

# Ontologies and Knowledge Representation

Gilles Falquet  
Université de Genève,  
Centre universitaire d'informatique

1

## Outline

Introduction, definitions  
Ontologies in NLP: Wordnet  
Ontologies in other computing domains

2

## Ontology and Ontologies

**Ontology** (philo.) The branch of metaphysics dealing with the nature of **being**. In particular:

- Categories of being
- Entities and types of entities
- Relationships between entities

**An ontology.** Enumeration/description/organization of existing entities.

- hierarchies of concepts
- definitions / axioms
- etc.

3

## Some Definitions

"a shared and common understanding of a domain that can be communicated between people and heterogeneous and distributed systems." [Fensel, 2000];

"a computer model of some portion of the world." [Huhns & Singh, 1997];

"a specification of a conceptualization." [Gruber, 2000].

"a catalog of the types of things that are assumed to exist in a domain of interest **D** from the perspective of a person who uses a language **L** for the purpose of talking about **D**. [Sowa 2003]

4

## Formal and Informal Ontologies

**Informal ontology:** specified by a catalog of types that are either undefined or defined only by statements in a natural language.

**Formal ontology:**

- a collection of names for concept and relation types organized in a partial ordering by the type–subtype relation.
- a formal way to distinguish subtypes from their supertype

from [Sowa 2003]

5

## Concept

A class of objects grouped according to their properties.  
≠ usual sense of ‘concept’ = general notion, abstract idea

Concept **extension**: all the objects having the desired properties = the **instances** of the concept.

Concept **intension**: the properties that define the concept.

6

## The is-a Semantic relation

Generic/specific: A is more specific than B (A is-a B) if

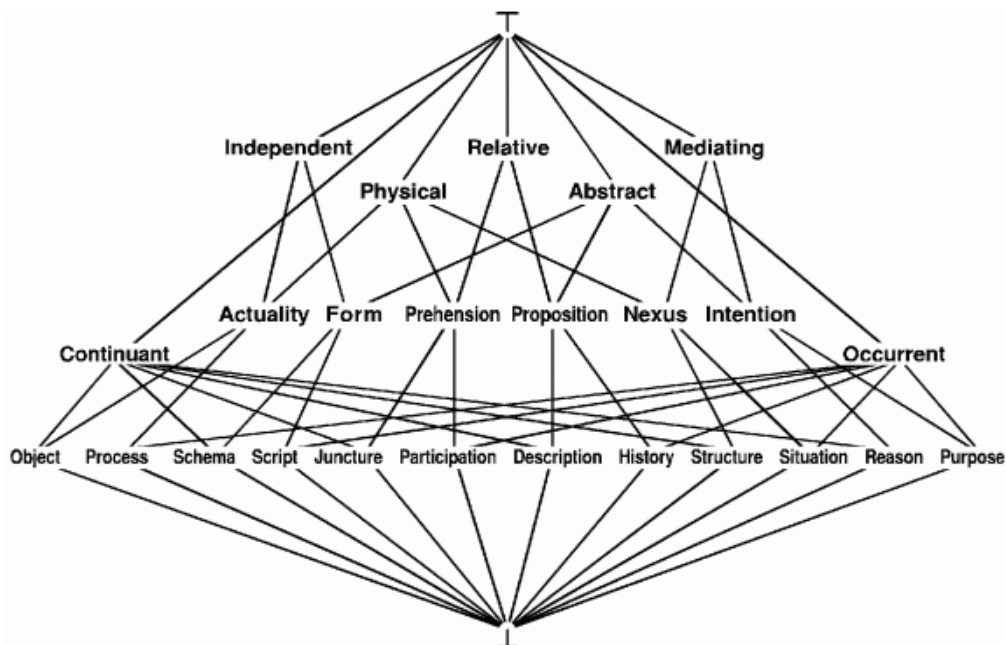
every instance of A is an instance of B  
(inclusion of the extensions)

“all cars are vehicles”  
“all humans are animals”

Ontologies are usually organized according to the is-a relation.

