semantic Web is about **two things**. It is about **common formats for integration and combination of data** drawn from diverse sources, where on the original Web mainly concentrated on the **interchange of documents**. It is also about **language for recording how the data relates to real world objects**. That allows a person, or a machine, to start off in one database, and then move through an unending set of databases which are connected not by wires but by being about the same thing.”*

SOURCE: W3C Semantic Web Initiative

Semantic Web

Principles SW would like to preserve:

- Globality
- Information distribution
- Information inconsistency
 - Content inconsistency
 - Link inconsistency
- Information incompleteness
 - ... of contents
 - ... of routing information (links)

Adding information about the content

Adding information is not enough

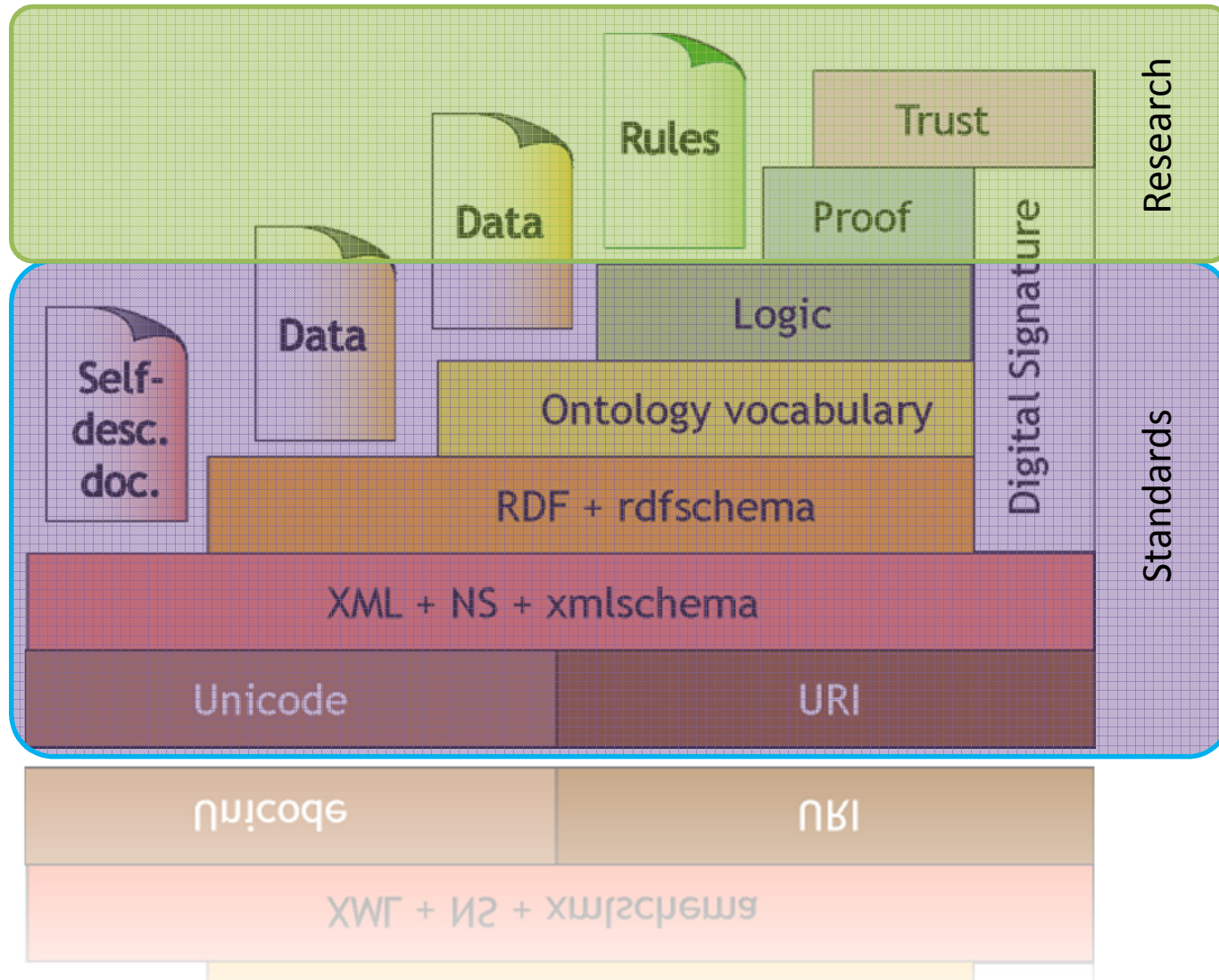
- Information should be structured (e.g., Linneo classification for the living world)
 - *Ontologies?*
- There is the need of some inference mechanism (e.g., sillogism, FOL, DL algorithm)
 - *Logic?*
- We should be able to infer new knowledge
 - We need the *proofs* that originated such new knowledge

Proof and Trust

We could exchange the proofs to ...

- ... justify new inferred knowledge
- ... overcome the definitory aspect of IT
- ... reason upon the trust...

Semantic Web Architecture



SW – Applications?

SW is cross-domain (as ICT): standards and tools have application fields in every possible domain.

To cite some:

- Search engines
- Intelligent Assistant
- Database Integration
- Digital libraries (XMP Adobe)
- Web services and cloud computing (Semantic Web Services)

Applications

Document search

- Industries (mid-size and more) needs to index and easily access/retrieve all the documentation
 - GSA - Google Search Appliance
 - (2007 prices: \$1,995 up to 50.000 docs, \$30,000 up to 500.000 docs)
 - (2015 prices: not available, 2-3 yrs contract, depend on the number of indexed documents)
 - Microsoft Sharepoint Search Services/Server

Applications

Other portals ...

- Sun's White Paper and System Handbook
- Harper's Online magazine – papers linked by means of an internal ontology
- Oracle - virtual press room
- Opera's community site
- Yahoo! Food
- FAO's Food
- Nutrition and Agriculture Journal

