Ontology Engineering: How can we build ontologies? Methods, Techniques and Methodologies

6th November 2007

Mari Carmen Suárez-Figueroa
mcsuarez@fi.upm.es
Ontology Engineering Group. Laboratorio de Inteligencia Artificial
Facultad de Informática
Universidad Politécnica de Madrid

Oscar Corcho
ocorcho@fi.upm.es
Ontology Engineering Group
Laboratorio de Inteligencia Artificial
Facultad de Informática
Universidad Politécnica de Madrid

6th November 2007
Main References

Gómez-Pérez, A.; Fernández-López, M.; Corcho, O.  **Ontological Engineering.**  *Springer Verlag.* 2003

http://www.ontoweb.org

- Deliverables
  - D1.4
  - D1.5

http://www.seemp.org

- Research deliverables

http://knowledgeweb.semanticweb.org

- Research deliverables
- Industry deliverables

http://www.neon-project.org

- Research deliverables
Acknowledgements

• Asunción Gómez-Pérez, Mariano Fernández-López, and Boris Villazón
  – Most of the slides have been done jointly with them
• Jeremy Roberts (University of Manchester)
  – Knowledge elicitation techniques
Ontology Definition

“An ontology is a formal, explicit specification of a shared conceptualization”

- Machine-readable
- Concepts, properties, relations, functions, constraints, axioms, are explicitly defined
- Consensual Knowledge
- Abstract model and simplified view of some phenomenon in the world that we want to represent

I want to build my ontology

- Which one are the activities involved in the ontology development process?
- Which one is the goal of each activity?
- When should I carry out each activity?
- What is the relationship of one activity with the others?
- Where can I find ontologies with the goal of reusing them?
- How can I use the ontology in my application?
Ontology Engineering

It refers to the set of activities that concern

the ontology development process,

the ontology life cycle,

the methods and methodologies for building ontologies,

and the tool suites

and languages that support them
Three aspects of ontology development

Development Process: Which activities

Life Cycle: Order of activities
1. Intra-ontology dependencies
2. Inter-ontology dependencies

Methodologies: How to carry out the activities
1. Input and outputs
2. Methods, tasks, techniques, tools
Outline

- Ontology Development Process
- Ontology Development Lifecycle
- Methodologies for Building Ontologies
- An example: An ontology about human resources
The Framework

The world of ontologies

ONTOGRAPHY
Can be public
Define Ontology
(Imported ontologies . . .)

METHODOLOGY
Item 1: It is necessary . . .
Item 2: Since . . .

Tools

• To set up a life cycle
• Development process

Ontologies are available anywhere in Internet
The NeOn Glossary of Activities

Collaboratively Ontology Activity identification and definition
- 53 activity definitions consensuated

On-going Steps:
- Publication in the NeOn website (http://www.neon-project.org)
- Procedure for getting feedback from the community (http://cicero.uni-koblenz.de)
The NeOn Glossary of Activities

- Ontology Adaptation
- Ontology Alignment / Aligning
- Ontology Annotation
- Ontology Articulation
- Ontology Assessment
- Ontology Combining
- Ontology Comparing
- Ontology Conceptualization
- Ontology Configuration Management
- Ontology Coordination
- Ontology Customization
- Ontology Diagnosis
- Ontology Documentation
- Ontology Elicitation
- Ontology Enrichment
- Ontology Evaluation

- Ontology Evolution
- Ontology Extension
- Ontology Formalization
- Ontology Generation
- Ontology Implementation
- Ontology Integration

- Knowledge Acquisition for Ontologies
  - Ontology Learning
  - Ontology Localization
  - Ontology Mapping
  - Ontology Matching
  - Ontology Mediation
  - Ontology Merging
  - Ontology Modification
  - Ontology Modularization
  - Ontology Module Extraction
  - Ontology Partitioning
  - Ontology Personalization
  - Ontology Population
  - Ontology Pruning
  - Ontology Reconciliation
  - Ontology Reengineering
  - Ontology Repair
  - Ontology Reuse