7

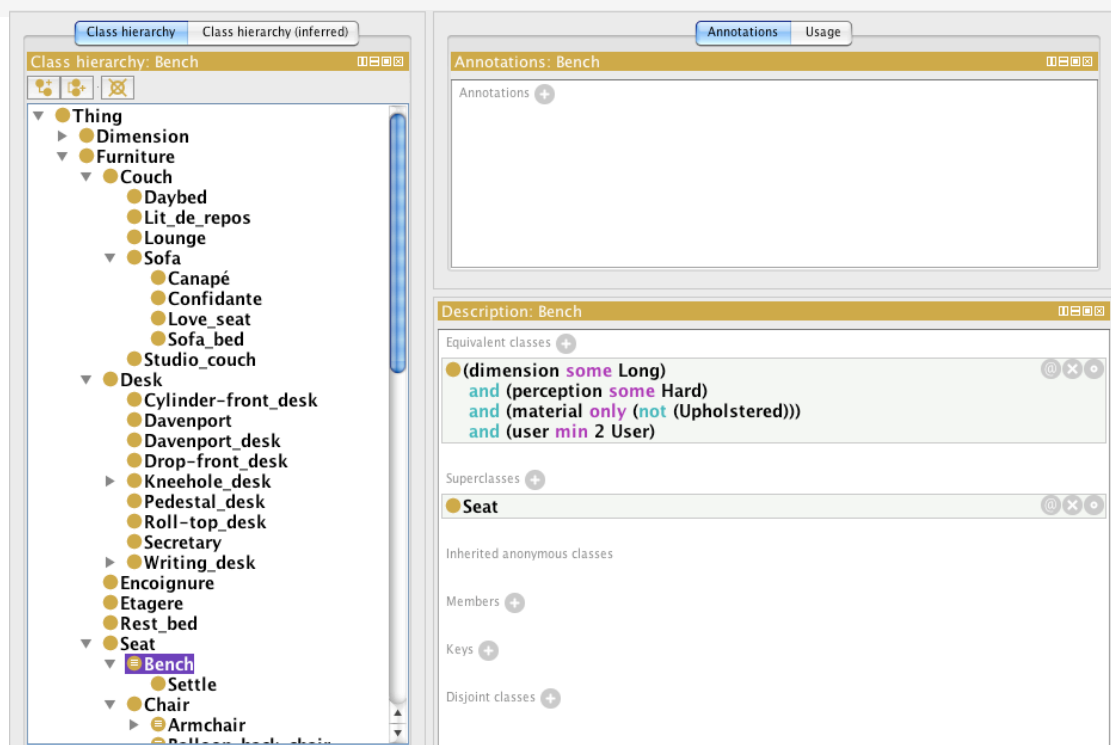
## A Top Level Ontology



John F. Sowa. *Knowledge Representation*

8

## A Domain Ontology



## Categories of Formal Ontologies

**Axiomatized ontology:** distinguishes subtypes by axioms and definitions stated in a formal language, (logic or some computer-oriented notation that can be translated to logic)

**Prototype-based ontology:** distinguishes subtypes by a comparison with a typical member or *prototype* for each subtype.

from [Sowa 2003]

Sowa, J. (2003). Ontology. Retrieved from <http://www.jfsowa.com/ontology/> on 2010-10-27.

## Ontologies and knowledge representation in Computing Sciences

Computational linguistics

⇒ Lexical ontologies

Artificial intelligence

⇒ reasoning on world states

Information systems

⇒ object-based analysis and design

⇒ system interoperability

⇒ semantic web

11

## Why do we use ontologies?

To offer a shared and common understanding of some domain the can be communicated across people and application systems [[Fensel, 2000](#)]

To enable knowledge sharing [[Gruber, 2000](#)]

[Fensel, 2000] Fensel D, 2000: Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce, URL: <http://www.cs.vu.nl/%7E Dieter/ftp/paper/silverbullet.pdf>

[Gruber, 2000] Gruber T, 2000: What is an Ontology?, URL: <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>

12

## Computational Linguistics: Lexical Ontologies

Goal: associate senses with words

The lexicon “determines” the ontology (what is not named does not exist).

