

# Ontology Engineering

How to design complex ontologies

# RDF and OWL

- Not able to model every real-world problem
- How to make an approach?
  - provide methodologies
  - auxiliary tools for development, evaluation and maintenance.

# Requirement Analysis

- semantic representation?
  - easy exchangeable
  - integrate with knowledge from other sources
  - deduct implicit knowledge
- Available tool support
  - commitment to one specific tool?
  - how mature is it?
  - support from the tool vendor?
  - interoperable with other tools?

# Requirement Analysis

- Choosing the semantic language
  - RDF? OWL? OWL DL? OWL2?
  - Depends how complex the domain is
- Choosing the domain-specific primitives:
  - What domain has to be modeled?
  - What is the needed granularity of the specification?
  - What are the tasks to be accomplished with the help of the ontology?

# Ontology Creation

- Where is the knowledge?
  - Human sources
  - Unstructured sources
  - Semistructured sources
  - Structured sources

# Human sources

- A domain expert
- Problem:
  - no knowledge of the representation languages in semantics
- Solution:
  - Knowledge engineer interview the expert and making logical specifications
    - Misunderstandings, information loss
  - Hard to express cases based on clear definitions/rules
    - apply automated techniques

# Unstructured sources

- Written text
- How to parse and extract information from a text like:  
*"Markus does not like animal food. But he ordered a Thai dish that contains fish."*  
to get a structural representation of the grammatical interdependencies (parse tree)
- Syntactic analysis is hard if sentences are ambiguous

# Unstructured sources

- Formalization
  - transform to logical description
  - parse tree of text -> logical statements by applying a set of rules
  - statements should not contain incorrect information
- Background knowledge
  - "fishes are animals"
  - use thesauri to get the knowledge



# Semistructured sources

- Hyperlinks, wiki articles, etc.
- Transferred directly into RDF or OWL representation
- Example: File systems
  - metadata accessible
  - folder structure
  - location of the stored files

# Structured sources

- Database, ontologies
- Reusing parts of existing ontologies
- Database to fill in information about each primitive, cardinality, etc.

# Quality Assurance of Ontologies

- How to improve existing ontologies
- What makes an ontology good?
  - Fulfill its purpose?
- Criteria
  - Logical
    - Modeled correctly?
  - Structural and formal
    - explicit taxonomic cycles?
  - Accuracy
    - capture aspects of the modeled domain it has been designed for?

# How to (not) model correctly

- Don't forget disjointness
  - man/woman, can't be both!
- Don't forget role characteristics
- Don't choose too specific domains or ranges
- Be careful with quantifiers
- Don't mistake parts of subclasses
- Watch the direction of roles
- Don't confuse class subsumption and class equivalence
  - Is every being living in water a fish?
- Don't translate too verbally