

Semantic Web

Ing. Federico Chesani

Corso di Fondamenti di Intelligenza Artificiale M

a.a. 2009/2010

7 Maggio 2010

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
- d) Few examples of Semantic Web applications

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

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

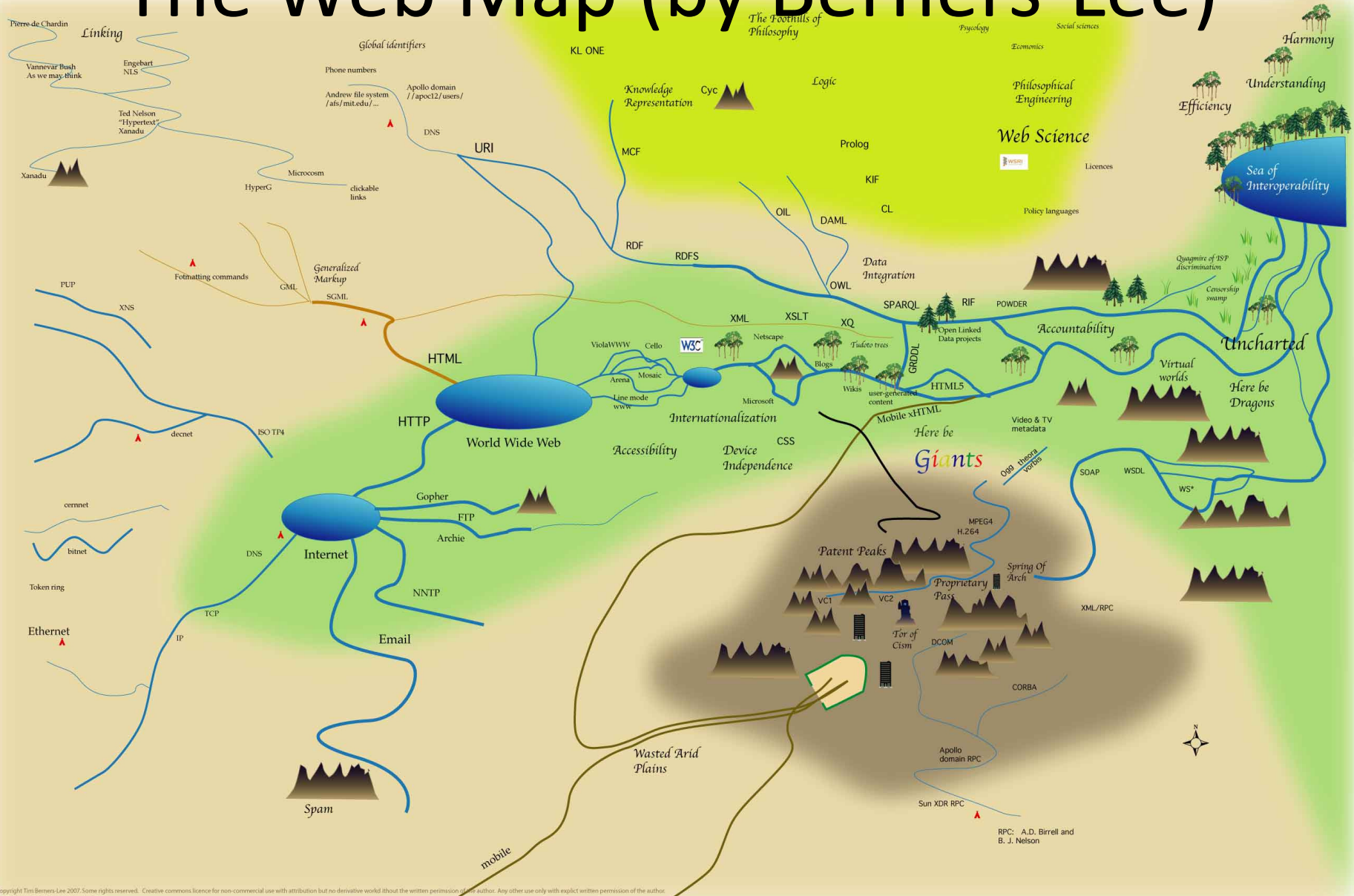
3. Semantic Web Tools

- 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

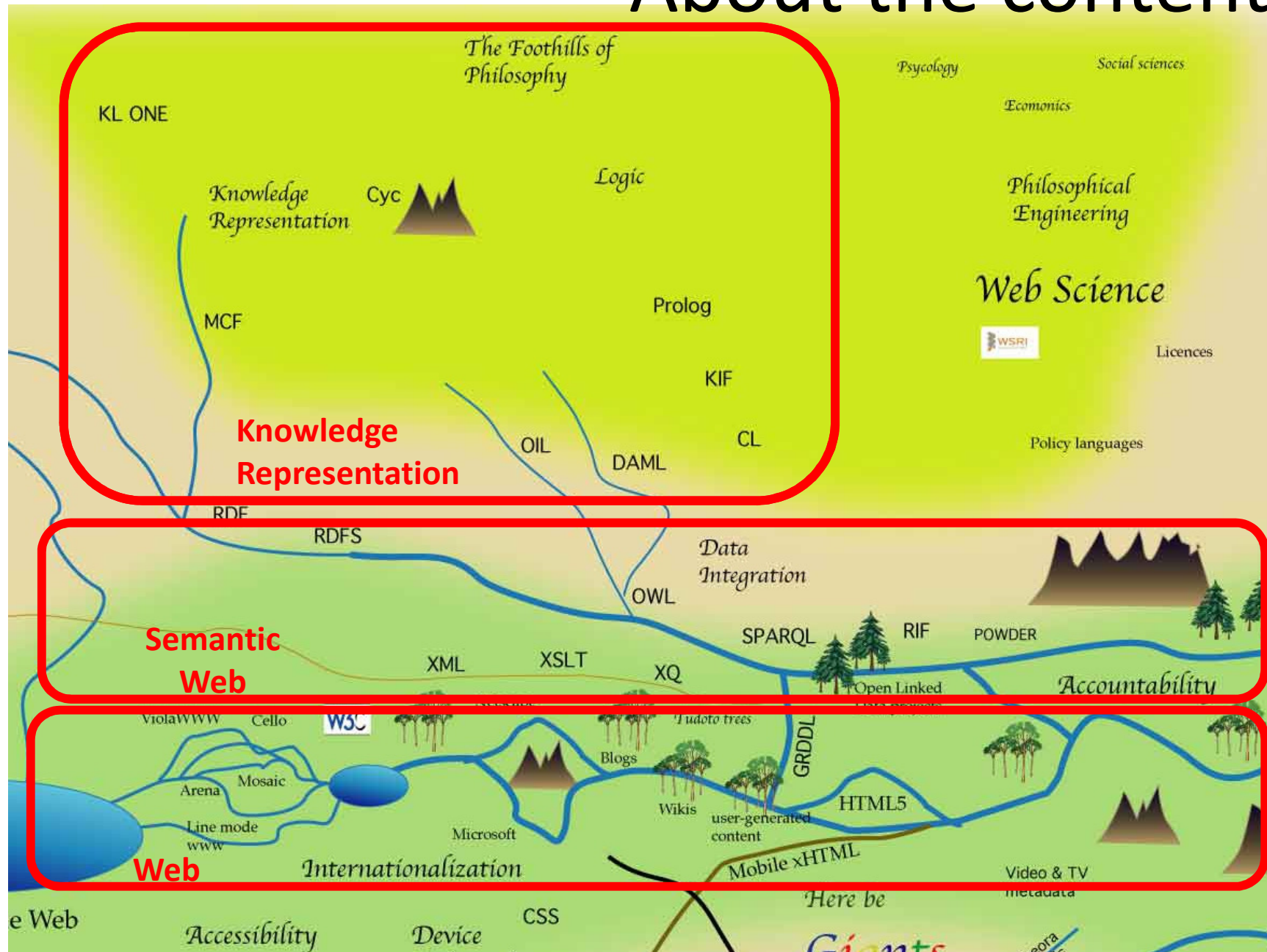
The Web Map (by Berners-Lee)



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 ...