Show semantic relationships between senses

13

## Wordnet: a Lexical Ontology

Based on the English (Spanish, ...) lexicon

Connect each form (sequence of letters) to its senses

A sense (concept) is a ‘synset’

≥ 166 000 terms (word, sense)

118 000 forms (words)

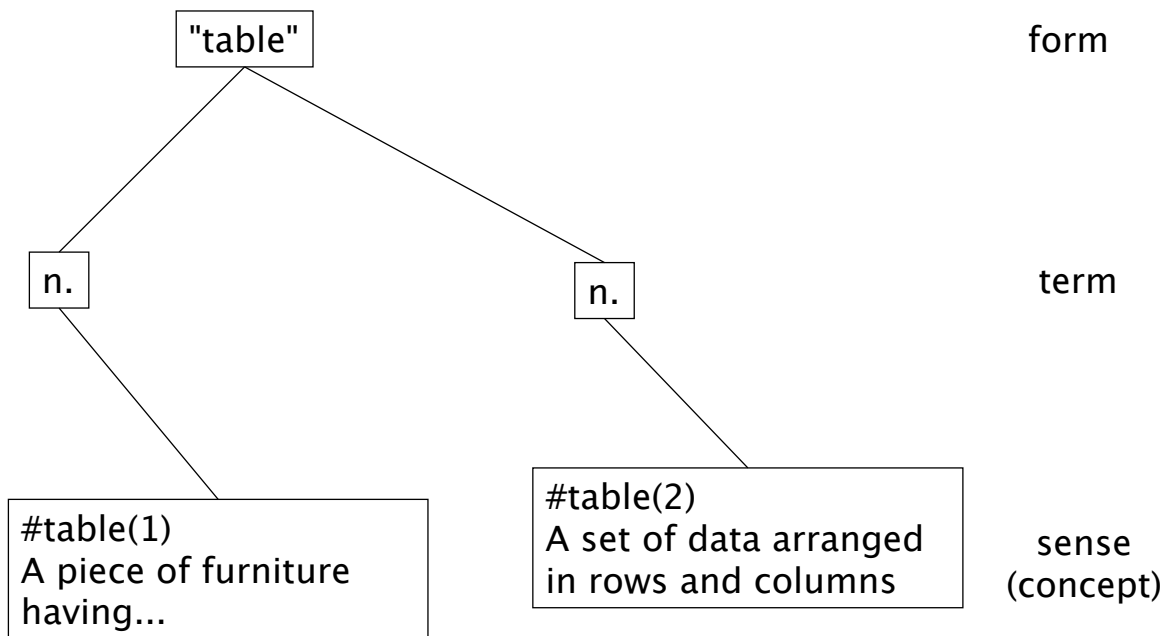
90 000 senses

17% of the words are polysemous

40% of the words have a synonym

14

# Senses



Enter a word to search for:

KEY: "S:" = Show Synset (semantic) relations, "W:" = Show Word (lexical) relations