Applications

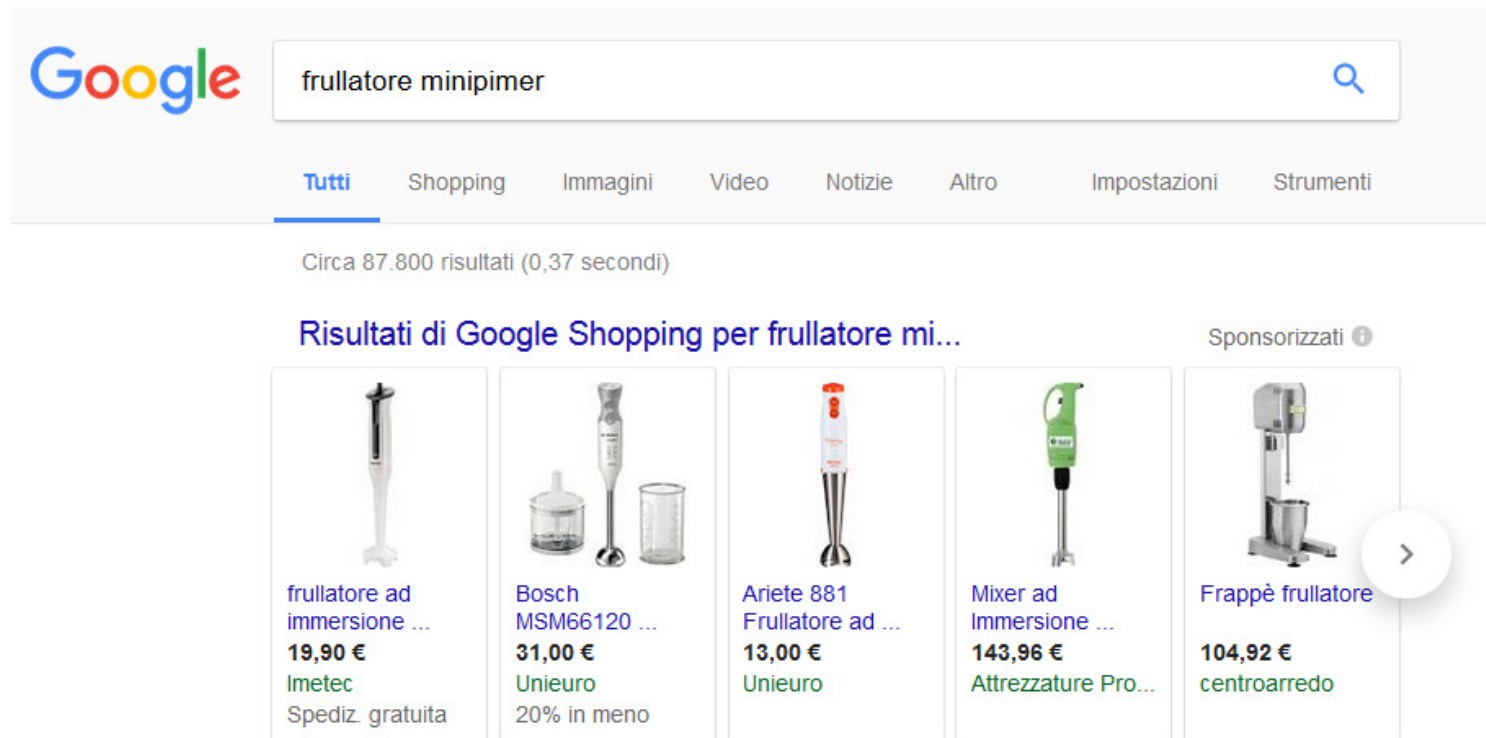
Microdata and micro formats

- Google and shopping...
- <https://support.google.com/merchants/answer/6069143?hl=en>
- schema.org (ownership: Google, Microsoft, Yahoo and Yandex)
 - Creative works: CreativeWork, Book, Movie, MusicRecording, Recipe, TVSeries ...
 - Embedded non-text objects: AudioObject, ImageObject, VideoObject
 - Event
 - Health and medical types: notes on the health and medical types under MedicalEntity.
 - Organization
 - Person
 - Place, LocalBusiness, Restaurant ...
 - Product, Offer, AggregateOffer
 - Review, AggregateRating
 - Action

Applications

Microdata and micro formats

- Google shopping example:



The image shows a Google search interface for 'frullatore minipimer'. The search bar contains the text 'frullatore minipimer' and a magnifying glass icon. Below the search bar, there are navigation tabs: 'Tutti' (selected), 'Shopping', 'Immagini', 'Video', 'Notizie', 'Altro', 'Impostazioni', and 'Strumenti'. Below the tabs, it says 'Circa 87.800 risultati (0,37 secondi)'. The main content area is titled 'Risultati di Google Shopping per frullatore mi...' and shows five product listings. Each listing includes an image of the product, the product name, the price, the brand, and a small promotional note. A 'Sponsorizzati' label is visible above the last product. A right arrow button is overlaid on the right side of the product grid.

Product Name	Price	Brand	Notes
frullatore ad immersione ...	19,90 €	Imetec	Spediz. gratuita
Bosch MSM66120 ...	31,00 €	Unieuro	20% in meno
Ariete 881 Frullatore ad ...	13,00 €	Unieuro	
Mixer ad Immersione ...	143,96 €	Attrezzature Pro...	
Frappè frullatore	104,92 €	centroarredo	

Applications

Microdata and micro formats

The screenshot shows the unieuro website interface. At the top, there's a navigation bar with the unieuro logo and a search bar. Below the navigation, there are several menu items: Punt vendita, Offerte, Faq, Unieuro club, Naviga per marca, and Guida all'acquisto. The main content area features a product image of a Bosch immersion blender, its specifications (600W power, 2 speeds), and a large orange 'ACQUISTA' button. The price is prominently displayed as €31,00 IVA INCLUSA. Below the product, there are shipping options: 'SPEDIZIONE STANDARD €9,99' and 'RITIRO IN NEGOZIO €0,00'. The bottom of the page has a dark blue banner with 'SPEDIZIONE GRATUITA sopra i 99€' and 'Domande Frequenti'.

The browser's developer tools are open, showing the HTML structure. The selected element is a `<div class="price" itemprop="price">` tag, which contains a `€ ` tag followed by the price `31` and a `,00` tag. The `itemprop="price"` attribute is highlighted in blue. The `itemscope="http://schema.org/Offerz"` attribute is also visible on the parent `<div class="prices">` tag. The CSS rules panel on the right shows the styling for the price element, including `font-size: 3.4285714286rem;` and `font-weight: 700;`.

Applications

Microdata and micro formats

The screenshot shows the unieuro website interface. At the top, there is a navigation menu with items like 'Punti vendita', 'Offerte', 'Faq', and 'Unieuro club'. A search bar contains the text 'COSA STAI CERCANDO?'. The main content area displays a Bosch immersion blender with its accessories. The price is prominently shown as €31,00 IVA INCLUSA. Below the price, there is an orange 'ACQUISTA' button. Shipping options are listed as 'SPEDIZIONE STANDARD €9,99' and 'RITIRO IN NEGOZIO €0,00'. The page footer includes 'SPEDIZIONE GRATUITA sopra i 99€' and 'Domande Frequenti'.

The screenshot shows the Chrome DevTools Inspector. The 'Inspector' tab is active, displaying the HTML structure of the page. The 'DOM' tree shows the following structure:

```
<span class="hr"></span>
<div class="actions-icon-reevoo"></div>
<div class="promotion-container-tablet"></div>
::after
</section>
<!--use logged or nologged-->
<section class="container-right">
  <article class="promo-container">
    <span class="hr"></span>
    <article class="price-container">
      <div class="table-mobile">
        <div class="prices" itemscope="">
          <div class="discounted-pr">
            <div class="price" itemprop="price">
              <span class="currency">€&nbsp;</span>
              31
              <span class="decimal">,</span>
              <span class="decimal">00</span>
            </div>
            <div class="iva">Iva incl</div>
          </div>
        </div>
      </div>
    </article>
  </section>
</div>
```

```
<div class="price" itemprop="price">
  <span class="currency">€&nbsp;</span>
  31
  <span class="decimal">,</span>
  <span class="decimal">00</span>
</div>
<div class="iva">Iva inclusa</div>
```

