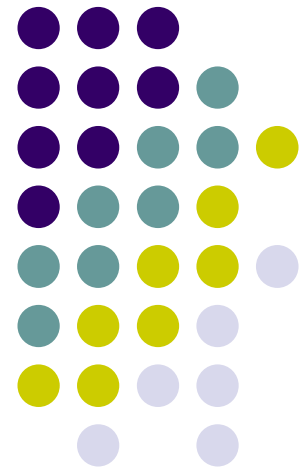


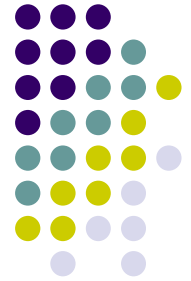
Ontologies

Lina Al-Jadir
UNIGE - SYINF

Cours Bases d'informations
Printemps 2007



Outline



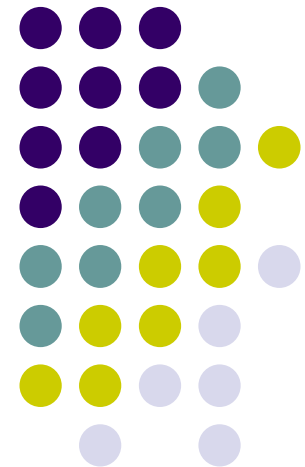
- Introduction
 - Motivations
 - Definition
 - Classification
 - Requirements
- Ontology representations
 - Logic based models:
 - RDF(S)
 - DL
 - OWL
 - Database like models: Kaon
- Conclusion

Ontologies

Introduction

Ontology representations

Conclusion





Motivations

- Information sharing & communication
 - People, software systems need to understand each other to be able to share knowledge, communicate, and interoperate
- Applications:
 - Web → semantic web
 - more semantics in web pages
 - semantic-based search (instead of keyword-based search engines)
 - agent-agent communication in e-business, e-commerce
 - Information systems integration
 - add semantics to understand data to integrate



Motivations

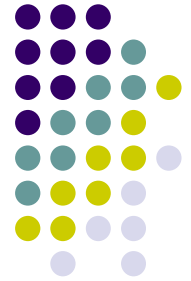
→ Add **more semantics** to information

- From the less to the more structured solutions:
 - Natural language with some markups
 - Meta-data such as in XML-based languages
 - or data models such as RDF (Resource Description Framework)
 - ↓ ● ONTOLOGIES



Definition

- What is an ontology?
 - Specification of a conceptualization [Gruber 1993]
 - Representation of a shared understanding of a domain of interest
 - Formal and machine processable representation of a domain of interest



Definition

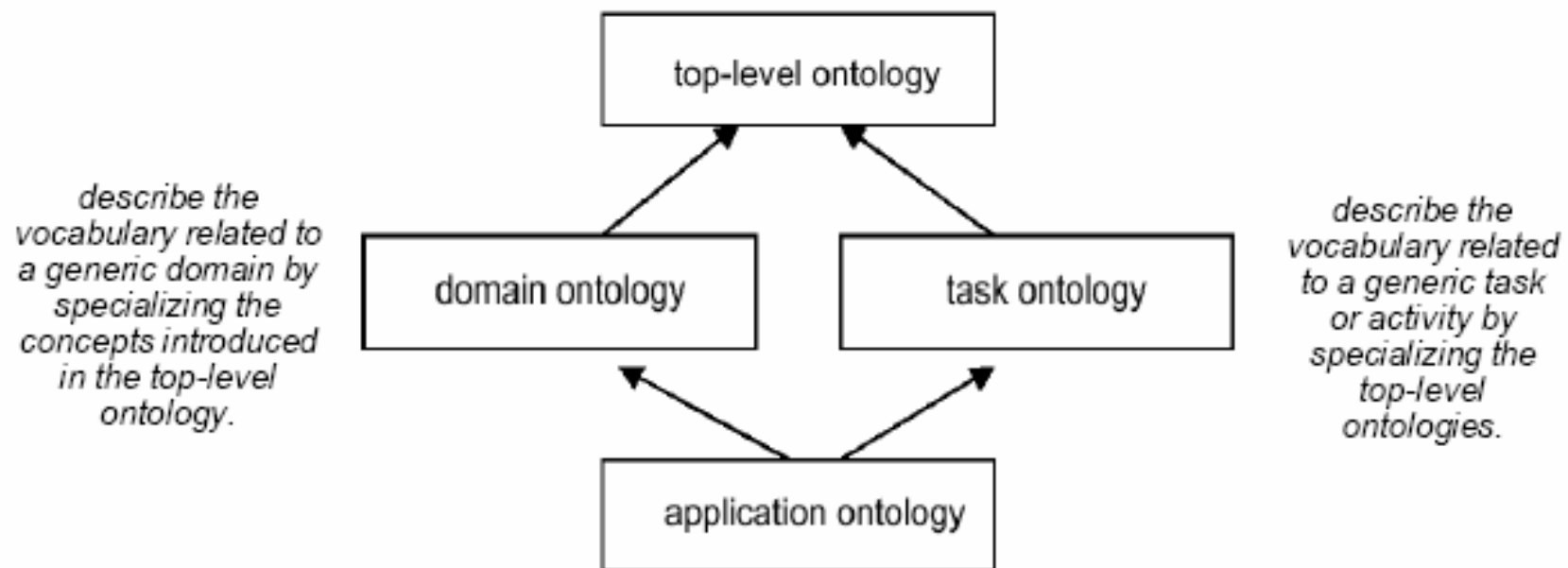
- **Hierarchical description** of important **concepts** in a domain along with descriptions of their **properties**
 - descriptions have been **agreed upon by** the user community of the ontology
- Ontologies can be used as a **reference** for the interpretation of the semantics of data
 - applications refer to the ontology to choose and interpret the concepts they are using