**Noun**

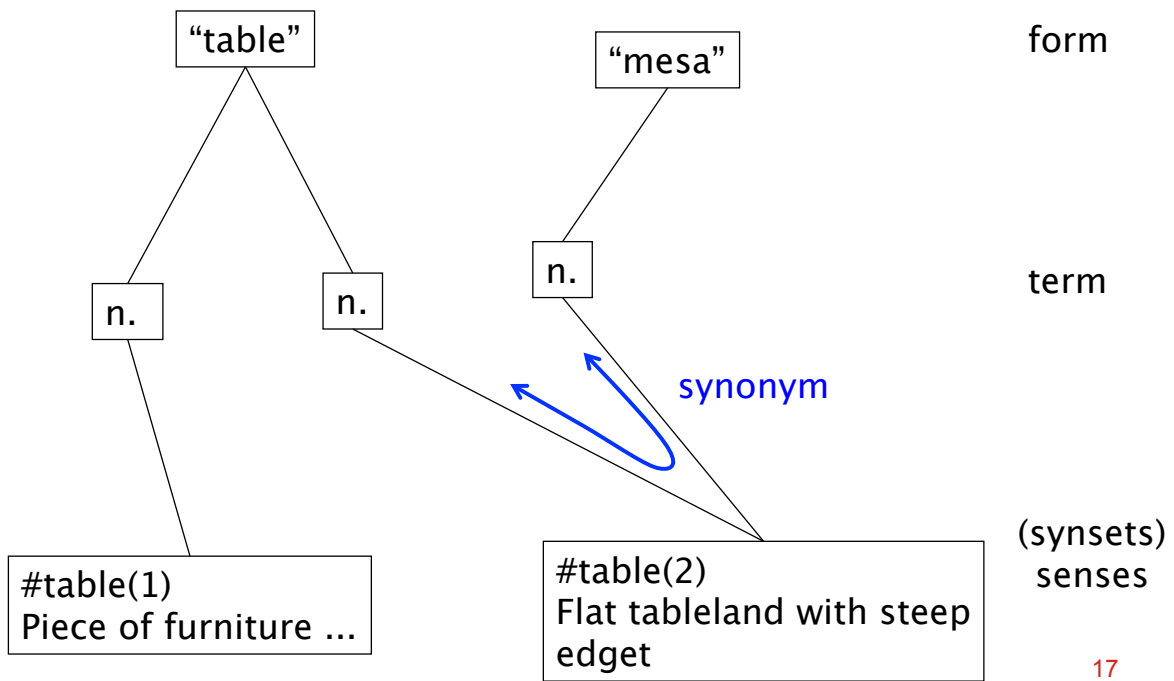
- **S: (n) table, tabular array** (a set of data arranged in rows and columns) *"see table 1"*
- **S: (n) table** (a piece of furniture having a smooth flat top that is usually supported by one or more vertical legs) *"it was a sturdy table"*
- **S: (n) table** (a piece of furniture with tableware for a meal laid out on it) *"I reserved a table at my favorite restaurant"*
- **S: (n) mesa, table** (flat tableland with steep edges) *"the tribe was relatively safe on the mesa but they had to descend into the valley for water"*
- **S: (n) table** (a company of people assembled at a table for a meal or game) *"he entertained the whole table with his witty remarks"*
- **S: (n) board, table** (food or meals in general) *"she sets a fine table"; "room and board"*