Semantic Information

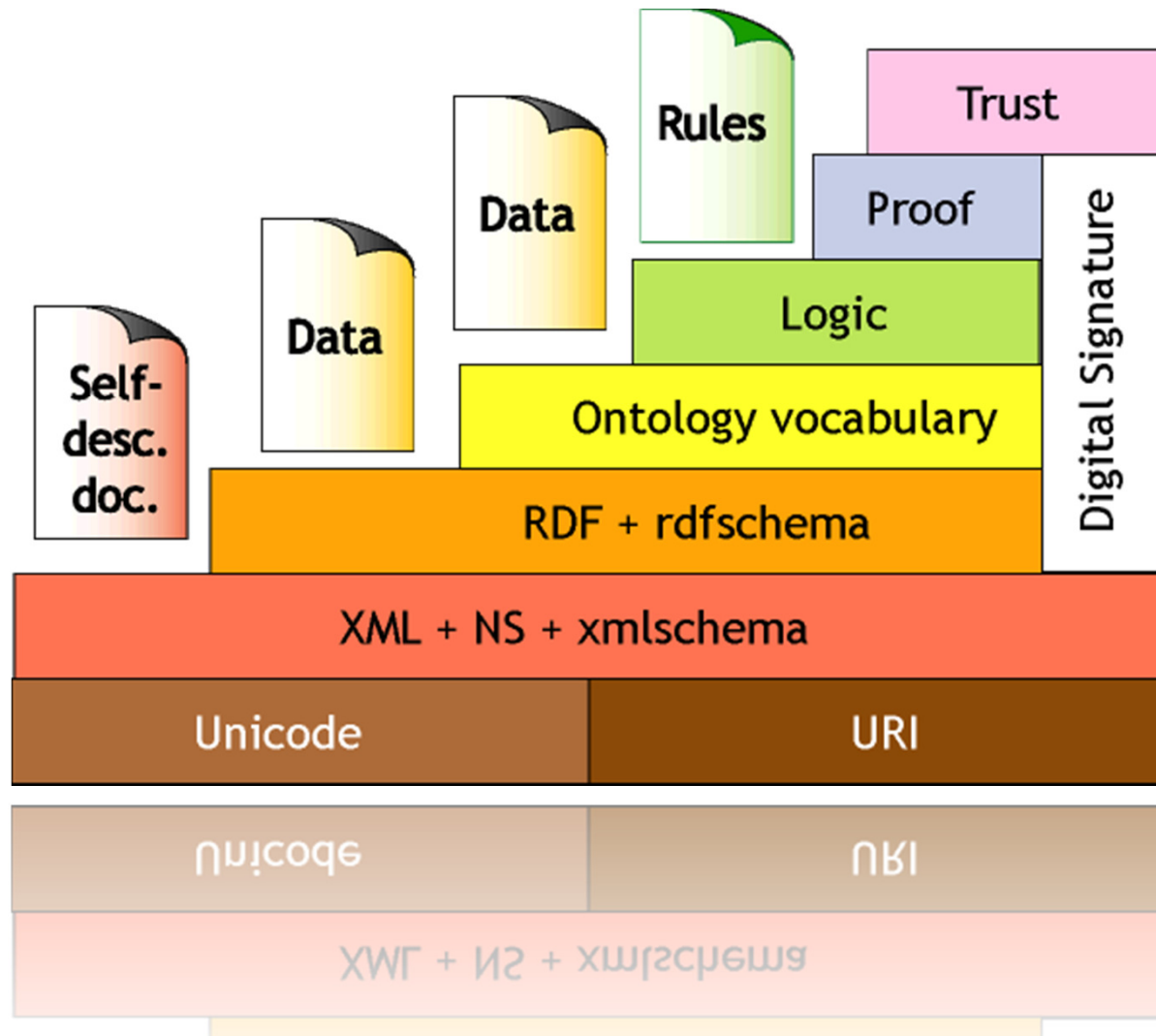
Semantic Models

How to represent semantic information?

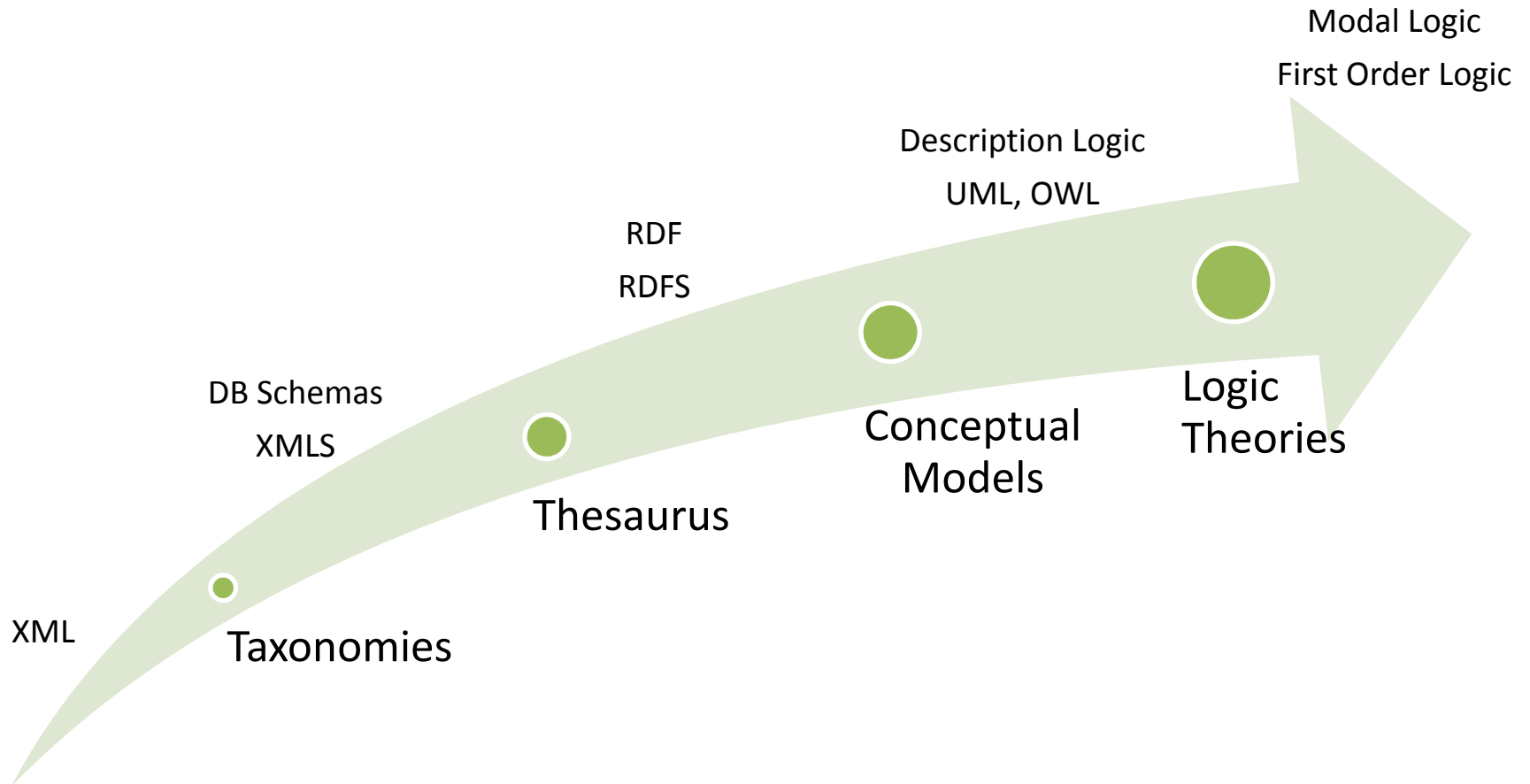
- Which language?
- Which expressivity?
- Reasoning? What about performances?

At this point, Semantic Web meets the
Knowledge Representation research field
(from AI)

Semantic Web Architecture



Semantic Models



Semantic Models

- **Taxonomy**: a set of *terms*, *hierarchically* organized
 - Allows to represent that there are relations among terms ...
 - ... but does not permit to describe the nature of such relations
 - Typically, father/child node relation
 - Search of a term is efficient only if you already know where to look for....

Semantic Models

An example of taxonomy we have to deal with: IEEE Computer Society Keywords

<http://www.computer.org/portal/web/publications/acmtaxonomy> , approximately 1766 terms hierarchically structured ...

Category: Artificial Intelligence

IV. Knowledge Representation Formalisms and Methods

- I. Agent communication languages
- II. Distributed representations
- III. Frames and scripts
- IV. Knowledge base management
- V. Knowledge base verification
- VI. Modal logic
- VII. Predicate logic
- VIII. Relation systems
- IX. Representation languages
- X. Representations (procedural and rule-based)
- XI. Semantic networks
- XII. Storage mechanisms
- XIII. Temporal logic

V. Programming Languages and Software

- I. Expert and knowledge-intensive system tools and techniques

Semantic Models

- **Thesaurus**: originally from linguistic research field, it is a set of terms together with (linguistic) relations among them:
 - Synonym
 - Hyperonyms
 - Hyponims
 - Holonyms
 - Meronyms
 - ...

