

## 12 – The Semantic Web and RDF

In the previous episodes...

A (video) summary:

**Michael Wesch: “Web2.0... The Machine is Us/ing Us”**

<http://www.youtube.com/watch?v=6gmP4nk0EOE>

Università della Svizzera italiana	Facoltà di scienze della comunicazione	I
		3

- Describing syntax ("classic" HTML) → Describing data (XML)
- Describing data (XML) → Describing *knowledge* (???)

"The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined *meaning*, better enabling computers and people to work in cooperation"

Tim Berners-Lee  
**The Semantic Web**  
Scientific American, 2001

According to [Jim Hendler](#), two main directions:

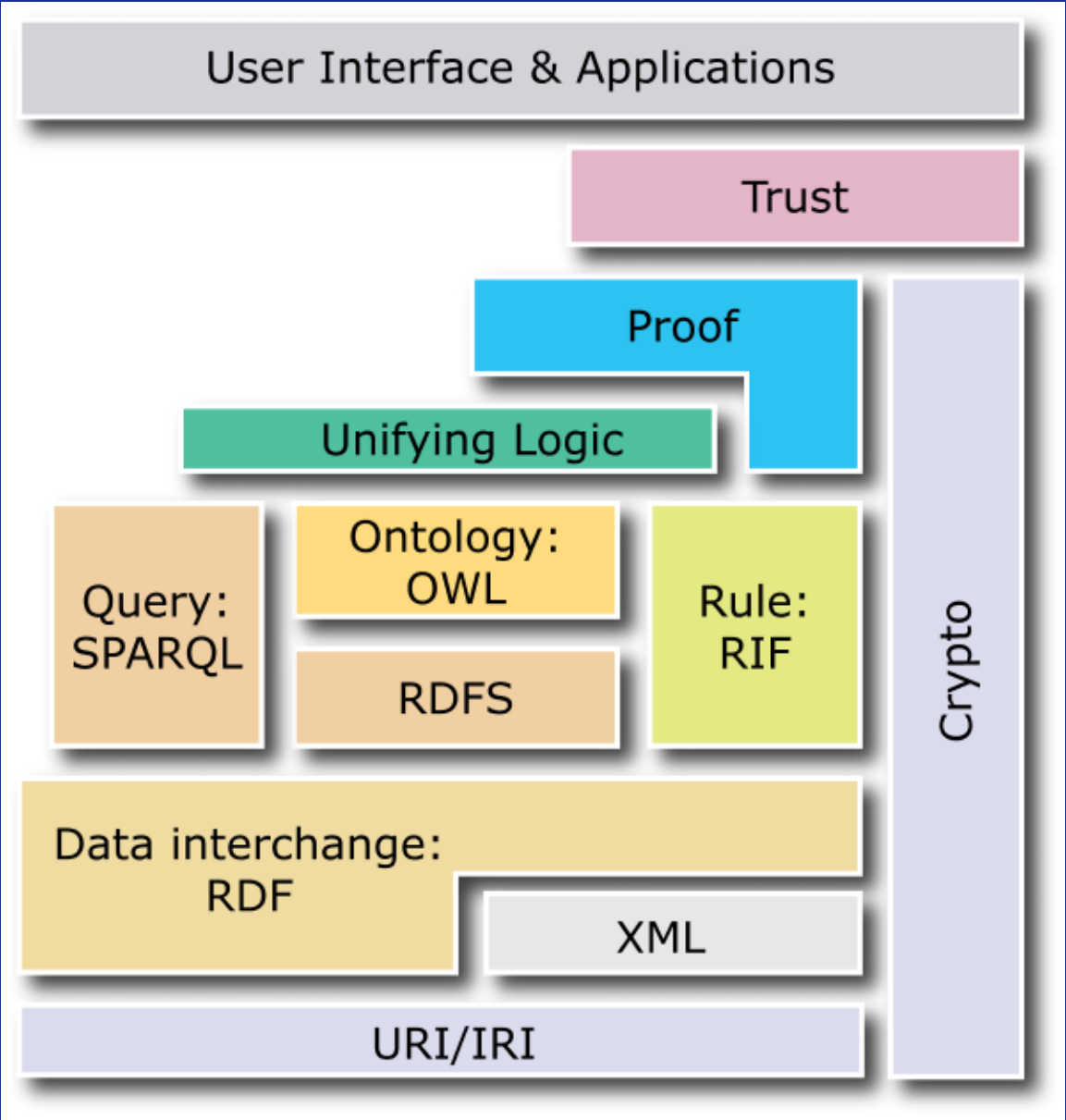
## ■ Web

- Web-based apps with little semantics
- Emphasis on *linking data* using URIs
- Standards: RDF (Resource Description Framework) and SPARQL