**Verb**

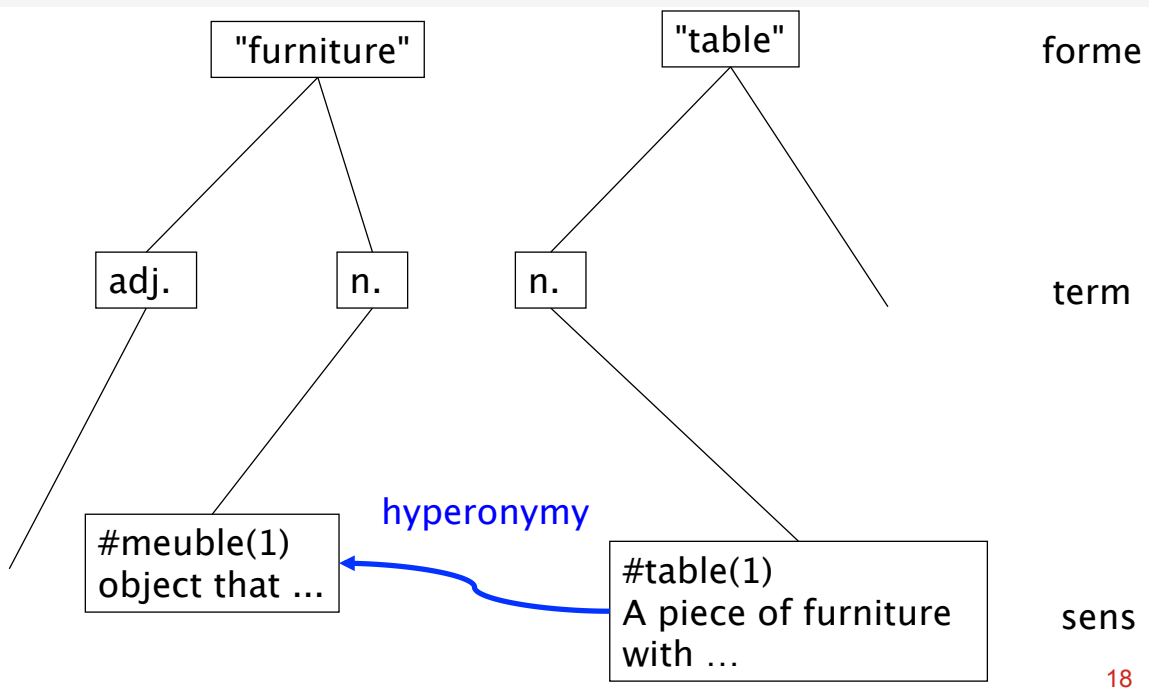
- **S: (v) postpone, prorogue, hold over, put over, table, shelve, set back, defer, remit, put off** (hold back to a later time) *"let's postpone the exam"*