Classification



[Guarino, 98]

describe very general concepts like space, time, event, which are independent of a particular problem or domain. It seems reasonable to have unified top-level ontologies for large communities of users.

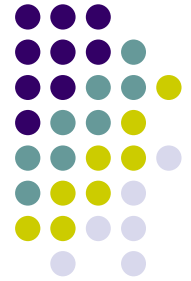


These are the most specific ontologies. Concepts in application ontologies often correspond to roles played by domain entities while performing a certain activity.



Other Classification

- 2 kinds:
 - Thesaurus-like ontologies
 - Descriptive ontologies
- In both cases, an ontology has reasoning capabilities.

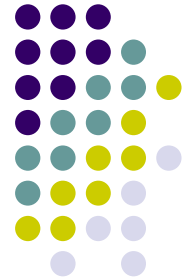


Thesaurus-like Ontology

- Thesaurus-like ontology
 - definitions of terms
 - terms organized in a **subsumption** hierarchy and linked by other relationships to express **synonymy**, **composition**, etc....
 - provides a reference vocabulary
- First generation of ontologies
- Example: Wordnet

Wordnet I

(<http://wordnet.princeton.edu>)



WordNet 2.0 Search

Search word:

Overview for "horse"

The **noun** "horse" has 5 senses in WordNet.

1. **horse**, Equus caballus -- (solid-hoofed herbivorous quadruped domesticated since prehistoric times)
2. **horse** -- (a padded gymnastic apparatus on legs)
3. cavalry, horse cavalry, **horse** -- (troops trained to fight on horseback; "500 horse led the attack")
4. sawhorse, **horse**, sawbuck, buck -- (a framework for holding wood that is being sawed)
5. knight, **horse** -- (a chessman in the shape of a horse's head; can move two squares horizontally and one vertically (or vice versa))

Search for of senses

☒ Show glosses

☐ Show contextual help



Wordnet II

Search word:

Results for "Synonyms, ordered by estimated frequency" search of noun "horse"

5 senses of horse

Sense 1

horse, *Equus caballus* -- (solid-hoofed herbivorous quadruped domesticated since prehistoric times)

=> equine, equid -- (hoofed mammals having slender legs and a flat coat with a narrow mane along the back of the neck)

Sense 2

horse -- (a padded gymnastic apparatus on legs)

=> gymnastic apparatus, exerciser -- (sports equipment used in gymnastic exercises)

Sense 3

cavalry, horse cavalry, horse -- (troops trained to fight on horseback; "500 horse led the attack")

=> military personnel, soldiery, troops -- (soldiers collectively)

Sense 4

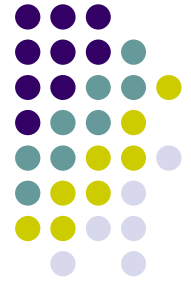
sawhorse, horse, sawbuck, buck -- (a framework for holding wood that is being sawed)