- 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)

And the Web 2.0 ?

- Term referring to O'Reilly Media Web 2.0 Conference, 2004 (but no 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)

And the Web 3.0 ?

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



The "Press Any Key" Dilemma



And recently I heard
also of Web 4.0...

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

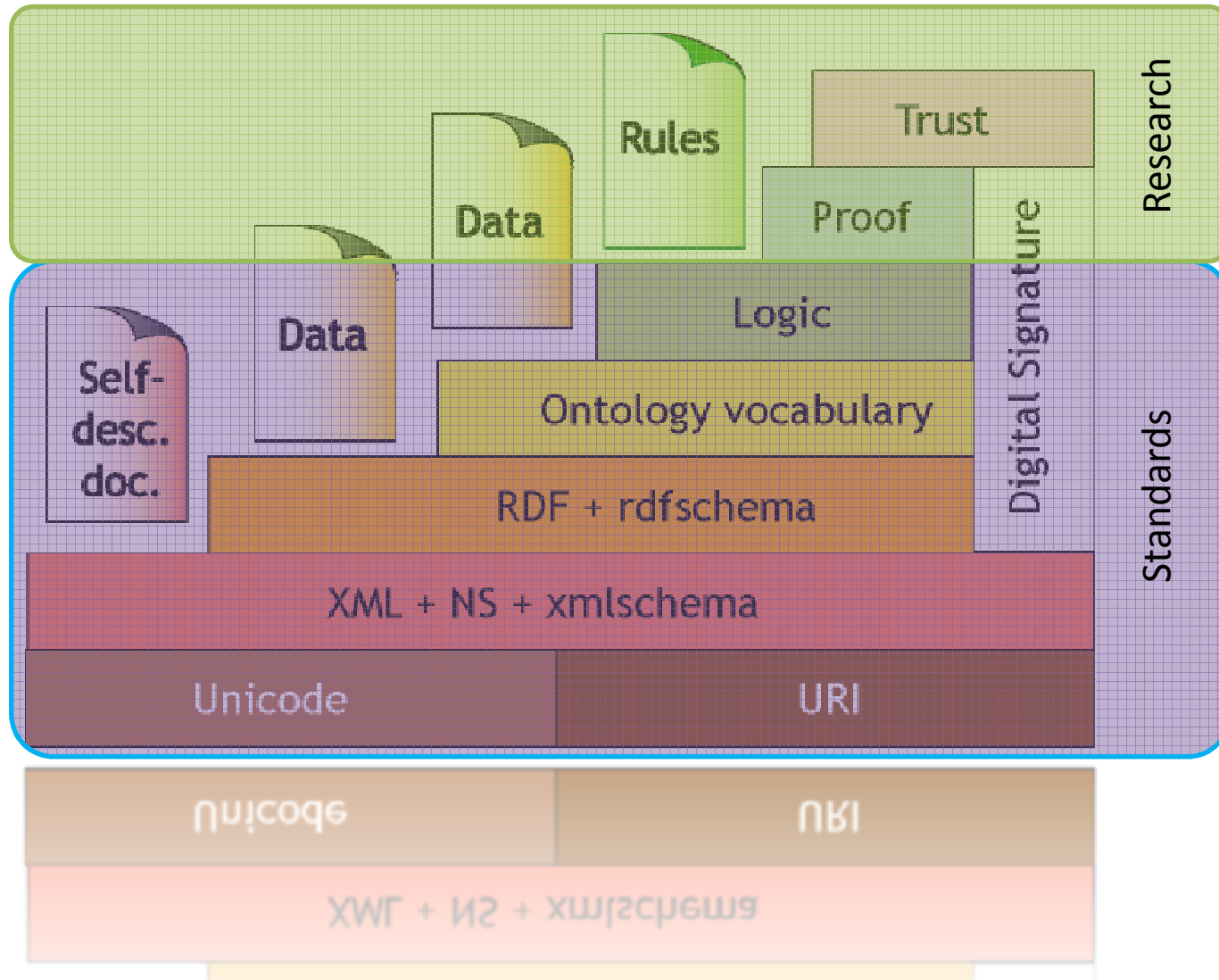
- Information should be structured (e.g., Linneo classification for the living world)
 - *Ontologies?*
- There is the need of some inference mechanism (e.g., sillogism, FOL, DL algorithm)
 - *Logic?*
- We should be able to infer new knowledge
 - We need the *proofs* that originated such new knowledge

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)
 - (2010 prices: not available, 2-3 yrs contract, depend on the number of indexed documents)
 - Microsoft Sharepoint Search Services/Server
- Vodafone Live! Mobile Portal based on RDF (SW)
 - For each download, 50% less pages accessed
 - 20% increment of downloaded stuff in 2 months (source: Ivan Herman, SW lead)

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

Intelligent Assistant

- Original Scenario proposed by Tim Berners-Lee
- Two users, by means of a “Semantic Web Agent” (running on your phone):
 - Synchronize the agenda
 - Generate and agree upon a plan
 - Delegate tasks each other
 - Exploit business contracts to support cost-based decisions

Applications

Intelligent Assistant - revisited

- We can access on-line to:
 - Bank account
 - agenda
 - Photo album
- ... But there is no real integration:
 - We can't merge the agenda with the payments enrolled a certain day...
- Less reasoning and more data integration!!!