# Synsets

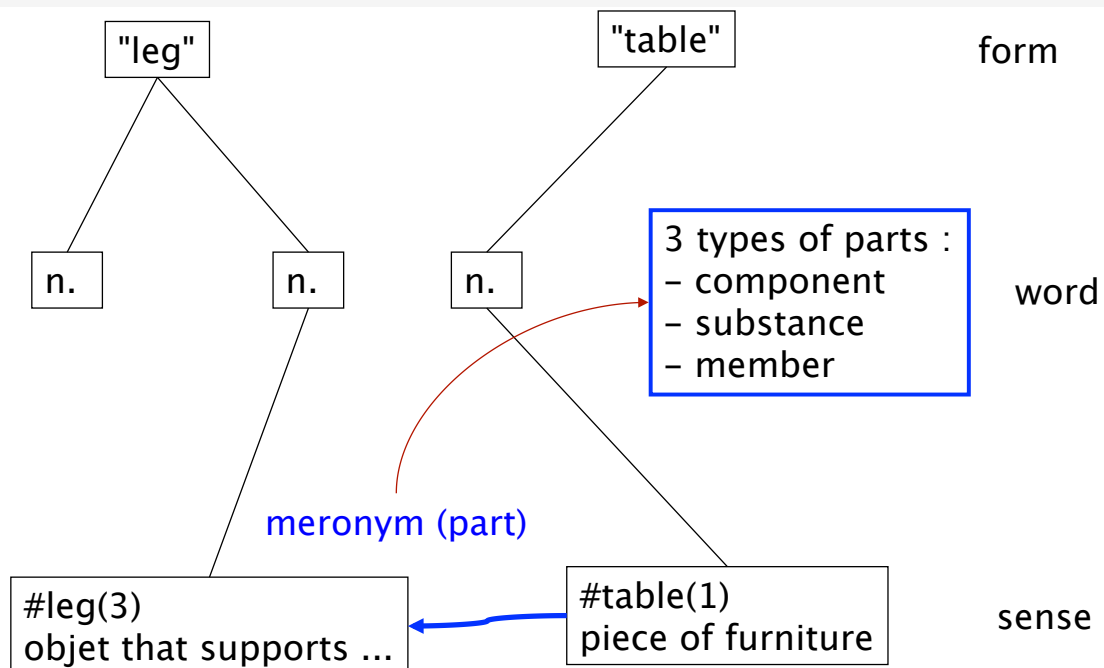


# Semantic Relations



- **S: (n)** [administration](#), [governance](#), [governing body](#), [establishment](#), [brass](#), [organization](#), [organisation](#) (the persons (or committees or departments etc.) who make up a body for the purpose of administering something) *"he claims that the present administration is corrupt"; "the governance of an association is responsible to its members"; "he quickly became recognized as a member of the establishment"*
  - *direct hyponym / full hyponym*
    - **S: (n)** [Curia](#) ((Roman Catholic Church) the central administration governing the Roman Catholic Church)
    - **S: (n)** [top brass](#) (the most important persons in a governing body)
    - **S: (n)** [executive](#) (persons who administer the law)
      - **S: (n)** [Bush administration](#) (the executive under President George W. Bush)
      - **S: (n)** [Clinton administration](#) (the executive under President Clinton)
      - **S: (n)** [Bush administration](#) (the executive under President George H. W. Bush)
      - **S: (n)** [Reagan administration](#) (the executive under President Reagan)
      - **S: (n)** [Carter administration](#) (the executive under President Carter)
    - **S: (n)** [judiciary](#), [bench](#) (persons who administer justice)
    - **S: (n)** [county council](#) (the elected governing body of a county)
    - **S: (n)** [government officials](#), [officialdom](#) (people elected or appointed to administer a government)
      - **S: (n)** [civil service](#) (government workers; usually hired on the basis of competitive examinations)
        - **S: (n)** [Whitehall](#) (the British civil service)
      - **S: (n)** [bureaucracy](#), [bureaucratism](#) (nonelective government officials)
        - **S: (n)** [Pentagon](#) (the United States military establishment)
    - **S: (n)** [management](#) (those in charge of running a business)
      - **S: (n)** [house](#) (the management of a gambling house or casino) *"the house gets a percentage of every bet"*

## Meronymy



## Applications

In natural language processing

- Normalization (unify synonyms)
- Word sense disambiguation

In information retrieval

- Query expansion (with synonyms, hyponyms, ...)

21

## Other Language-Related Resources

### **Terminological database**

- Set of entries comprised of Term, Definition, Source, Reliability, Synonyms, Generic Term, Translation, etc.
- Represent the language used in a specific domain

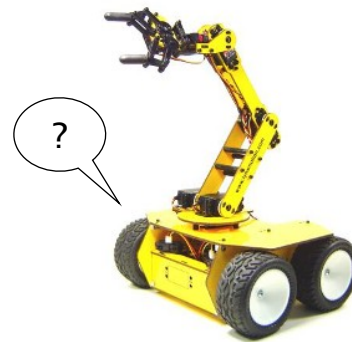
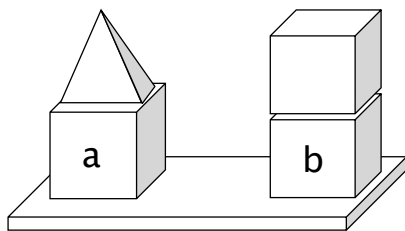
### **Thesaurus**

- Define a controlled vocabulary (e.g. for indexing)
- Relations are not formally defined ("broader term", "narrower term", "see also", "used for", ...)
- Some thesauri are nevertheless formally organized in a is-a hierarchy

22

## Ontologies in Artificial Intelligence

Intelligent systems must be able to “understand” the world, to infer implicit facts, etc.



“Put a on b !”