=> framework, frame, framing -- (a structure supporting or containing something)

Sense 5

knight, horse -- (a chessman in the shape of a horse's head; can move two squares horizontally and one vertically (or vice versa))

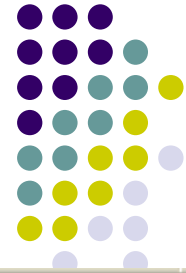
=> chessman, chess piece -- (any of 16 white and 16 black pieces used in playing the game of chess)



Descriptive Ontology

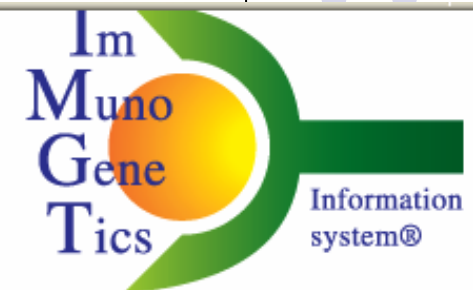
- **Descriptive ontology**
 - goes beyond terminology
 - defines **conceptualizations** that include **properties** of concepts and **relationships** between concepts
 - similar to a database schema
 - achieves a semantically rich representation of the intended domain
- Ex: ImMunoGeneTics international medical ontology

ImMunoGeneTics



WELCOME !
to IMGT/GENE-DB

THE
INTERNATIONAL
IMMUNOGENETICS
INFORMATION SYSTEM®



<http://imgt.cines.fr>

IMGT/GENE-DB Query page

GENERAL CRITERIA:

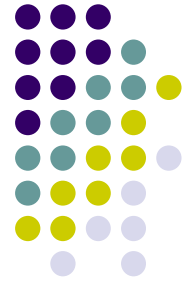
- ☒ Species
- ☒ Locus
- ☒ Gene type
- ☒ Group
- ☒ Subgroup
- ☒ Functionality
- ☒ Selection of genes which have been found

☒ Chromosomal orphon set

☒ any ☐ rearranged ☐ transcribed ☐ translated

do the search

reset



Requirements for ontologies

- Ontologies look like databases but ...
- Requirements for ontologies are different from those for databases
 - Data modeling
 - Instance handling
 - Reasoning
 - Querying

Requirements for ontologies I

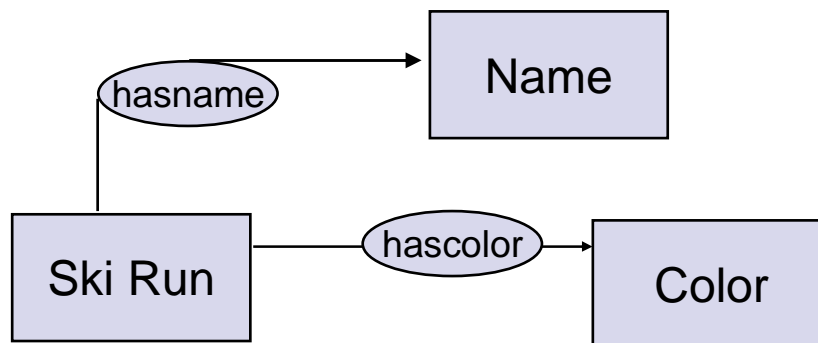


1) Data modeling

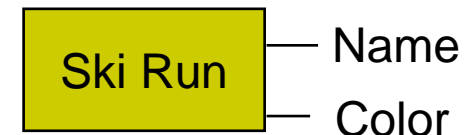
- Need for rich modeling concepts

BUT most (all?) of the time:

- objects have an identity and binary relationships with objects (roles) or values (attributes)

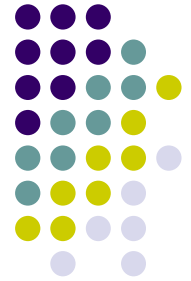


ontology schema



database schema

Requirements for ontologies I



- Modeling concepts:
 - Objects / Instances / **Individuals**
 - Elements of the domain of discourse
 - Types / Classes / **Concepts**
 - Sets of objects sharing certain characteristics
ex: {skirun1, skirun2, ...} for skirun
 - Relations / Properties / **Roles**
 - Sets of pairs (tuples) of objects
ex: {<skirun1, blue>, <skirun2, black>, ...} for color