Applications

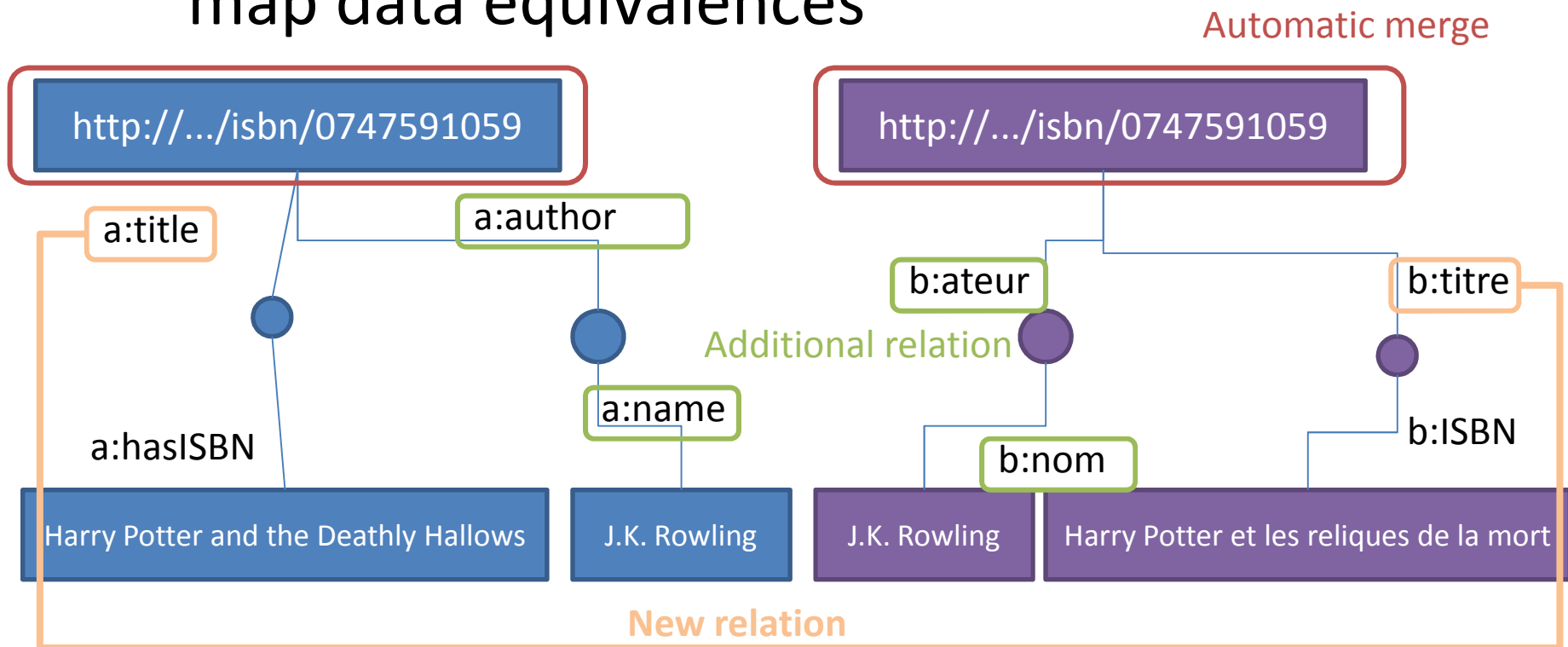
Data Integration

- Frequent need to integrate several different databases
- Roughly:
 1. Define new, more abstract data structures, to capture the data heterogeneity
 2. Merge of such abstract representations
 - Real Merge vs. Virtual merge
 3. Finally, more complex and more expressive queries

Applications

Data Integration

- The merge is successful if the data abstraction process correctly identifies and map data equivalences



Applications

Data Integration

- SW intrinsically supports such data abstraction process
- Differences:
 - The entire web is the background
 - Data are considered as distributed
 - Use of ontologies (more expressivity of the E/R model)
- Tools:
 - RDF
 - GRDDL
 - **Open Data Link Initiative**

Applications

Libraries and Digital Libraries

- Indexing happens by means of categories
 - We can access a library by means of its content/topics catalog
- This is not true with the current web
 - E.g.: looking for something related to the concept of “Artificial Intelligence”
 - ... we get as a result a list of resources containing the keywords...
 - ... no guarantee about such resources are indeed about AI...

 - Another example: suppose you want to download a disney movie for your children...
 - ... you have many chances to download an adult movie without knowing it ...

Applications

Libraries and Digital Libraries

- Digital Libraries are evolving quickly
 - Many attempts at the EU level to standardize digital content search and access (e.g. [EUROPEANA](http://www.europeana.eu/portal/), <http://www.europeana.eu/portal/>)
- New, complete different models to organize the knowledge
 - The [FRBRoo](#) ontology proposal, already supported at the national level in some EU countries

Applications

Semantic Web Services

"A Web service is a software system identified by a *URI*, whose *public interfaces* and *bindings* are defined and described using *XML*. Its definition can be *discovered* by other software systems. These systems may then *interact* with the Web service in a manner prescribed by its definition, using *XML based messages* conveyed by *internet protocols*."

("Web Services Architecture"

<http://www.w3c.org/TR/2002/WD-ws-arch-20021114>)

Applications

Semantic Web Services

- Through UDDI and WSDL it is possible to dynamically retrieve **binding** (location) and **interface** of a service ...
- ... but no way of knowing **what** such service does ...
- Es: invoke a service called `sum`, accepting two `integers` as input parameters, and an `integer` as output parameter. **How do we know that such service calculate the sum?**

Applications

Semantic Web Services

We need two extra information:

- Semantic description of the functionality offered by the service
 - E.g.: **preconditions**, **input**, **output** and **effects**
- Rules about:
 - How to invoke the service
 - Which constraints about the data

Applications

Semantic Web Services

A lot of interest about them, from both industrial and academic fields

- Two contests in the academic arena

- IEEE Contest

- <http://insel.flp.cs.tu-berlin.de/wsc06/>

- SWS Challenge

- http://sws-challenge.org/wiki/index.php/Main_Page

SWS tools

- Several proposals, no winner yet ...
 - Semantic Annotations WSDL (SAWSDL) (W3C)
 - Semantic Web Service Language (SWSL) (W3C)
 - OWL Web Services Ontology (OWL-S)
 - Supported in US by DAML (DARPA Agent Markup Language)
 - Web Service Modeling Ontology (WSMO)
 - Supported in EU by DERI (Digital Enterprise Research Institute: Galway, Seoul, Stanford)