## ■ Semantic

- Models to represent knowledge in an expressive way
- *Inference of new knowledge* by using reasoners
- Standards: OWL (Web Ontology Language)

# The Semantic Web Layer Cake



# RDF at a glance

ROW SELECTION



1	<i>Bleach</i>	Nirvana	1989	4
2	<i>Nevermind</i>	Nirvana	1991	4,5



8	<i>My World 2.0</i>	Justin Bieber	2010	N/A
---	---------------------	---------------	------	-----



5	<i>Black Album</i>	Metallica	1991	4
6	<i>Master of Puppets</i>	Metallica	1986	5
7	<i>The Number of the Beast</i>	Iron Maiden	1982	5

ID	Title	Band	Year	Rating
1	<i>Bleach</i>	Nirvana	1989	4
2	<i>Nevermind</i>	Nirvana	1991	4,5
3	<i>Ten</i>	Pearl Jam	1991	4,5
4	<i>Vitalogy</i>	Pearl Jam	1994	4
5	<i>Black Album</i>	Metallica	1991	4
6	<i>Master of Puppets</i>	Metallica	1986	5
7	<i>The Number of the Beast</i>	Iron Maiden	1982	5
8	<i>My World 2.0</i>	Justin Bieber	2010	N/A

COLUMN SELECTION



Title	Band
<i>Bleach</i>	Nirvana
<i>Nevermind</i>	Nirvana
<i>Ten</i>	Pearl Jam
<i>Vitalogy</i>	Pearl Jam
<i>Black Album</i>	Metallica
<i>Master of Puppets</i>	Metallica
<i>The Number of the Beast</i>	Iron Maiden
<i>My World 2.0</i>	Justin Bieber



Band
Nirvana
Nirvana
Pearl Jam
Pearl Jam
Metallica
Metallica
Iron Maiden
Justin Bieber



Year	Rating
1989	4
1991	4,5
1991	4,5
1994	4
1991	4
1986	5
1982	5
2010	N/A

## CELL SELECTION !!!

ID	Title	Band	Year	Rating
1	<i>Bleach</i>	Nirvana	1989	4
2	<i>Nevermind</i>	Nirvana	1991	4,5
3	<i>Ten</i>	Pearl Jam	1991	4,5
4	<i>Vitalogy</i>	Pearl Jam	1994	4
5	<i>Black Album</i>	Metallica	1991	4
6	<i>Master of Puppets</i>	Metallica	1986	5
7	<i>The Number of the Beast</i>	Iron Maiden	1982	5
8	<i>My World 2.0</i>	Justin Bieber	2010	N/A