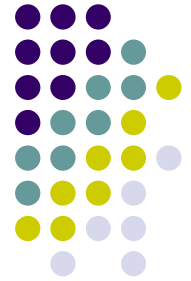
Requirements for ontologies II



2) Instance handling

- in DB:
 - instances created after the schema is completed
 - **closed-world assumption**: Instances have to comply with the schema and its constraints (ex. students should be linked to a university)
- in ontologies:
 - instances do not necessarily exist, may be added anytime
 - **open-world assumption**: Information in the ontology may not be complete. Instances are accepted as long as they do not contradict the knowledge already described (ex. students should also be linked to a university, but if they are not we suppose it will be done later)

Requirements for ontologies III



3) Reasoning

- Define **derived concepts**: specify the conditions that their instances must satisfy
 - ex: SwissStudent is a student from Switzerland
 - managed as the other concepts
- Definition/building of the **generalization (subsumption) hierarchy** includes derived concepts

Requirements for ontologies IV



4) Ontology querying:

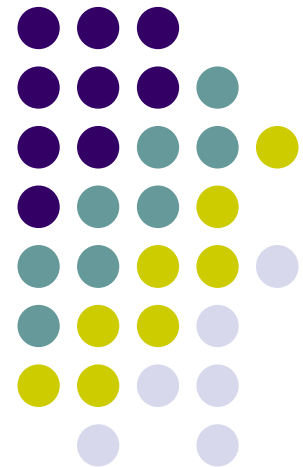
- Schema exploration:
 - Give the characteristics of a role (transitive, symmetric, inverse)
 - Give the roles going from (or to) a concept
- Reasoning on the schema:
 - Are two concepts equivalent or disjoint? Does a concept (or role) subsume another one? Classify the whole set of concepts. What are the super or sub-concepts (at any level) of a concept?
- Reasoning on the data:
 - To which most specific concepts does this individual belong?