Semantic Information

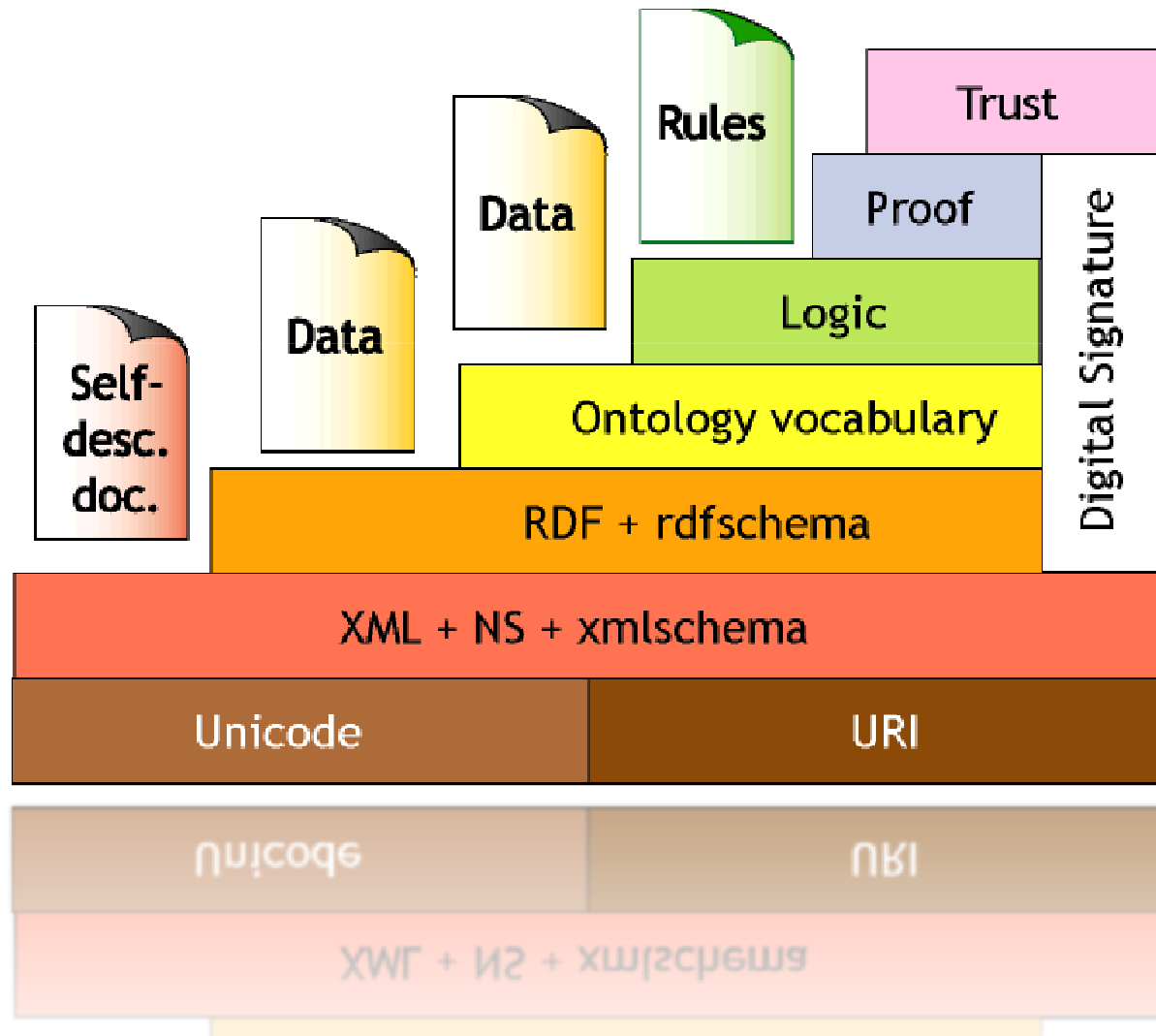
Semantic Models

How to represent semantic information?

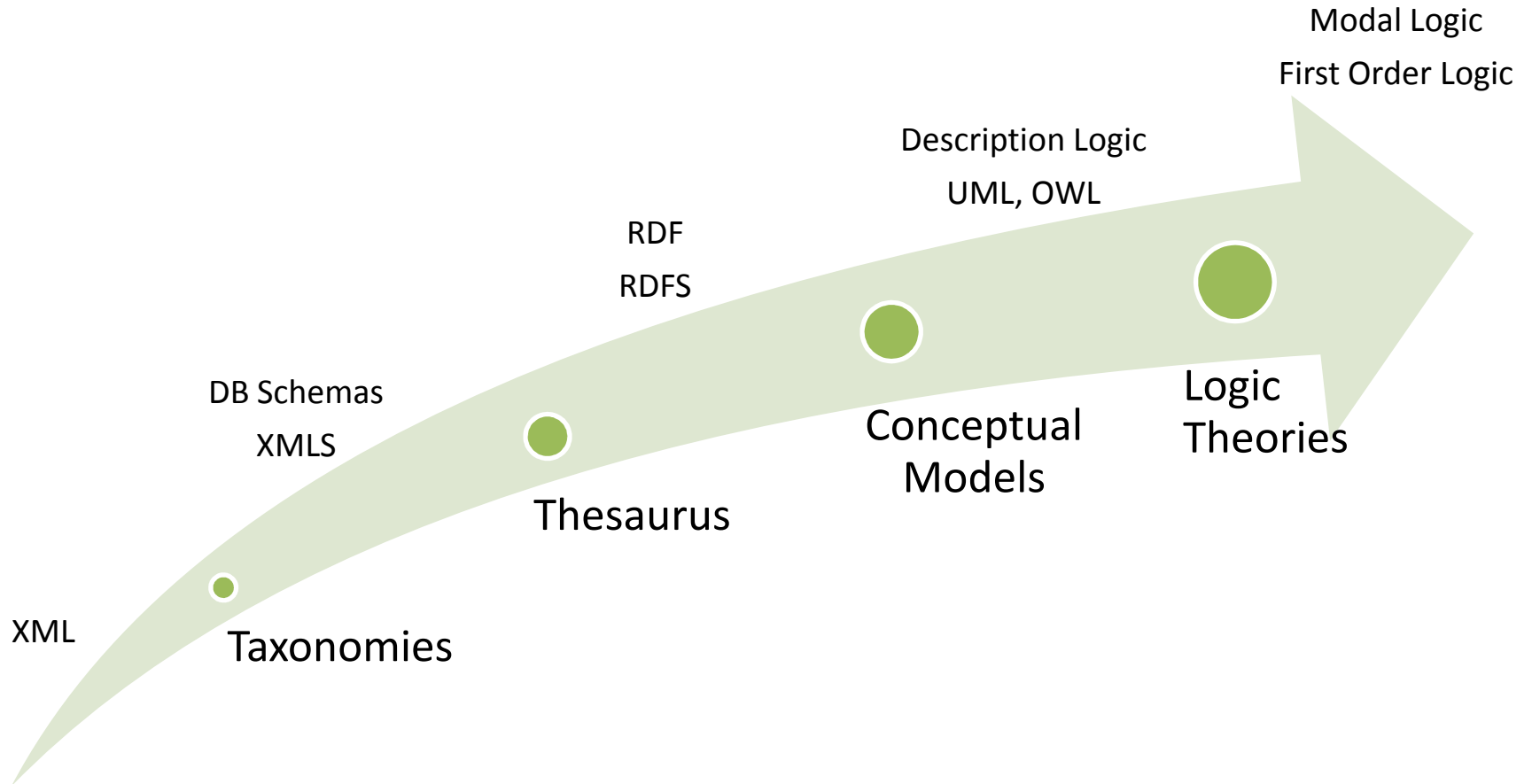
- Which language?
- Which expressivity?
- Reasoning? What about performances?

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

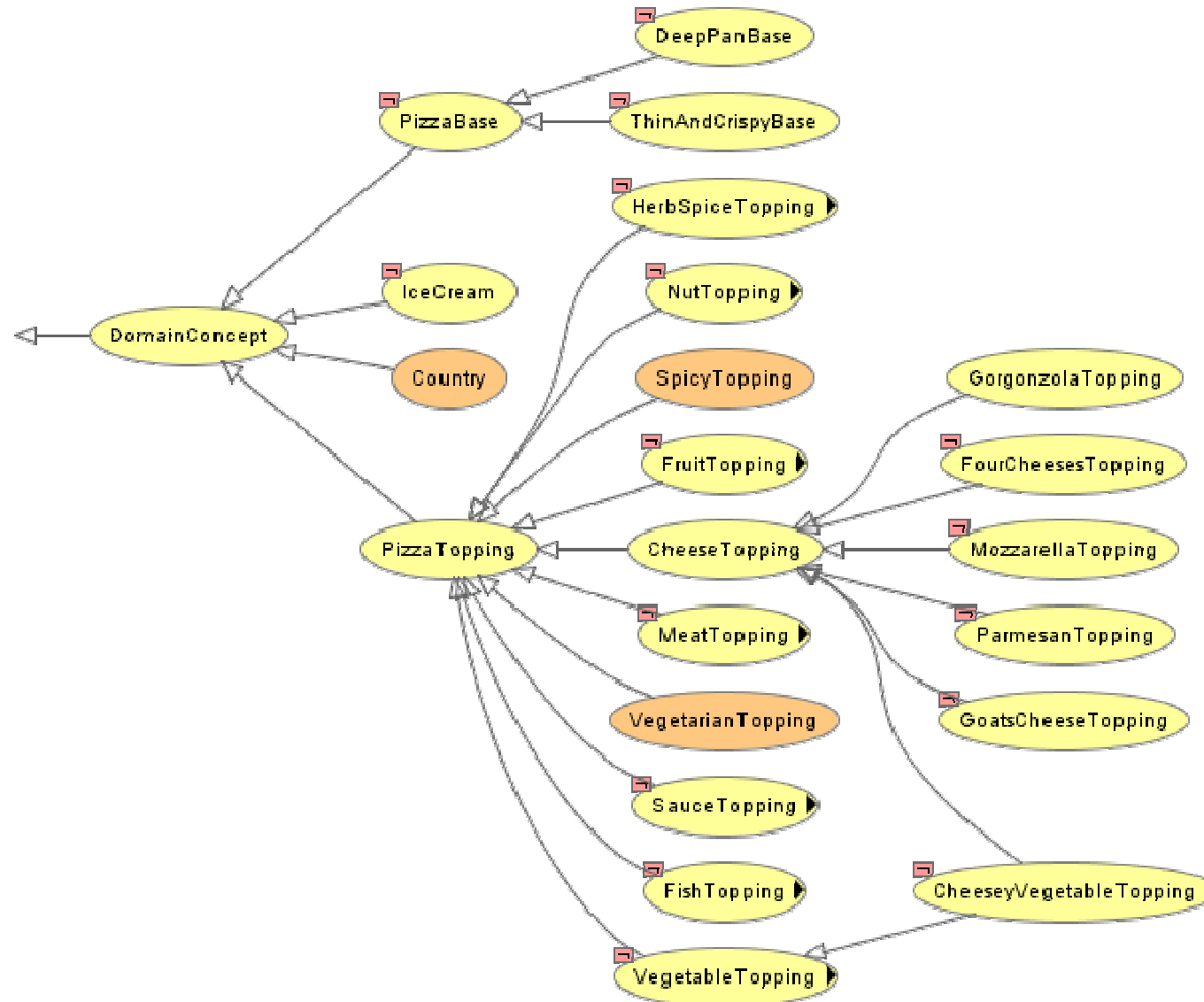
- **Conceptual models**: focused on a particular *domain area*. They specify:
 - Domain *entitites*
 - **Relations** between the entitites (properties and attributes)
 - Rules about classes, roles and relations
 - **Inference mechanisms** -> *Logic theories!!!*