They address typical problems in natural language, such as ambiguity and redundancy

- WordNet (©Princeton University),
<http://wordnet.princeton.edu/>

Semantic Model: thesaurus

WordNet 2.1 Browser

File History Options Help

Search Word: fox

Searches for fox: Noun Verb Senses: []

The noun fox has 7 senses (first 2 from tagged texts)

1. (3) **fox** -- (alert carnivorous mammal with pointed muzzle and ears; do not hunt in packs)
2. (1) dodger, **fox**, slyboots -- (a shifty deceptive person)
3. **fox** -- (the grey or reddish-brown fur of a fox)
4. **Fox**, Charles James Fox -- (English statesman who supported American Revolution (1749-1806))
5. **Fox**, George Fox -- (English religious leader who founded the Society of Friends)
6. **Fox** -- (a member of an Algonquian people formerly living west of the Great Lakes)
7. **Fox** -- (the Algonquian language of the Fox)

The verb fox has 3 senses (no senses from tagged texts)

1. flim-flam, trick, fob, **fox**, pull a fast one on, play a trick on -- (cause to believe something that is not true)
2. confuse, throw, **fox**, befuddle, fuddle, bedevil, confound, disconcert, cause to be unable to think clearly; "These questions confuse even the best of us"; "This question befuddled even the teacher"
3. **fox** -- (become discolored with, or as if with, mildew spots)

Overview of fox

File History Options Help

Search Word: fox Redisplay Overview

Searches for fox: Noun Verb Senses: []

Sense 2
dodger, **fox**, slyboots -- (a shifty deceptive person)

- => deceiver, cheat, cheater, trickster, beguiler, slicker -- (someone who leads you to believe something that is not true)
- => wrongdoer, offender -- (a person who transgresses moral or civil law)
- => bad person -- (a person who does harm to others)
- => person, individual, someone, somebody, mortal, soul -- (a human being; "there was too much for one person to do")
- => organism, being -- (a living thing that has (or can develop) the ability to act or function independently)
- => living thing, animate thing -- (a living (or once living) entity)
- => object, physical object -- (a tangible and visible entity; an entity that can cast a shadow; "it was full of rackets, balls and other objects")
- => physical entity -- (an entity that has physical existence)
- => entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
- => causal agent, cause, causal agency -- (any entity that produces an effect or is responsible for events or results)
- => physical entity -- (an entity that has physical existence)
- => entity -- (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Sense 3
fox -- (the grey or reddish-brown fur of a fox)

- => fur, pelt -- (the dressed hairy coat of a mammal)

"Hypernyms (this is a kind of...)" search for noun "fox"

Semantic Models

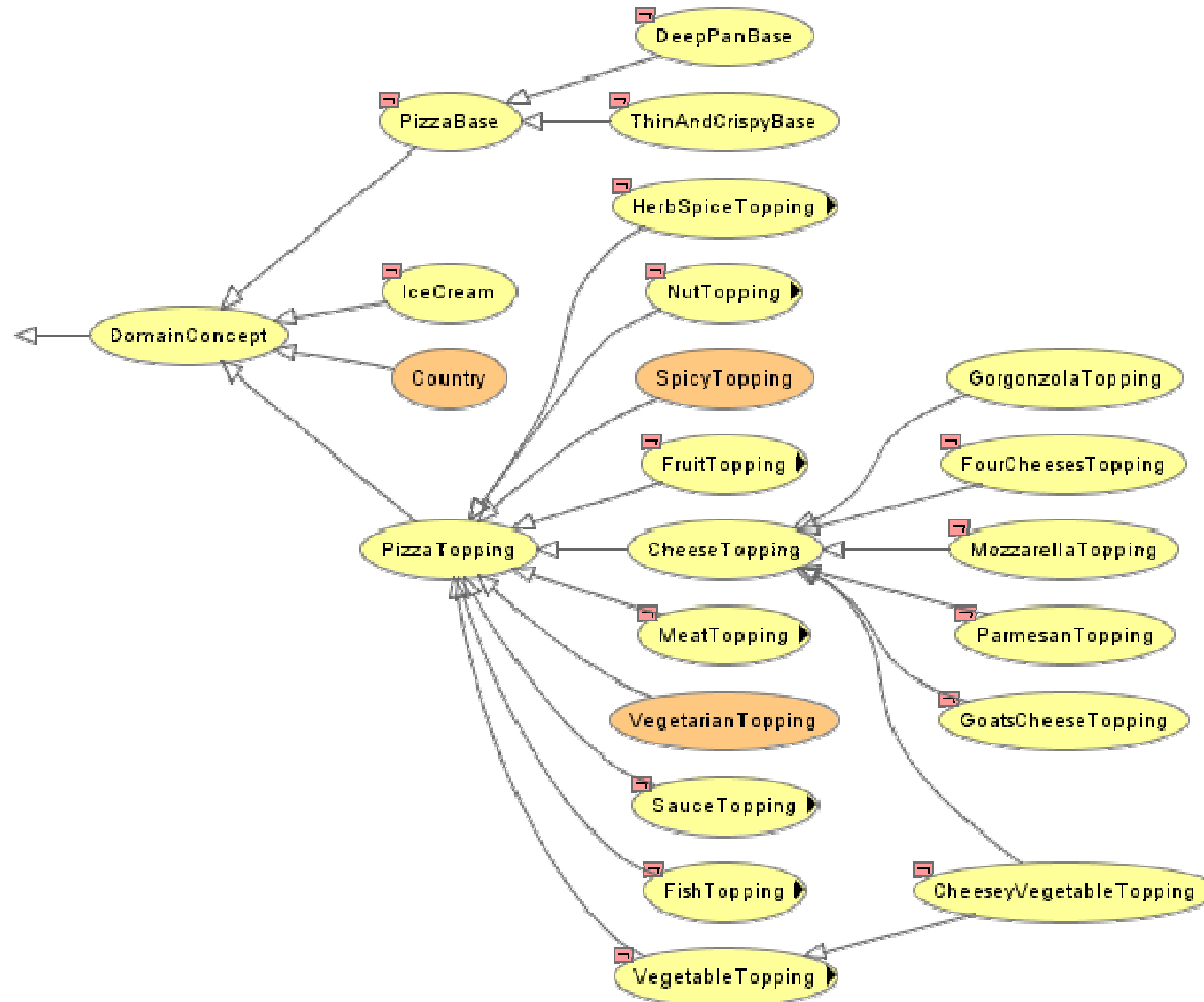
- **Conceptual models**: focused on a particular *domain area*. They specify:
 - Domain *entitites*
 - **Relations** between the entitites (properties and attributes)
 - Rules about classes, roles and relations
 - **Inference mechanisms** -> *Logic theories!!!*

Ontologies – a definition

An ontology is a **formal, explicit description** of a **domain** of interest

- Classes
- Semantic relation between classes (roles)
- Properties associated to a concept (e.g., restrictions)
- Logic (axioms, inference rules)

Ontologies – an example



Ontologies

An ontology is a **formal, explicit description** of a **domain** of interest

- They are a fundamental piece, independently of Semantic Web
- The issues are in the “subtle distinction of meaning”
- They have been a research field in AI since the beginning

Ontologies

Do we really miss them?

Crisis of dotCom market (2001)