23

## Logical Models

Use propositional and predicate logic

- to represent world state
- to represent inference rules

Inference engines to deduce implicit facts, to find solutions,

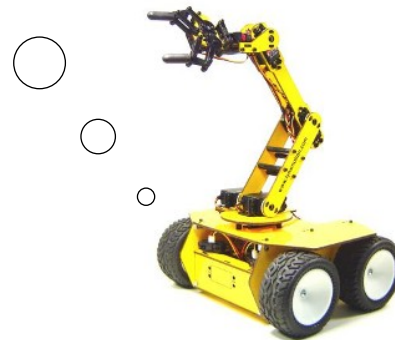
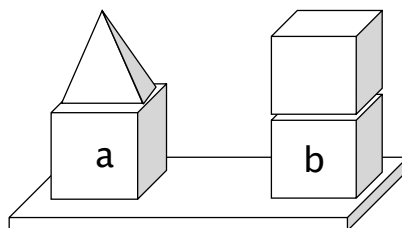
...

Axiom 1.  $\forall x \text{ Cube}(x) \Rightarrow \text{Object}(x)$

Axiom 2.  $\forall x \forall y \text{ Object}(x) \wedge \text{Object}(y) \wedge x \neq y$

$\Rightarrow \text{location}(x) \neq \text{location}(y)$

Axiom 3.  $\forall x \forall y \text{ On}(x, y) \Rightarrow \neg \text{Movable}(y)$



25

## The CyC Project

Build a theory of commonsense, to add AI to all computer programs

- In first order logic
- Currently millions of axioms
- Grouped in coherent “microtheories”: geometry, physics, movement, transport, ...
- Partial reasoning system (computes logical entailment)

(the top level is freely available, <http://www.cyc.com/>)

26

# Information Systems

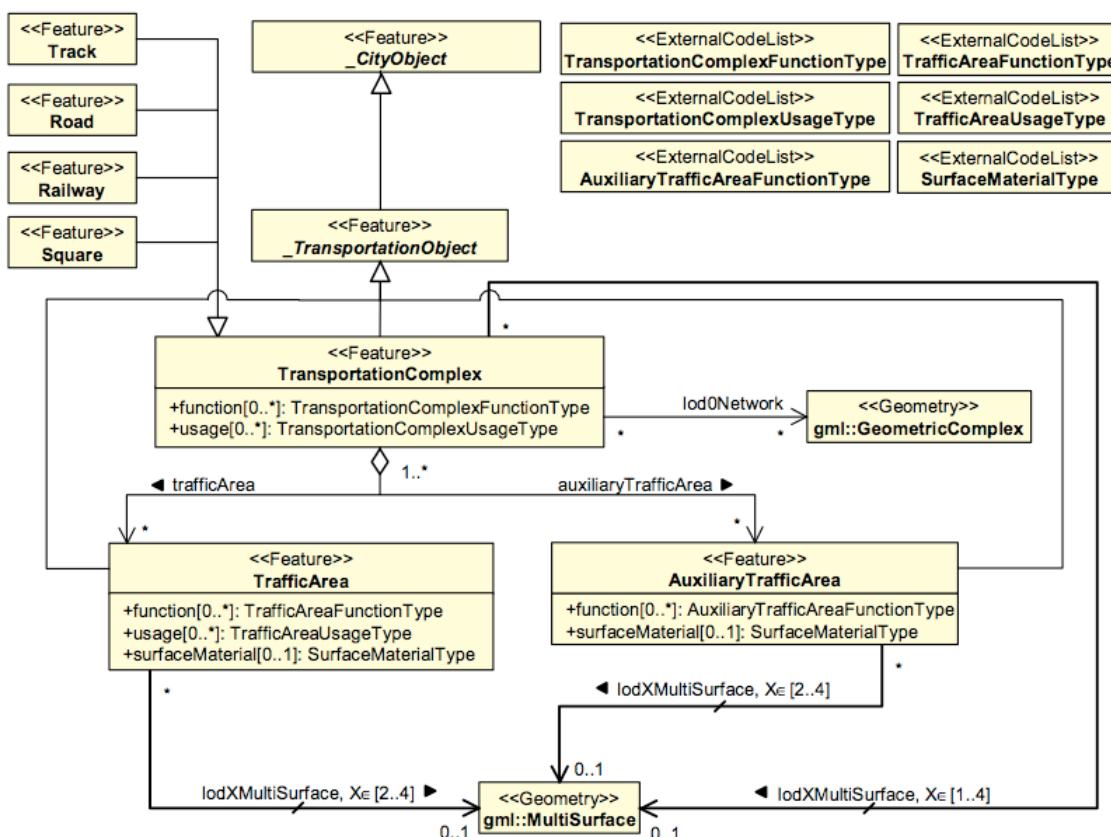
Collect, store, process, retrieve information that is required to manage an organization.