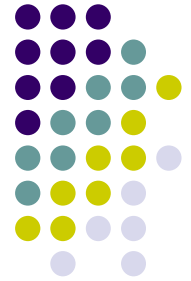
Ontologies

Introduction

Ontology representations

Conclusion





Languages for ontologies

- Logic-based languages

- RDF(S)
- Description Logics (DL)
- DAML+OIL / OWL

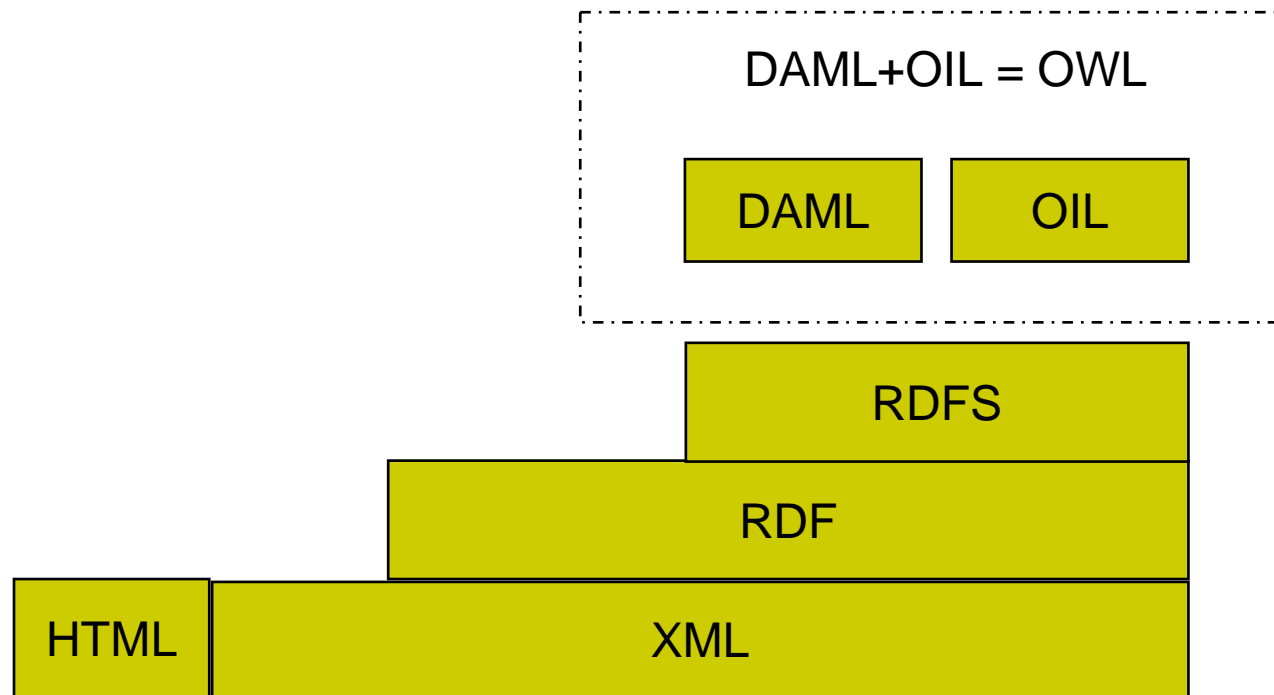
- Database-based languages

- Kaon
- Dogma

- Semantic web languages

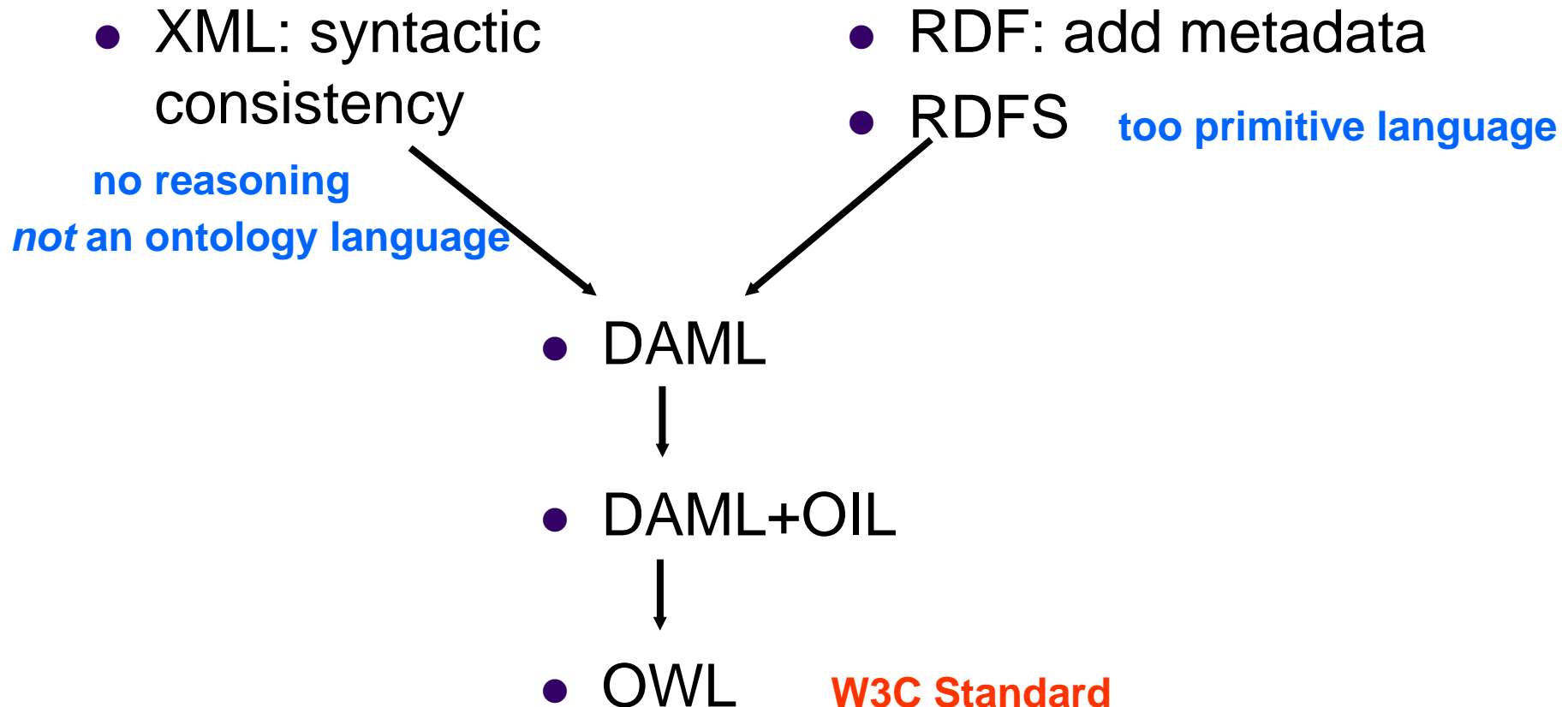
- RDF(S)
- DAML+OIL / OWL

Layer Language Model for the www



["Spinning the semantic web", D. Fensel, 2003]