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)

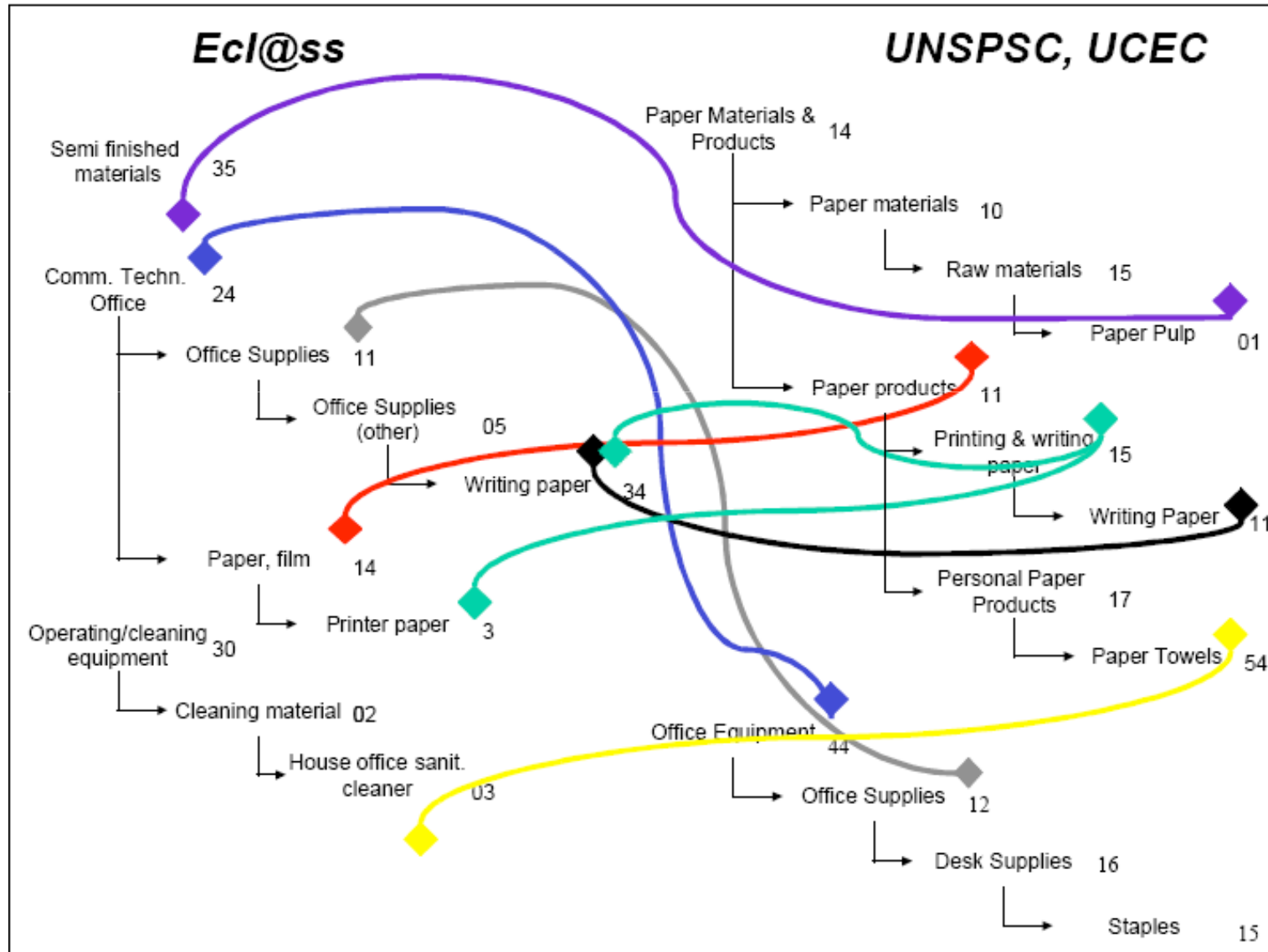
Harvard Business Review, October 2001:

“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 !