Row 1 Title  
Bleach



Row 2 Band  
Nirvana

Row 4 Rating  
4

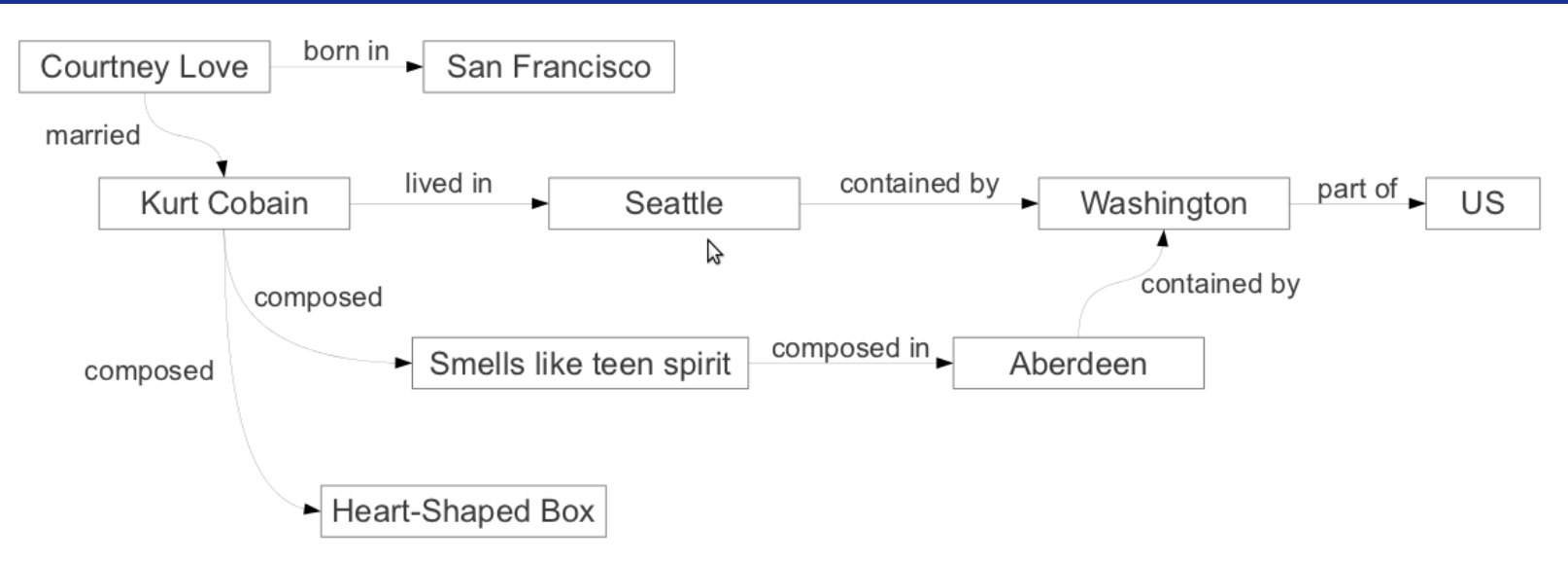
Row 6 Band  
Metallica



Row 8 Year  
2010

# RDF at a glance

<b>Subject</b>	<b>Predicate</b>	<b>Object</b>
Courtney Love	born in	San Francisco
Courtney Love	married	Kurt Cobain
Kurt Cobain	lived in	Seattle
Kurt Cobain	composed	Smells like teen spirit
Kurt Cobain	composed	Heart-Shaped Box
Seattle	contained by	Washington
Washington	part of	US
Smells like teen spirit	composed in	Aberdeen
Aberdeen	contained by	Washington





## ■ Not this one

### Reality distortion field

From Wikipedia, the free encyclopedia



It has been suggested that this article or section be [merged](#) into *Steve Jobs*. ([Discuss](#))

**Reality distortion field** is a term coined by [Bud Tribble](#) at [Apple Inc.](#) in 1981, to describe company co-founder [Steve Jobs'](#) [charisma](#) and its effects on the developers working on the Mac project.<sup>[1]</sup> Later the term has also been used to refer to perceptions of his keynote (or [Stevenote](#)) by observers and devoted users of [Apple computers](#) and products.<sup>[2]</sup>

Bud Tribble claimed that the term came from [Star Trek](#).

In essence, RDF is the idea that Steve Jobs is able to convince people to believe almost anything with a mix of [charm](#), [charisma](#), [bluster](#), [exaggeration](#), and [marketing](#). RDF is said to distort an audience's sense of proportion or scale. Small advances are

## ■ Resource Description Framework

### ■ W3C recommendation (2004)

### ■ a *Semantic Web specification* together with OWL

## ■ Graph data model

## ■ Abstract syntax based on the concept of *triple*

## ■ Serialization in different text-based formats (including XML)

- RDF is based on the following ideas:
  - things being described have *properties* which have *values*
  - resources can be described by making *statements* that specify those properties and values
- These statements are called *triples*:
  - the *Subject* is the resource the statement is about
  - the *Predicate* identifies the property or the characteristic
  - the *Object* identifies the value of the property
- Example:

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

Subject  
 Predicate  
 Object



# Literals, Resources, URIs

- Each element in a triple can belong to two different types:
  - Resource
    - `http://www.whatever.com/index.html#me`
    - `dc:creator`
  - Literal
    - Plain: "666", "English", "April, 8 2009"
    - Typed: "27"^^xsd:integer, "2009-04-08"^^xsd:date
- Subjects and predicates can only be resources, while objects can be resources or literals
- Resources are identified by Uniform Resource Identifiers (URIs)
  - URLs are a particular kind of URI
  - URI reference = URI + *fragment identifier*
    - *i.e.* `http://www.example.org/index.html#section2`

- As in any XML document, all elements have to belong to a given namespace
  - NOTE: in the XML serialization, properties can become either elements or attributes
- Example (RDF header + CD description):

```
<?xml version="1.0"?>
```

```
<rdf:RDF
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:cd="http://www.recshop.fake/cd#">
```

```
<rdf:Description
```

```
  rdf:about="http://www.recshop.fake/cd/Empire_Burlesque">
```

```
  <cd:artist>Bob Dylan</cd:artist>
```

```
  <cd:country>USA</cd:country>
```

```
  <cd:company>Columbia</cd:company>
```

```
  <cd:price>10.90</cd:price>
```

```
  <cd:year>1985</cd:year>
```

```
</rdf:Description>
```

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

- Look at this address:

```

Students:123456                                     (subj)
students:address                                    (pred)
"765 San Antonio Ave, Palo Alto, CA 94304".        (obj)
  
```

- What if we want to be able to access the single elements of the address?

```

(subj)          (pred)          (obj)
students:123456 students:address studaddrid:654321 .
studaddrid:654321 students:street "765 San Antonio Ave" .
studaddrid:654321 students:city  "Palo Alto" .
studaddrid:654321 students:state "CA" .
studaddrid:654321 students:zip   "94304" .
  