- Ontology Searching
- Ontology Selection
- Ontology Specialization
- Ontology Specification
- Ontology Summarization
- Ontology Transforming
- Ontology Translating
- Ontology Update
- Ontology Upgrade
- Ontology Validation
- Ontology Valuation
- Ontology Verification
- Ontology Versioning
- Scheduling
- Control
- Quality Assurance
- Environment Study
- Feasibility Study
- Reverse Engineering
- Reengineering
- Forward Engineering
The NeOn Glossary of Activities

WP5WorkingArea: Knowledge Acquisition for Ontologies

- **Final Definition**: Knowledge Acquisition for Ontologies comprises activities for capturing knowledge (e.g., T-Box and A-Box) from a variety of sources. We distinguish between: Ontology Elicitation, Ontology Learning and Ontology Population.

WP5WorkingArea: Ontology Elicitation

- **Final Definition**: Ontology Elicitation is a knowledge acquisition activity in which conceptual structures (e.g. T-Box) and their instances (e.g. A-Box) are acquired from domain experts.

WP5WorkingArea: Ontology Learning

- **Final Definition**: Ontology Learning is a knowledge acquisition activity that relies on (semi-) automatic methods to transform unstructured (e.g. corpora), semi-structured (e.g. folksonomies, html pages, etc.) and structured data sources (e.g. data bases) into conceptual structures (e.g. T-Box).

WP5WorkingArea: Ontology Population

(Redirected from WP5WorkingArea: Ontology Population/Grounding)

- **Final Definition**: Ontology Population is a knowledge acquisition activity that relies on (semi-) automatic methods to transform unstructured (e.g. corpora), semi-structured (e.g. folksonomies, html pages, etc) and structured data sources (e.g. data bases) into instance data (e.g. A-Box).
Table of “Required and If-Applicable” Activities

- **Required activities** refer to those activities that should be carried out when developing networks of ontologies.

- **If Applicable activities** refer to those activities that can be carried out or not, depending on the case, when developing ontology networks.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Required</th>
<th>If Applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontology Enrichment</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Ontology Environment Study</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Ontology Evaluation</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Ontology Evolution</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Ontology Extension</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Ontology Feasibility Study</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Ontology Formalization</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Ontology Forward Engineering</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Ontology Implementation</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Ontology Integration</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Knowledge Acquisition for Ontologies</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Ontology Learning</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Ontology Localization</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Ontology Matching</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Ontology Network Development Process

Development-oriented

Pre-Development
- O. Environment Study
- O. Feasibility Study

Development
- O. Translation
- O. Restructuring Resource Reengineering
- O. Reengineering
- O. Customization
- O. Alignment

Post-Development
- O. Upgrade
- O. Versioning
- O. Evolution

Support
- Knowledge Acquisition
- O. Documentation
- O. Summarization
- O. Evaluation (V&V)
- O. Configuration Management

Ontological Engineering
Ontology Life Cycle. Intra-dependencies

Pre-Development
- O. Environment Study
- O. Feasibility Study

Management
- Scheduling

Development-oriented
- Development
- Post-Development

Support
- Knowledge Acquisition
- O. Evaluation (V&V)
- O. Configuration Management
- O. Documentation
- O. Summarization
- O. Assessment

Ontology Engineering
Ontology Life Cycle. Intra-dependencies

Development activities

<table>
<thead>
<tr>
<th>Activity</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specification</td>
<td>Conceptualization</td>
</tr>
</tbody>
</table>

Management activities

<table>
<thead>
<tr>
<th>Activity</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheduling</td>
<td>Control</td>
</tr>
</tbody>
</table>

Support activities

<table>
<thead>
<tr>
<th>Activity</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge acquisition</td>
<td>Integration</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ontology Life Cycle. Inter-dependencies

Inter-dependencies refer the relationship between activities carried out when building different ontologies.