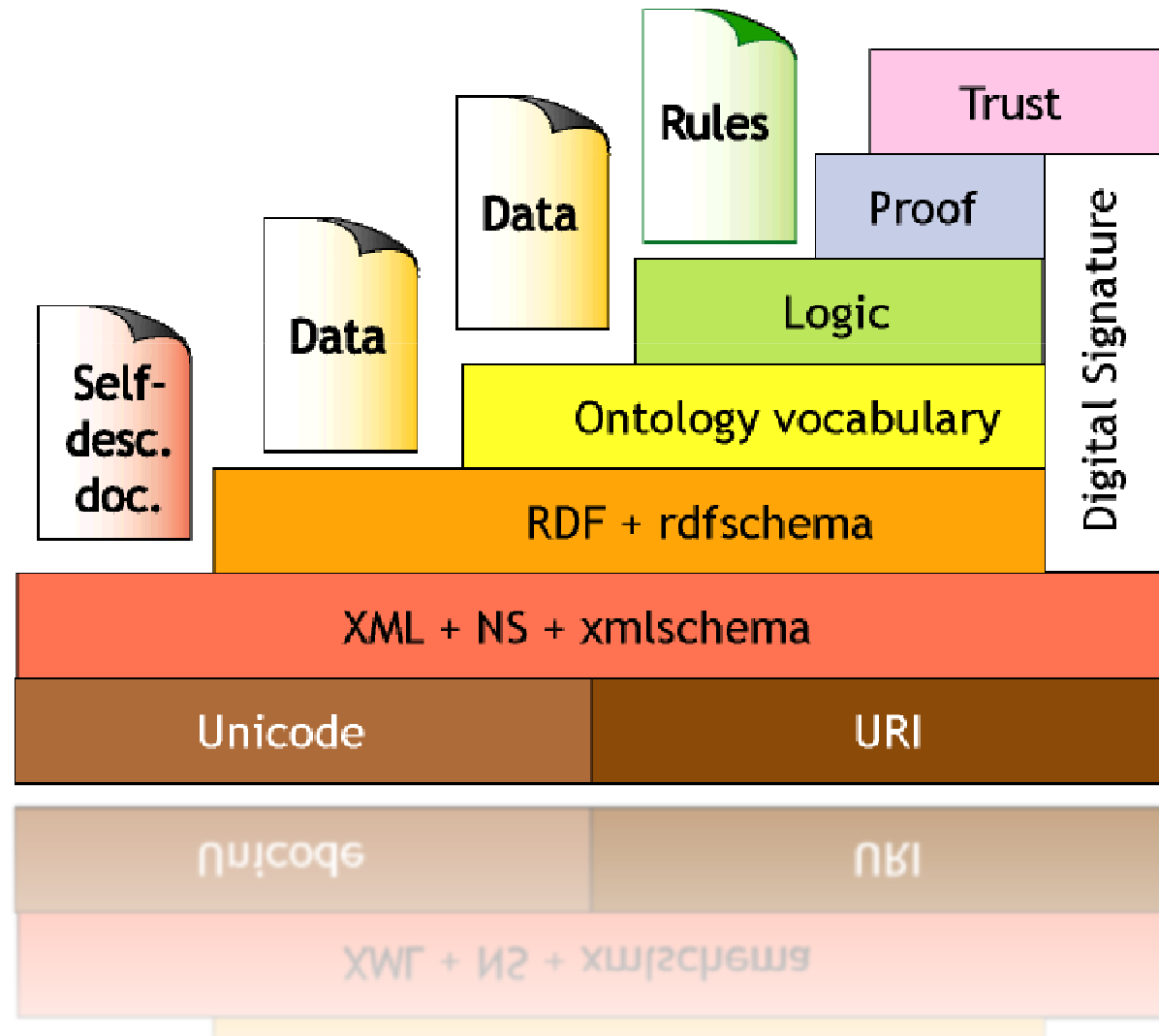
Ontologies

Few examples:

- Dublin Core, focussed on documents
- WordNet
- Gene Ontology, genomics
- Protein Ontology, proteomics
- SnoMed, a very important ontology in the medical field
- 41 use case and examples available at <http://www.w3.org/2001/sw/sweo/public/UseCases/>

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

Uniform Resource Identifiers (URI)

- The URI is used as name system by all the internet
- Beginning of the '90 distinction between
 - Uniform Resource Locator (URL)
 - Uniform Resource Name (URN)
- Nowadays it does not matter anymore, and they are used in an equivalent manner

Uniform Resource Locator

- A particular type of URI
- Identify the resource by expliciting also the access mechanism:

<http://lia.deis.unibo.it/~fc/LIAIndex.html>

Access scheme

Network location

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!

eXtensible Markup Language - XML

- XML is equipped with two different dialects for specifying the grammar of a XML dialect
 - Document Type Definition (DTD)
 - XML Schema
 - Data types
 - Namespaces
 - Uses XML syntax itself
 - Higher expressive power with respect to DTD
- Concept of **valid document** (no grammar) and **well formed document** (defined by means of a grammar)

eXtensible Markup Language - XML

- Complete frameworks for XML Document managing (**DOM** e **SAX**) are available
- **XSL** (XML Stylesheet language)
 - XSLT (XSL Transformation)
 - XPath, language for defining expressions (query/matching XML documents)
 - FO (Formatting Objects)

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:

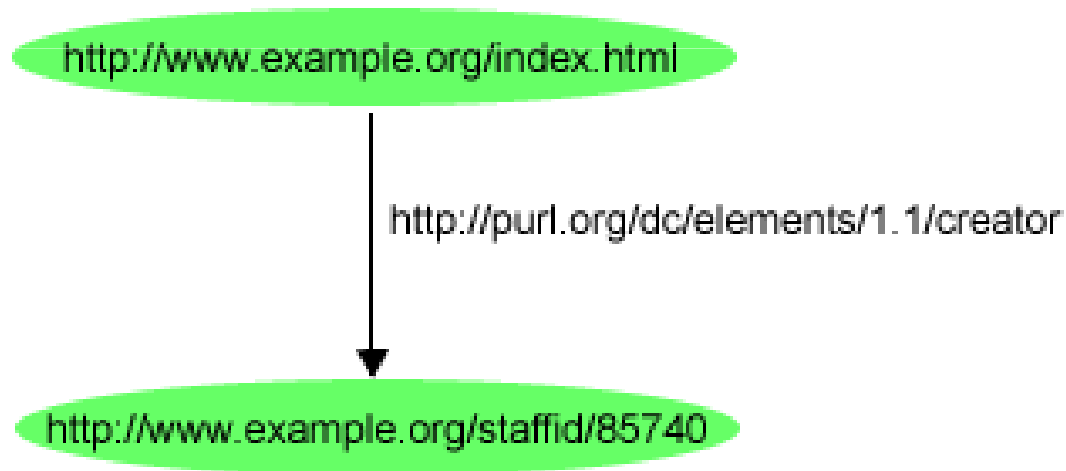
< subject, predicate, object >

< resource, attribute, value >

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

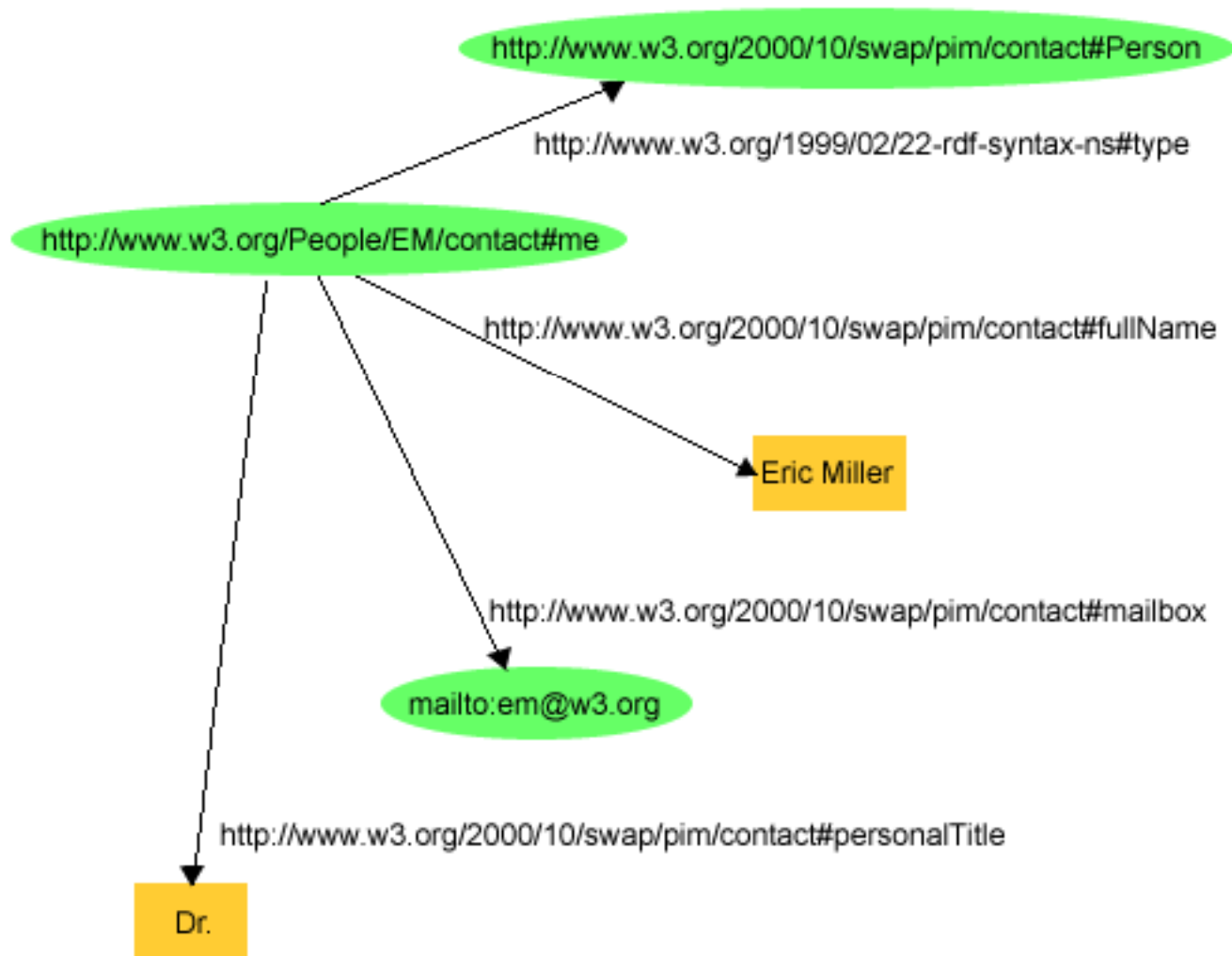
RDF – Graph Representation

- 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"/>
```

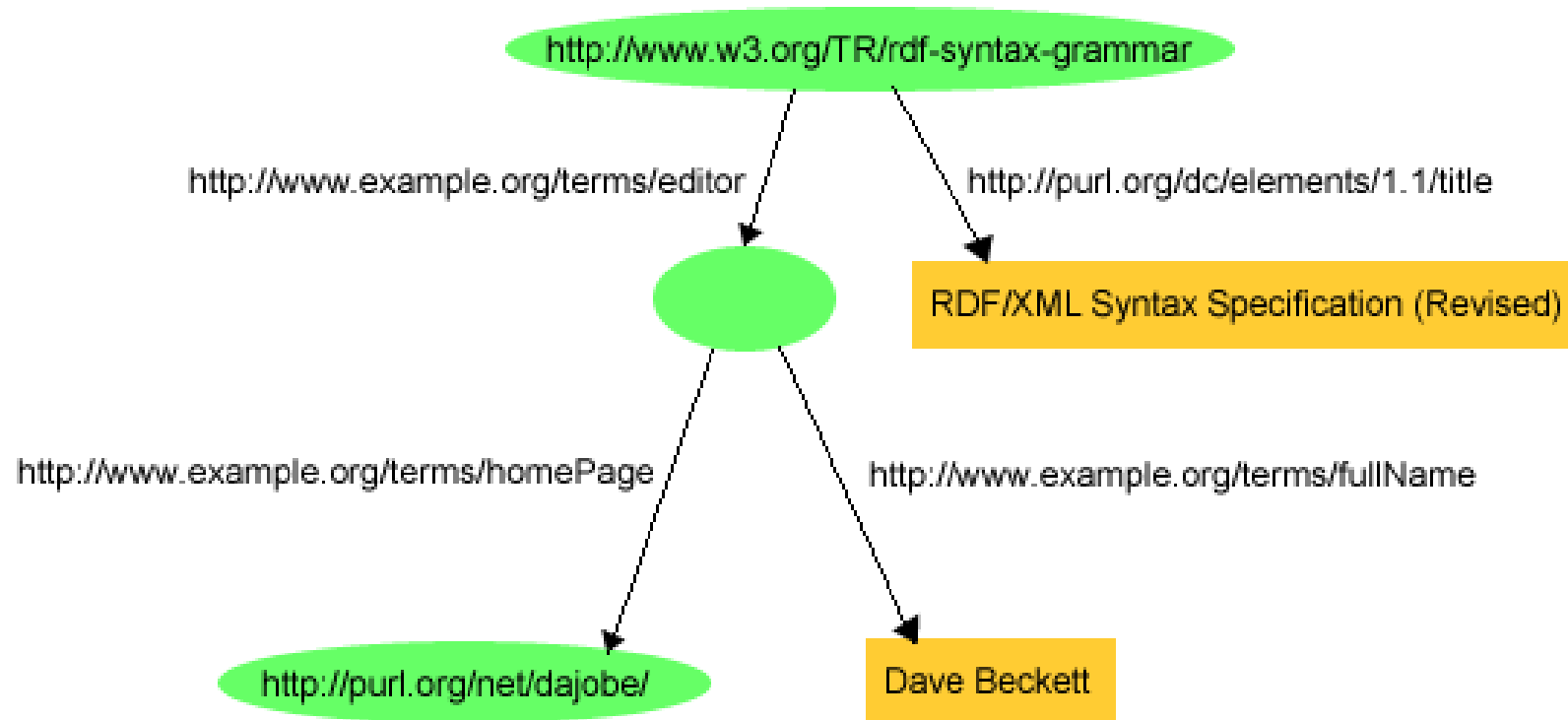
```
    <contact:personalTitle>Dr.</contact:personalTitle>
```

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

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

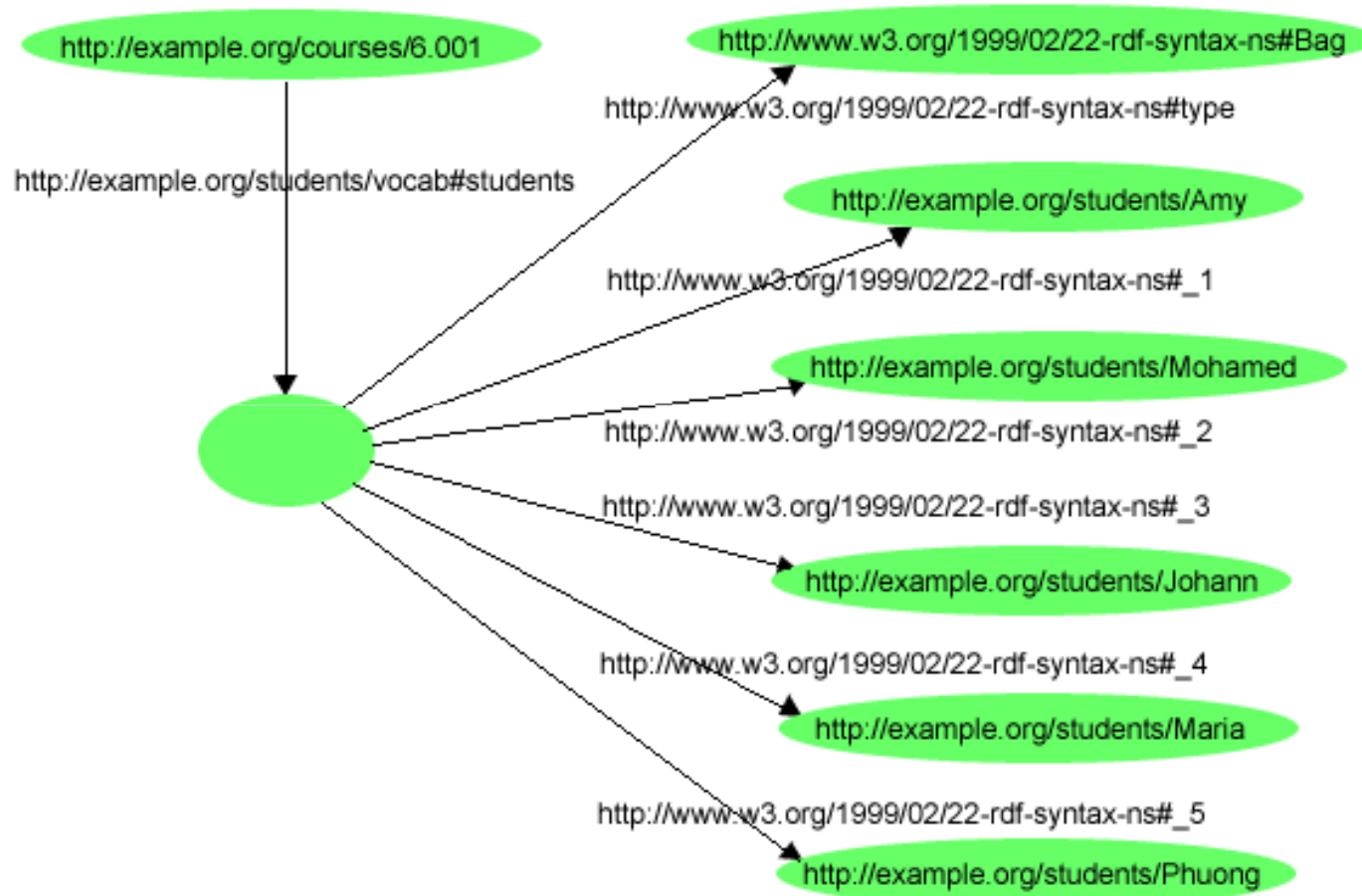
I can query for the mailbox of Eric Miller, without knowing a priori if he uses mailbox or email ...
... if Eric Miller will change mailbox, search result will be coherent!

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 Frameworks

- **JENA**, Java framework for representing and managing RDF statements
- **SPARQL**, Query language for RDF Repository, that supports graph navigation also to different/distributed rdf repositories
 - Agnostic w.r.t. Implementation

RDF Tools

Many tools already available ...

Only in the W3C wiki there are citations for:

- 25 Frameworks/reasoners
- 27 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/>

GRDDL

- GRDDL is a mechanism for **G**leaning **R**esource **D**escriptions from **D**ialects of **L**anguages.
- A technique for obtaining RDF data from XML documents and in particular XHTML pages.
- Authors may explicitly associate documents with transformation algorithms, typically represented in XSLT, using a link element in the head of the document.

Source: GRDDL Primer

<http://www.w3.org/TR/2007/NOTE-grddl-primer-20070628/>

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

POWDER

- POWDER — the Protocol for Web Description Resources — provides a mechanism to describe and discover Web resources and helps the users to make a decision whether a given resource is of interest.
- There are a variety of use cases: from providing a better means to describing Web resources and creating trustmarks to aiding content discovery, child protection and Semantic Web searches.

Source: GRDDL Primer

<http://www.w3.org/TR/2009/NOTE-powder-primer-20090901/>

RIF

- RIF defines a standard for exchanging rules among rule systems, in particular among Web rule engines.
- RIF focuses on exchange rather than defining a single one-fits-all rule language
 - a single language would not cover all popular paradigms of using rules for knowledge representation and business modeling.