```

- studaddrid:654321 is a *universal* identifier, but we won't need it again in other documents

- solution is to use a *local, anonymous* node which does not need an identifier

```
studaddrid:654321 => _:anon123
```

RDF and XML models are fundamentally different

- RDF has a very simple model which consists of *labeled arcs*
- Any specific group of RDF declarations forms a *graph* that can be serialized in XML
- XML data model is a *labeled tree*, which is less flexible for describing metadata

## Resources used in RDF and XML Schema are different

- In RDF, nodes do not necessarily appear inside the document itself, but could be any resource which has a URI (typically *external*)
- RDF is a language *for metadata*
- The nodes an XML Schema refers to are *internal* to the XML document, in a specific location within the structure of a document

## The semantics of RDF and XML schemas are different

- RDF schemas have an interpretation which is primarily *semantic*
- XML schemas have an interpretation which is primarily *syntactic*
- RDF is used to build (*model*) *knowledge*, where tree-based representation structures are not sufficient
- XML schemas are used to model *documents*



## ■ An example:



## ■ How would you render it in XML?

```
<author>
  <uri>book</uri>
  <name>Bob</name>
</author>
```

```
<document href="book">
  <author>Bob</author>
</document>
```

```
<document>
  <details>
    <uri>href="book"</uri>
    <author>
      <name>Bob</name>
    </author>
  </details>
</document>
```

- Meaning is not hardcoded in tag names. What you have here:

```
<document>
  <author>
    <uri>href="book"</uri>
    <details>
      <name>Bob</name>
    </details>
  </author>
</document>
```

is interpreted by a machine as a meaningless text:

```
<v>
  <x>
    <y> a="ppppp"</y>
    <z>
      <w>qqqqq</w>
    </z>
  </x>
</v>
```

... what are the relations between elements, now?

## ■ RDF/XML

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">

  <rdf:Description
    rdf:about="http://en.wikipedia.org/wiki/Lugano">
    <dc:title>Lugano</dc:title>
    <dc:publisher>Wikipedia</dc:publisher>
  </rdf:Description>
</rdf:RDF>
```

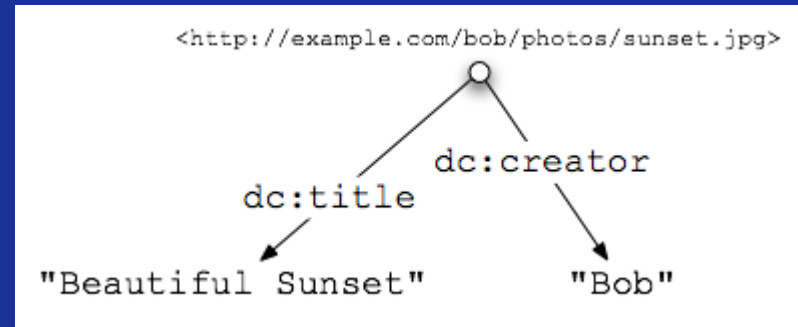


## ■ RDFa

```
<div about="/alice/posts/trouble_with_bob">  
  <h2 property="dc:title">The trouble with Bob</h2>
```

The trouble with Bob is that he takes much better photos than I do:

```
<div about="http://example.com/bob/photos/sunset.jpg">  
    
  <span property="dc:title">Beautiful Sunset</span>  
  by <span property="dc:creator">Bob</span>.  
</div>  
</div>
```



## ■ Some Web references:

- Why RDF is different from XML: <http://www.w3.org/DesignIssues/RDF-XML>
- RDF Primer: <http://www.w3.org/TR/REC-rdf-syntax>
- Dean Allemang, Jim Hendler: "Semantic Web for the Working Ontologist".  
<http://workingontologist.org>
- FOAF: <http://www.foaf-project.org>
- Dublin Core: <http://dublincore.org>

## ■ Tools:

- W3C RDF Validator: <http://www.w3.org/RDF/Validator>
- Morla RDF editor: <http://www.morlardf.net>
- FOAF-o-matic: <http://www.foaf-o-matic.org>