Fernández-López, M.; Gómez-Pérez, A.; Rojas M.D.  
*Ontology’s Crossed Life Cycle.*  
*Lectures Notes in Artificial Intelligence* Nº 1937. October 2000
Outline

• Ontology Development Process

• **Ontology Development Lifecycle**

• Methodologies for Building Ontologies

• An example: An ontology about human resources
Life Cycle Models and Life Cycles in Ontological Engineering

• An **ontology (network) life cycle model** is the framework (waterfall, evolving prototyping, spiral, etc.), selected by each organization, on which to map the activities identified in the ontology development process.

• The **ontology (network) life cycle** is the project specific sequence of activities, created by mapping the activities identified in the ontology development process onto a selected ontology life cycle model.

Example: three versions of the waterfall ontology network life cycle model
Several Ontology Life Cycle Models are possible

There is **no a unique life cycle model** valid for all the ontology development projects and that each life cycle model is appropriate for a concrete project, depending on several features.

For example, sometimes it is better a simple one (like waterfall), whereas other times it is most suitable a spiral one (if the analysis of the risk is needed within the project).

- **Assumption**: Ontology requirements are known at the begining of the ontology development project.

- **Assumption**: Ontology requirements can be not known at the begining of the ontology development project and can change during the project.
Several Ontology Life Cycle Models are possible

- Assumption: Uncertainties in the ontology requirements can derive into risks in the project.

Risks can be:
- Properties became classes
- Move from frames to DL
- Reuse new existing resources

- **Planning**: in this phase it is carried out the whole schedule for the ontology network development and the specification of the ontology network requirements.

- **Risk analysis**: after analysing the possible risk within the ontology network development, the decision of continuing or not with a new iteration around the spiral is taken.

- **Engineering**: in this phase it is developed a prototype of the ontology network based on the specified requirements, following any type of waterfall ontology network life cycle model.
How software developers and ontology practitioners decide which **ontology network life cycle model** is the most appropriate for their ontology network and which **concrete activities** should be carried out in their ontology network life cycle?

**Proposed steps:**

1. Identify ontology network development requirements.
2. Select the ontology network life cycle model (ONLCM) to be used.
3. Select activities to be carried out from the “Required—if Applicable” table.
4. Map the selected activities into the selected ontology network life cycle model.
5. Set the order of the activities: the result is the ontology network life cycle for the ontology network.
Step 2: Decision Tree for Selecting the Ontology Network Life Cycle Model

Do you think that ontology network requirements will change during the development?
- No
- Yes

Do you want to produce intermediate results?
- Yes
- No

Do you want to produce each intermediate result in a complete manner?
- Yes
- No

Do you want to include risk control in your ontology network development?
- Yes
- No

Waterfall
- Iterative
- Incremental

Evolving Prototyping
Spiral
Step 3: Decision Tree for Selecting Activities to be mapped in the Ontology Network Life Cycle Model

Have you developed more than 5 ontologies?

- No
- Yes

Set of “yes/no” natural language questions for identifying the ‘if-applicable’ activities to be carried out.

➢ Do you want to have your ontology network in different natural languages, as for example, in English, Spanish and French? YES → O. Localization.

➢ Do you want to take an existing and implemented ontology, in order to enhance it and implement it again? NO → O. Reengineering is not selected.

Software developers and ontology practitioners select the activities to be carried out from the “Required-If Applicable” table