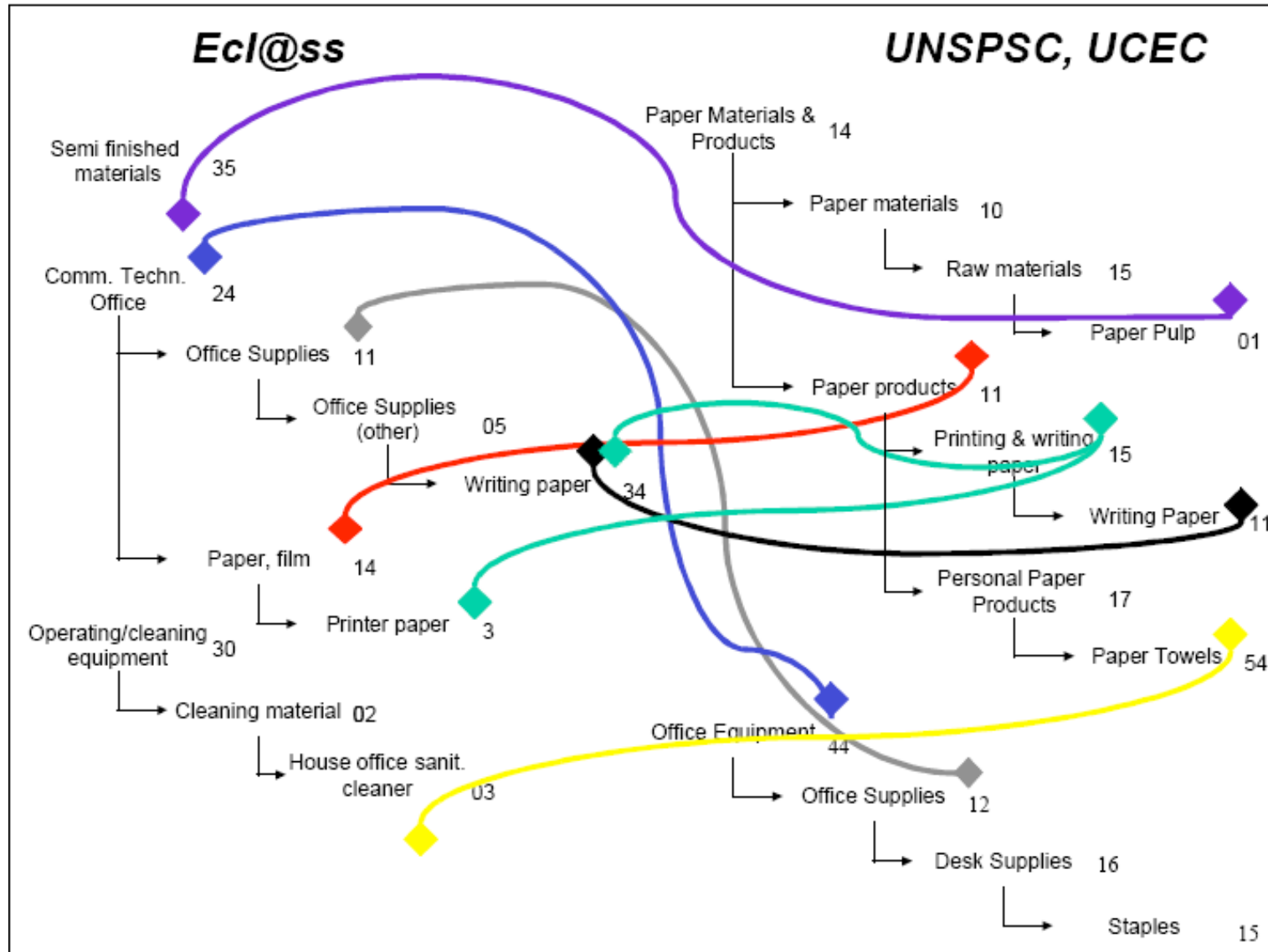
Harvard Business Review, October 2001:

“Trying to engage with too many partners too fast is one of the main reasons that ***so many online market makers have foundered.***

The transactions they had viewed as simple and routine actually involved many ***subtle distinctions in terminology and meaning***

Ontologies

Do we really miss them?



Ontologies

XML is not enough?

“XML is only the first step to ensuring that computers can communicate freely. *XML is an alphabet for computers* and as everyone who travels in Europe knows, knowing the alphabet doesn't mean you can speak Italian or French”

Business Week, March 18, 2002

Why ontologies?

- An ontology provides a structured model of a (business) domain
 - Solves term ambiguity
 - Clarifies/simplifies domain peculiarities
 - As a consequence, deep analysis and understanding of a (business) domain ...
 - ... high competitive advantage !

Ontologies

(now vocabularies in the W3C terminology)

Few examples:

- Dublin Core, focussed on documents
- WordNet (better example, BabelNet)
- Gene Ontology, genomic
- Protein Ontology, proteomics
- SnoMed, a very important ontology in the medical field

Linked Data

Ontologies comprise:

- the terms (the concepts, aka TBox)
- the instances (effective data, aka ABox)

Great emphasis in the last years for linking (open) data (LOD initiative) using SW tools

Example (and starting point): <http://dbpedia.org/>

Linked Data Principles

(Tim Berners-Lee, 2006)

- On the web
- Machine-readable
- Non-proprietary format
- RDF standards
- Linked RDF

Ontologies

From the formal viewpoint, based on Description Logics

- A family of logics
- A number of different logic operators
- "Recipes": depending on which logic operators are chosen, a different logic is obtained
- Up to almost 10.000 different logics...
- Complexity: <http://www.cs.man.ac.uk/~ezolin/dl/>

Ontologies

Key idea:

- Classes or concepts
- Individuals
- Relations

- A-Box and T-Box

- Subsumption as the main reasoning mechanism

Ontologies and OWA

- Differently by many other formalisms, Description Logics are based on an Open World Assumption.
- If a sentence cannot be inferred, then its truthness value is unknown.
- Note: somehow this characteristics is linked to the idea of distributed information in the Web.

Ontologies and OWA

- Oedipus killed his father...
- ... married his mother Iokaste ...
- ... had children with her, among them Polyneikes.

- Finally, also Polyneikes had children, among them Thersandros.

Ontologies and OWA

- Oedipus killed his father...
- ... married his mother Iokaste ...
- ... had children with her, among them Polyneikes.
- Finally, also Polyneikes had children, among them Thersandros.

We want to know. . .

. . . if Iokaste has a child that is a patricide and that itself has a child that is not a patricide. . .

Ontologies and OWA

Hypothesis 1: the child we are looking for is Oedipus

Indeed, we know Oedipus is a patricide. . .
We know also he has a son, Polyneikes,
but we do not know if Polyneikes is a
patricide or not. . .

Ontologies and OWA

Hypothesis 2: the child we are looking for is Polyneikes. . .

Indeed, we know Polyneikes has a son, Thersandros, and we know Thersandros is not a patricide.

. . . but we do not know if Polyneikes is a patricide or not. . .

Ontologies and OWA

But indeed...

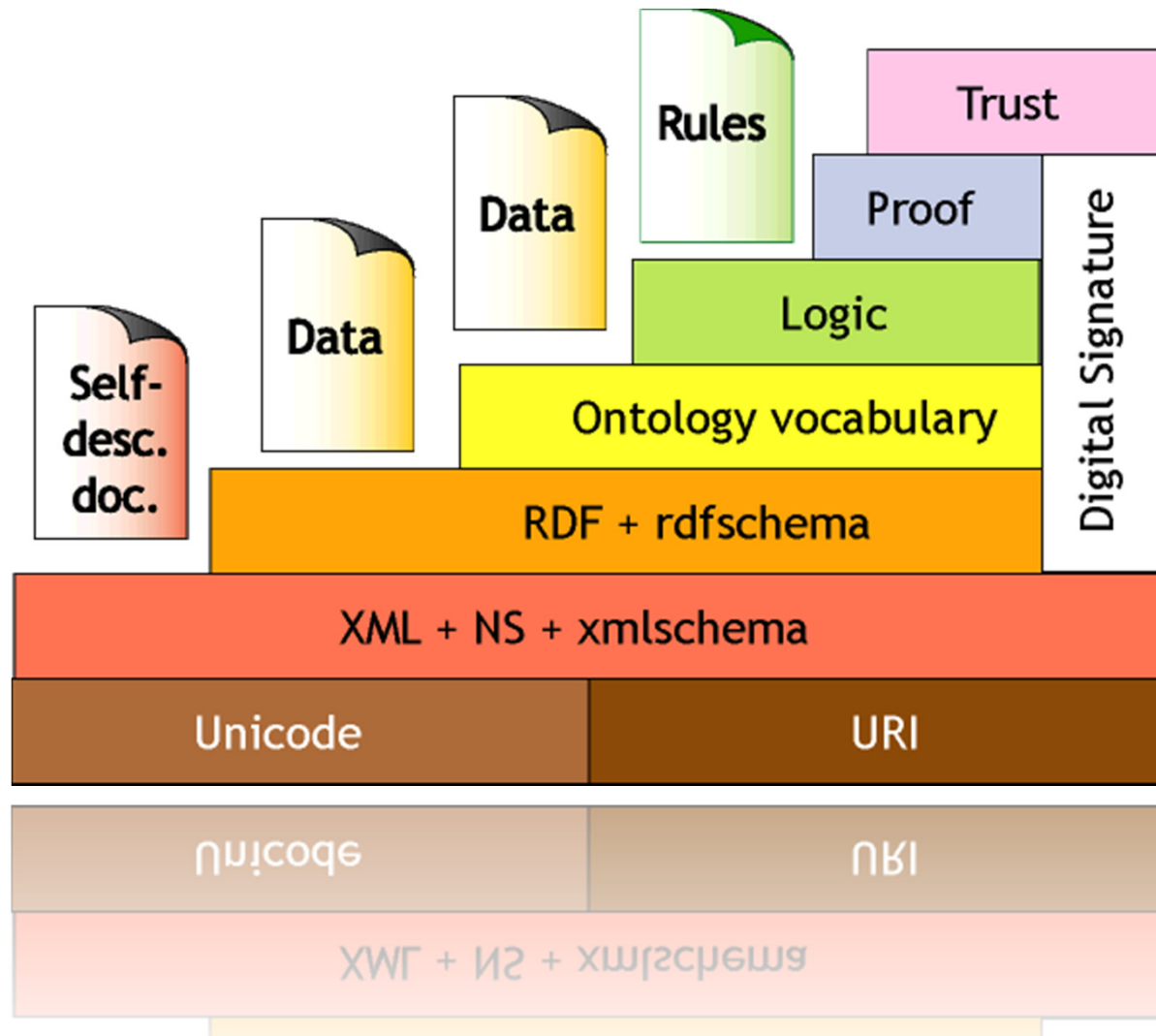
In all possible models, either Polyneikes is a patricide, or he is not. All possible models can be split up in two classes: one in which Polyneikes is a patricide, one where he is not a patricide.