Necessary to know “what exists” in the organization’s domain.

What type of information do we have?

What are the relations between these information types?

- ⇒ Object/Class-based models
- ⇒ the UML standard



# The Semantic Web Initiative

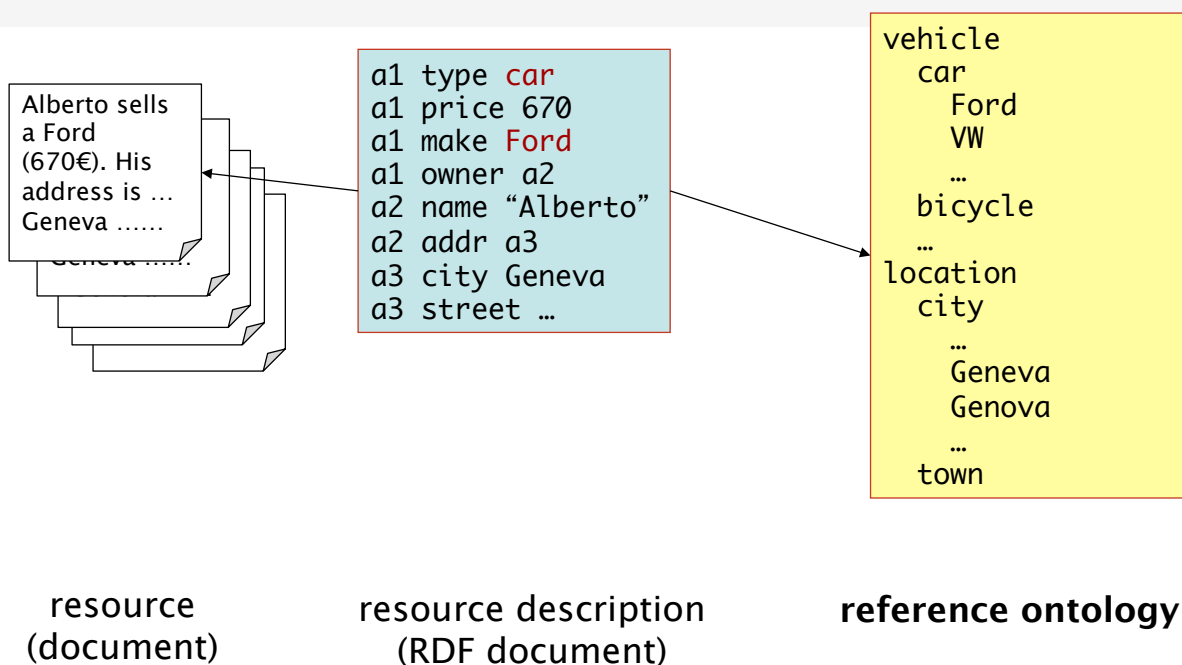
Machines cannot understand natural language

Hence, the web is not machine processable.

- impossible to write a program to “find a German car for sale at a price lower than 1000 €”

Idea: associate a formal representation to each web resource.

# The Semantic Web Initiative



## Ontologies for the Semantic Web

Provide a common vocabulary to describe things and to reason about things on the Web.

Ontology Languages for the SW

**RDFS:** class, subclass, property, domain, range

**OWL:** RDFS + Description Logics