Source: RIF W3C Activity

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

SAWSDL

- SAWSDL defines a set of extension attributes for the Web Services Description Language and XML Schema definition language that allows description of additional semantics of WSDL components.
- The specification defines how semantic annotation is accomplished using references to semantic models, e.g. ontologies.
- Semantic Annotations for WSDL and XML Schema (SAWSDL) does not specify a language for representing the semantic models ...
 - it provides mechanisms by which concepts from the semantic models, typically defined outside the WSDL document, can be referenced from within WSDL and XML Schema components using annotations.

Source: SAWSDL W3C Working group

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

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)

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

OWL – Features

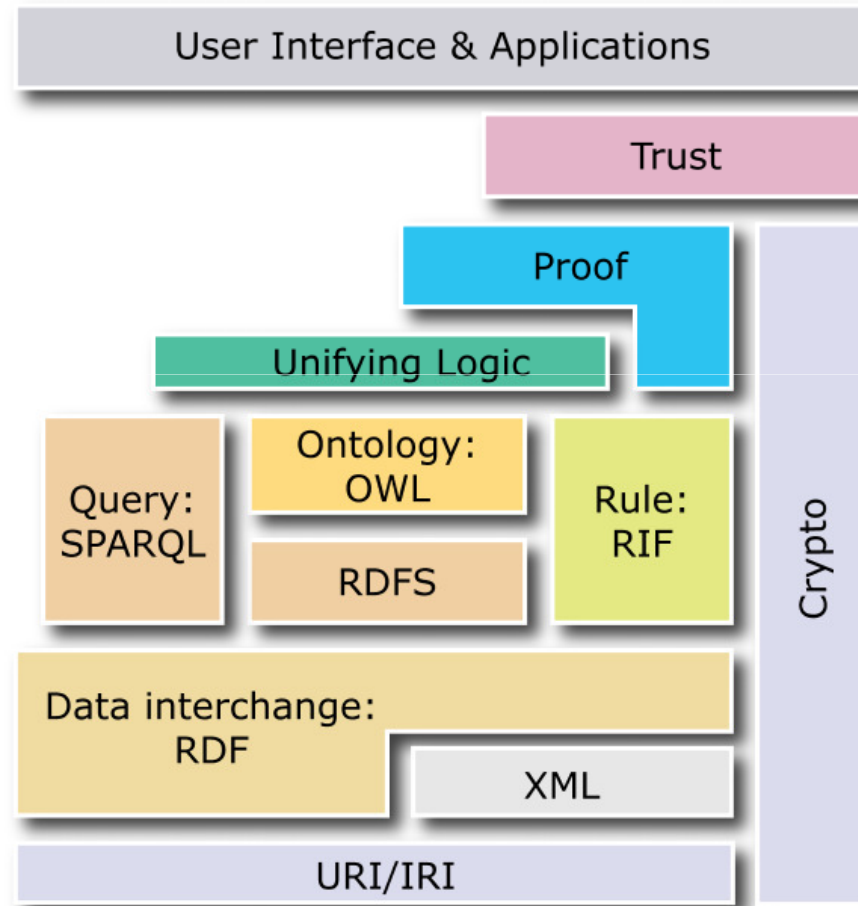
- **Classes (categories)**: subClassOf, intersectionOf, unionOf, complementOf, enumeration, equivalence, disjoint
- **Properties (Roles, Relations)**: symmetric, transitive, functional, inverse Functional, range, domain, subPropertyOf, inverseOf, equivalentProperty
- **Instances (Individuals)**: sameIndividualAs,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

OWL Tools

- Many tools for OWL
 - Editors (19 listed at <http://www.w3.org/2001/sw/wiki/Category:Editor>)
 - Reasoners (24 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!)
 - A lot of research about in the academic world
 - A bit less interest in the industrial world
 - Recently, some interest for Semantic Web Services
- 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
- 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)
- 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...

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

Concluding...

- Instead of Semantic Web, we should use the term Data Web (Berners-Lee)
- From the focus on the reasoning ...
- ... to the data, and especially the re-use of a big amount of data already available in the web

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 tio 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/2001/sw/>

Communities:

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

Thanks for the attention

Questions?

Ing. Federico Chesani, Phd

c/o DEIS – Facoltà di Ingegneria

Viale Risorgimento 2

40136 - Bologna

federico.chesani@unibo.it

Tel. 051 20 93086