In the first class, the child (we are looking for) is Polyneikes.
In the second class, the child (we are looking for) is Oedipus.

In ALL models Iokaste has a child that is a patricide and that itself has a child that is not a patricide. The answer to the previous question is YES.

Semantic Web Tools

Recalling the Semantic Web Cake



A unique way for identifying concepts

- How to uniquely identified concepts?
 - > by means of a name system ...
- SW exploits an already available name systems, URIs (*Uniform Resource Identifier*)
 - By definition, URI guarantees unicity of the names
 - To each URI corresponds *one and only one* concept ...
 - ... but more URI can refer to the *same* concept!
 - **NOTE: differently from the web, it is not necessary that to each URI corresponds some content!**

Examples:

<http://www.repubblica.it>

federico.chesani@unibo.it

ISBN 88-7750-483-8

eXtensible Markup Language - XML

- Created for supporting data exchange between heterogeneous systems (hardware and software)
 - No presentation information
 - Human readable and machine readable
- Extensible, so that anyone can represent any type of data
- Hierarchically structured by means of *tags*
- An XML document can contain, in a preamble, a description of the grammar used in such document (optional) (self-describing document!!!)
- Very mature technology!

Resource Description Framework (RDF/RDFS)

- Standard W3C
- XML-based language for representing “knowledge”
- A design criteria: provide a “minimalist” tool
- Based on the concept of triple:

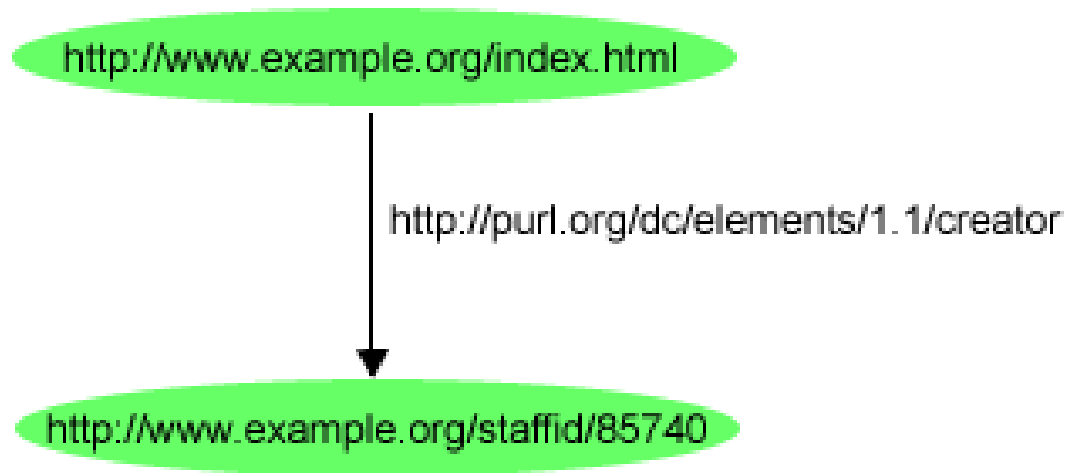
< subject, predicate, object >

< resource, attribute, value >

- Some different representations (N3, Graph, RDF/XML)

RDF – Graph Representation

- A node for the subject
- A node for the object
- A labeled arc for the predicate



`http://www.example.org/index.html` has a creator whose value is `John Smith`

RDF – Graph Representation



RDF – XML Representation

```
<rdf:RDF
```

```
  xmlns:rdf=http://www.w3.org/1999/02/22-rdf-syntax-ns#
```

```
  xmlns:contact=http://www.w3.org/2000/10/swap/pim/contact#
```

```
>
```

```
  <contact:Person   rdf:about="http://www.w3.org/People/EM/contact#me">
```

```
    <contact:fullName>Eric Miller</contact:fullName>
```

```
    <contact:mailbox rdf:resource="mailto:em@w3.org"/>
```

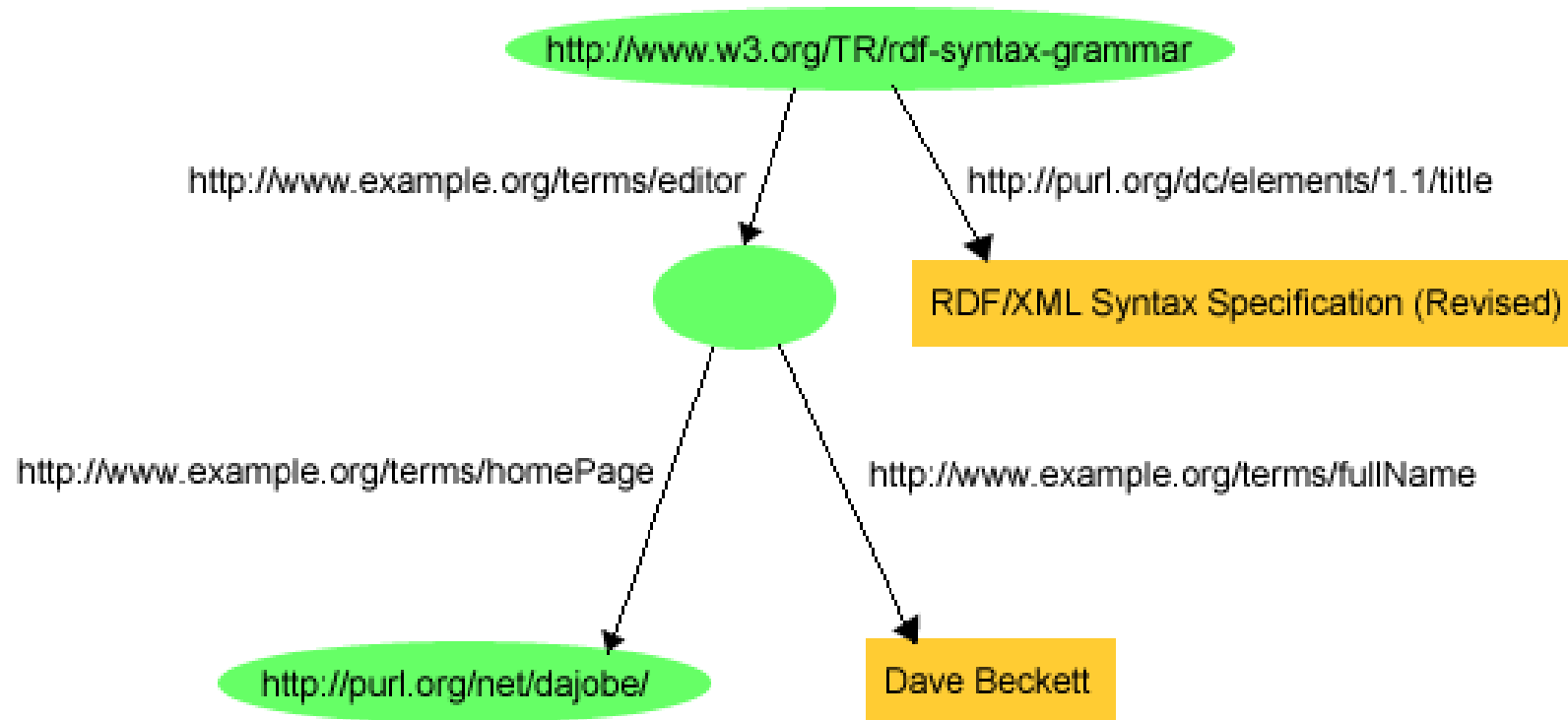
```
    <contact:personalTitle>Dr.</contact:personalTitle>
```

```
  </contact:Person>
```

```
</rdf:RDF>
```