<table>
<thead>
<tr>
<th>Activity</th>
<th>If Applicable</th>
<th>Selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontology Aligning</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ontology Customization</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ontology Learning</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ontology Localization</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ontology Matching</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ontology Modification</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ontology Reengineering</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Ontology Restructuring</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Automatically
Set of “yes/no” natural language questions for identifying the ‘if-applicable’ activities to be carried out.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Natural Language Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ontology Aligning</strong></td>
<td>Do you have two or more ontologies at your disposal that you want to examine to find correspondences and to take advantage of them?</td>
</tr>
<tr>
<td></td>
<td>Do you want to find out correspondences among ontologies to use them?</td>
</tr>
<tr>
<td><strong>Ontology Customization</strong></td>
<td>Do you want to adapt the ontology network to a specific user profile?</td>
</tr>
<tr>
<td></td>
<td>Do you want to modify the ontology network to meet specific user needs?</td>
</tr>
<tr>
<td><strong>Ontology Enrichment</strong></td>
<td>Do you want to widen/extend your current ontology network with additional elements (e.g., concepts, roles, axioms, etc.)?</td>
</tr>
<tr>
<td><strong>Ontology Extension</strong></td>
<td>Do you want to stretch, widen, broaden or expand your current ontology network by adding new concepts &quot;in a horizontal way/direction&quot; with the aim of widening its sphere of action?</td>
</tr>
<tr>
<td></td>
<td>(cf. Ontology Specialization)</td>
</tr>
<tr>
<td><strong>Ontology Forward Engineering</strong></td>
<td>Are you going to carry out a new implementation for a previously modified conceptual model?</td>
</tr>
<tr>
<td></td>
<td>Are you going to produce a new implementation for a modified conceptual model, whose previous version had already been implemented?</td>
</tr>
</tbody>
</table>
Scenarios for Building Ontology Networks

1. Specification
   - Ontological Resource Reuse
   - Ontology Design Pattern Reuse

2. Conceptualization
   - Non Ontological Resource Reuse
   - Ontology Design Pattern Reuse

3. Formalization
   - Ontology Reengineering
   - Ontology Repositories and Registries

4. Implementation
   - O. Aligning
   - O. Merging

5. Localization
   - Ontology Support Activities: Knowledge Acquisition (Elicitation), Documentation, Configuration Management, Evaluation (V&V), Assessment

6. Scenario 1: Ontological Resource Reuse
7. Scenario 2: Ontology Design Pattern Reuse
8. Scenario 3: Ontology Reengineering
9. Scenario 4: Ontology Repositories and Registries

Resources:
- Glossaries
- Lexicons
- Thesauri
- Ontology Design Patterns
- O. Repositories and Registries

Alignments:
- RDF(S)
- Flogic
- OWL

Ontological Engineering
Outline

• Ontology Development Process
• Ontology Development Lifecycle
• Methodologies for building ontologies
• An example: An ontology about human resources
Most relevant methodologies

- Cyc method
- Uschold and King’s method
- Grüninger and Fox’s methodology
- KACTUS approach
- METHONTOLOGY
- SENSUS method
- On-To-Knowledge
- DILIGENT

None deals with the three dimensions simultaneously

NeOn Methodology V1 will be available in February 2008
Uschold’s Methodology

1. Identify Purpose and Scope

2. Building the ontology
   - Ontology Capture
   - Ontology Coding
   - Integrating existing ontologies

   - Identify key concepts and relationships
   - Produce unambiguous text definitions
   - Identify terms to refer to such concepts and relations
   - Commit to a meta-ontology
   - Choose a representation language
   - Write the code

   How and whether to reuse ontologies that already exist

3. Evaluation

4. Documentation

TOVE Methodology

Motivating Scenarios → Informal Competency Questions → Formal Terminology → Formal CQ → Formal Axioms → Completeness Theorems

Identify intuitively possible applications and solutions

As an entailment of consistency problems with respect to the axioms in the ontology

Conditions under which the solutions to the questions are complete

Defined as a first-order sentence using the predicates of the ontology

Objects → Attributes → Relations

KIF

Constants
Variables

Functions
Predicates

Identify Queries:
- Answers: Axioms
  Formal definitions
- Questions: Terminology

Uschold, M.; Grüninger, M.
ONTOLOGIES: Principles, Methods and Applications.
Knowledge Engineering Review.
Vol. 11, N. 2, June 1996.
Methodology used on the KACTUS project

A bottom-up approach for building ontologies
Build a preliminary ontology for refinement and augment with new definitions

- Specification of the application
- Preliminary design based on relevant top-level ontological categories
- Domain Ontology

Redefine

---

SENSUS Method

Linking Domain Specific Terms to a broad Coverage Ontology

To identify the terms in SENSUS that are relevant to a particular domain and then prune the skeletal ontology using heuristics.