Evolution of Semantic Web (Ontology) Languages



Resource Description Framework (RDF)



- **RDF** is a general-purpose language for representing **metadata on the Web**.

<http://www.w3.org/RDF/>

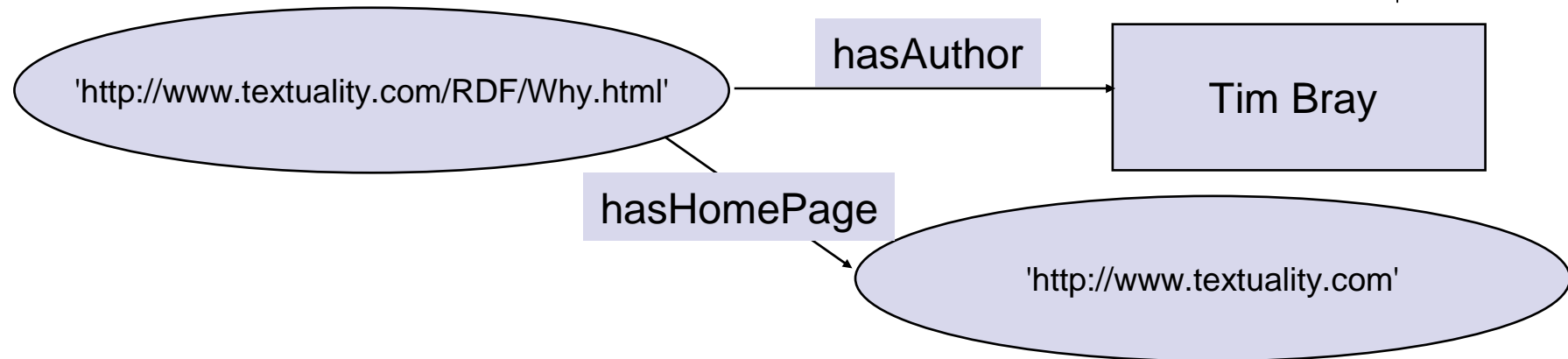
- It proposes a vocabulary which is XML-based to describe resources.



RDF Data Model

- **Resource**: any object that can be pointed to by a URI (Uniform Resource Identifier):
 - a document, a picture, a paragraph on the Web;
 - <http://www.textuality.com/RDF/Why.html> ...
- **Property**
 - Properties themselves are also resources (URIs)
 - ex: hasAuthor, hasTitle.
- **Statement** describes properties of resources
 - Triple: <Resource, Property, Value>.
 - ex:
 - <<http://www.textuality.com/RDF/Why.html>, hasAuthor, Tim Bray>
 - <<http://www.textuality.com/RDF/Why.html>, hasHomePage, <http://www.textuality.com>>

RDF example



```
<rdf:Description about='http://www.textuality.com/RDF/Why.html'>  
  <hasAuthor>Tim Bray</hasAuthor>  
  <hasHomePage rdf:resource='http://www.textuality.com' />  
</rdf:Description>
```

- Every **Description** element describes a resource
- Every **attribute** or **nested element** inside a Description is a property of that resource



RDF(S)

- **RDF(S)** – schema language
- Extends RDF with **schema elements**:
 - Class, Property
 - subClassOf, subPropertyOf
 - range, domain
 - But only primitive (non-derived) classes
- RDF(S) is considered as an ontology language



RDF(S) example

- ex:

<Person, **rdf:type**, **rdfs:Class**>

<hasAuthor, **rdf:type**, **rdf:Property**>

<TimBray, **rdf:type**, Person>

<hasAuthor, **rdfs:domain**, Page>

<hasAuthor, **rdfs:range**, Person>

<Professor, **rdfs:subClassOf**, Person>



RDF(S) Data Model