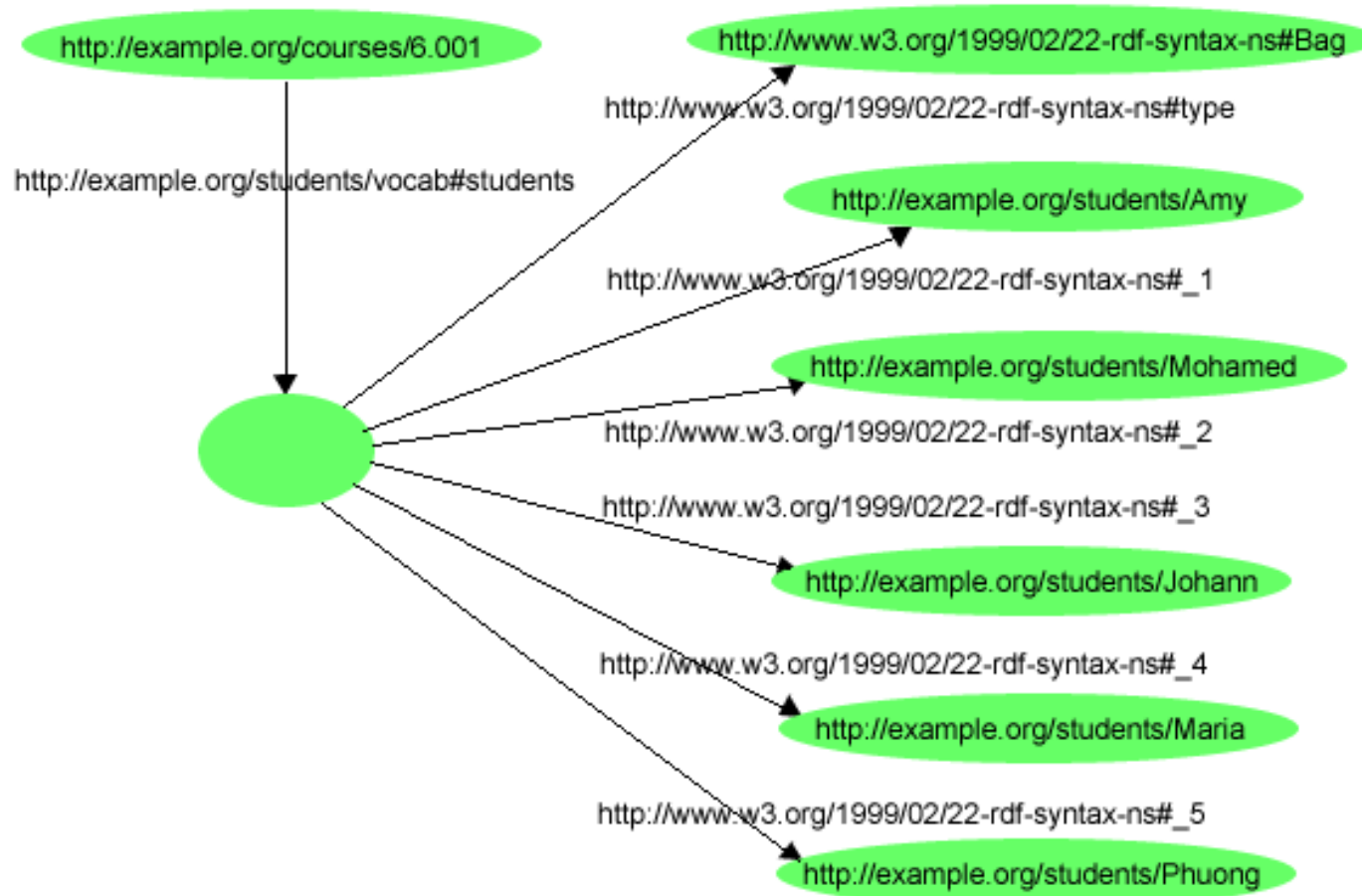
I can query for the mailbox of Eric Miller, without knowing a priori if he uses mailbox or email ...
... if Eric Miller will change mailbox, search result will be coherent!

RDF - Examples



Empty Nodes

RDF – Examples



Bags/Sets

RDF – Expressive Power

RDF supports:

- **Types** (classes) by means of the attribute **type** (that assume as value an URI)
- Subject/object of a sentence can be also **collections** (bag, sequence, alternative)
- **Meta-sentences**, through *reification* of teh sentences (“Marco says that Federico is the author of web page xy”)

RDF Schema

- RDF can be intended also as a description of resource attributes and of the values of such attributes
- RDFS allows to describe classes and relations with other classes/resources
 - *type*
 - *subClassOf*
 - *subPropertyOf*
 - *range*
 - *domain*

RDF and E/R Models

- Many similarities with E/R models ...
 - ... RDF is more expressive
- RDF to be intended as the “E/R” for the web
- Relations in RDF are “first class entities”
- In RDF the list of properties of an entity is not:
 - A priori determined by the developer
 - Centralized (DB)
 - Consequence of the fact that any one can assert anything about any one else

RDF and Relational Databases

There is a direct mapping with relational db

- A record is viewed as a RDF node
- The name of a table column is viewed as `rdf:propertyType`
- The corresponding field value is intended as the value of the property
- RDF aims to integrate different databases with different underlying model
 - Traditional DBMS are optimized for creating new data models within the same db or within a restricted set of dbs

RDF Tools

Many tools already available ...

Only in the W3C wiki there are citations for:

- 38 Frameworks/reasoners
- 47 RDF Triple Stores

Have a look to

<http://www.w3.org/2001/sw/wiki/Tools>

RDFa

- RDFa is a specification for attributes to express structured data in XHTML.
- The rendered, hypertext content of XHTML is reused by the RDFa markup
 - publishers don't need to repeat significant data in the document.

Source: RDFa Primer

<http://www.w3.org/TR/2008/NOTE-xhtml-rdfa-primer-20081014/>

RDFa

```
...  
All content on this site is licensed under  
<a href="http://creativecommons.org/licenses/by/3.0/">  
    a Creative Commons License  
</a>.
```

```
...  
All content on this site is licensed under  
<a rel="license" href="http://creativecommons.org/licenses/by/3.0/">  
    a Creative Commons License  
</a>.
```

This page has a **relation** of type **license** with the page at creative commons...

Source: RDFa Primer

<http://www.w3.org/TR/2008/NOTE-xhtml-rdfa-primer-20081014/>

RDFa

```
...  
<div>  
    <h2> The trouble with Bob </h2>  
    <h3> Alice </h3>  
    ...  
</div>
```

```
<div xmlns:dc="http://purl.org/dc/elements/1.1/">  
    <h2 property="dc:title"> The trouble with Bob </h2>  
    <h3 property="dc:creator"> Alice </h3>  
    ...  
</div>
```

Note the reference to the DC namespace, i.e. the Dublin Core initiative
<http://dublincore.org/>

SPARQL

- SPARQL can be used to express queries across diverse data sources, whether the data is stored natively as [RDF](#) or viewed as [RDF](#) via middleware.
- SPARQL contains capabilities for querying required and optional graph patterns along with their conjunctions and disjunctions.
- Supports extensible value testing and constraining queries by source [RDF](#) graph.
- The results of SPARQL queries can be results sets or [RDF](#) graphs.

Source: SPARQL W3C Working group

<http://www.w3.org/2001/sw/wiki/SPARQL>

<http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/>

SPARQL

Data:

```
<http://example.org/book/book1>  
  <http://purl.org/dc/elements/1.1/title>  
  "SPARQL Tutorial" .
```

Query:

```
SELECT ?title  
  WHERE { <http://example.org/book/book1>  
         <http://purl.org/dc/elements/1.1/title>  
         ?title .  
  }
```

Source: SPARQL W3C Working group

<http://www.w3.org/2001/sw/wiki/SPARQL>

<http://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/>

Ontology Web Language (OWL 1.0)

- Standard W3C
- Based upon/extend RDF/RDFS
- Formal Semantics (*Description Logic Fragments*)
- Three level of expressivity/complexity
 - OWL Lite
 - OWL DL
 - OWL Full
- OWL 2.0 è già standard affermato

OWL – Features

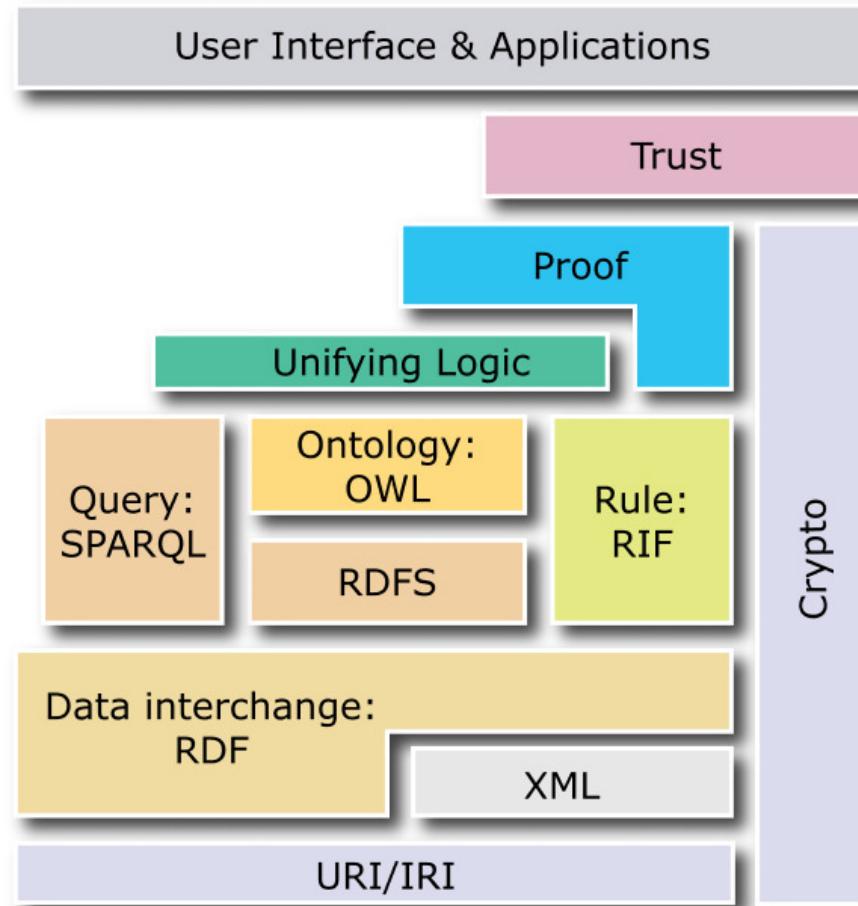
- **Classes (categories)**: subClassOf, intersectionOf, unionOf, complementOf, enumeration, equivalence, disjoint
- **Properties (Roles, Relations)**: symmetric, transitive, functional, inverse Functional, range, domain, subPropertyOf, inverseOf, equivalentProperty
- **Instances (Individuals)**: sameIndividualAs, differentFrom, allDifferent

OWL Tools

- Many tools for OWL
 - Editors (37 listed at <http://www.w3.org/2001/sw/wiki/Category:Editor>)
 - Reasoners (39 listed at <http://www.w3.org/2001/sw/wiki/Category:Reasoner>)
- Quite often integrated in a comprehensive framework

A well known (but not necessarily the best one) ontology editor:
Protégé <http://protege.stanford.edu/>

The Semantic Web Cake



Semantic Web: where are we?

Semantic Web – which problems?

- SW has been officially proposed in 2001 ...
- ... it has not *transformed* the web (yet!)
 - Are we really sure something is not changing behind?
 - A lot of research about in the academic world
- Roughly speaking, “it is difficult to understand the benefits”

Semantic Web – which problems?

- RDF adoption
 - Adding semantic content is expensive
 - Until a critical mass of semantic content is available on the web ... SW tools fail to convince.
 - W3C answer: many proposals in such directions
 - *Gleaning Resource Descriptions from Dialects of Languages* – GRDDL
 - RDFa with HTML5
 - and LINKED DATA!!!!!!!!!!
- Ontologies
 - To produce a new one is highly expensive and time-demanding
 - An ontology is “alive”, it changes in time
 - Updating costs
 - Managing costs

Semantic Web – critics?

- It cannot be done practically ... ??????
 - Metacrap problem
 - Wrong content (introduce with some bad purpose, see the wikipedia experience)
- Which use of the data?
 - Censorship problems & freedom
 - Privacy problems
- Data are already available on the web, it is sufficient to extract them
 - SW is not useful ???? But how to extract, and then represent data?
 - Mashups show some interesting results

Semantic Web – critics?

- Computationally expensive
 - ... but the adoption of a fragment of Description Logic is an answer
 - ... maybe we don't need in every application all the expressive power...

Semantic Web – but...

- A number of different things are happening anyway...
 - Linked Open Data <http://linkeddata.org/>
 - Microdata and micro formats (schema.org by Google and Microsoft)
 - Public bodies
 - <http://dati.senato.it/> (sparql endpoint and data batches, cc3.0)
 - <http://dati.comune.bologna.it/>

Concluding...

- Semantic Web: adding semantic information to web resources (data and whatever)
- Big perspectives ...
- ... we start seeing the results after 9 years, but no revolution has been really achieved yet

Few considerations (personal) ...

- Who is responsible to add semantic content on the web?
 - Single users (authors)
 - Metacrap ☹️
 - Folksonomies (e.g. flickr) 😊
 - Wikis 😊
 - Industry firms
 - They already have huge data collections, more or less organized ...
 - ... why they should share their knowledge?
 - Global market extremely hard and difficult... Knowledge is a key to competitive advantage in the Porter chain...

Few considerations (personal) ...

- Cultural issues ...
 - Not in all country there is such a desire of sharing information
- SW really appealing in a intra-business scenario
- Few doubts in a inter-business scenario
 - *Bussiness secrecy, NDA, and other commercial practices*
 - Usually, a firm wants to have a complete control over its data
 - Which data to publish?
 - Who is using them (competitors?)
 - What are they doing with our data?

Few links to start with...

Official site W3C:

- <http://www.w3.org/standards/semanticweb/>

Communities:

- <http://www.semanticweb.org/>

Thanks for the attention

Questions?

Ing. Federico Chesani, Phd

c/o DISI- Facoltà di Ingegneria

Viale Risorgimento 2

40136 - Bologna

federico.chesani@unibo.it

Tel. 051 20 93086