SANSUS method (II)

METHOD

1. Identify “seed” terms
2. Link seed terms to SENSUS by hand
3. Include nodes on the path to root
4. Add entire subtrees using the heuristic:
   If many nodes in a subtree are relevant, the other nodes in the subtree are relevant

SENSUS method (III)

Example: Travel Domain

SENSUS ontology

- Europe – Africa flight
  - seed term
- Europe – America flight
  - seed term
- London - Liverpool flight
  - seed term
- Madrid - Barcelona flight
  - seed term

Is hyponym

international flight
domestic flight
Ontological Engineering 36

SENSUS method (IV)

OB - THING

PROCESS
MATERIAL PROCESS
NON - DIRECTED - ACTION
MOTION - PROCESS
change of location, move
travelling
journeying
trip < journey

flight trip

international flight

Europe – Africa flight
seed term

Europe – America flight
seed term

London - Liverpool flight
seed term

Madrid - Barcelona flight
seed term

OBJECT
NON - CONSCIOUS - THIN
SPATIAL TEMPORAL
SPACE INTERVAL
point < location
root < point
goal < point

Is hyponym

SENSUS method (IV)
On-To-Knowledge

Proposes to build the ontology taking into account how the ontology will be used in further applications.

Feasibility study

- Identify problem and opportunity areas
- Select most promising focus area and target solution

Ontology Kickoff

- Requirement specification
- Analyze input sources
- Develop baseline taxonomy

Refinement

- Concept elicitation with domain experts
- Develop baseline taxonomy
- Conceptualize and formalize
- Add relations and axioms

Evaluation

- Identify problem and opportunity areas
- Select most promising focus area and target solution

Maintenance

- Manage organizational maintenance process
The identification of the ontology development process, a life cycle based on evolving prototypes, and methods and techniques to carry out in different activities.
## Summary of the Methodologies

With respect to the activities in the ontology development process

<table>
<thead>
<tr>
<th>Feature</th>
<th>Cyc</th>
<th>Uschold &amp; King</th>
<th>Grüniger &amp; Fox</th>
<th>KACTUS</th>
<th>METHONTOLOGY</th>
<th>SENSUS</th>
<th>On-To-Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ontology management activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scheduling</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Proposed</td>
<td>NP</td>
<td>Described</td>
</tr>
<tr>
<td>Control</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Proposed</td>
<td>NP</td>
<td>Described</td>
</tr>
<tr>
<td>Quality assurance</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Described</td>
</tr>
<tr>
<td><strong>Ontology development-oriented activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-development processes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environment study</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Proposed</td>
<td>NP</td>
<td>Proposed</td>
</tr>
<tr>
<td>Feasibility study</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Proposed</td>
<td>NP</td>
<td>Proposed</td>
</tr>
<tr>
<td>Development processes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specification</td>
<td>NP</td>
<td>Proposed</td>
<td>Described in detail</td>
<td>Proposed</td>
<td>Described in detail</td>
<td>NP</td>
<td>Proposed</td>
</tr>
<tr>
<td>Conceptualization</td>
<td>NP</td>
<td>NP</td>
<td>Described in detail</td>
<td>Proposed</td>
<td>Described in detail</td>
<td>NP</td>
<td>Proposed</td>
</tr>
<tr>
<td>Formalization</td>
<td>NP</td>
<td>Described in detail</td>
<td>Described in detail</td>
<td>Described in detail</td>
<td>Described in detail</td>
<td>NP</td>
<td>Described</td>
</tr>
<tr>
<td>Implementation</td>
<td>Proposed</td>
<td>Proposed</td>
<td>Described in detail</td>
<td>Proposed</td>
<td>Described in detail</td>
<td>Described</td>
<td>Described</td>
</tr>
<tr>
<td><strong>Post-development processes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Proposed</td>
<td>NP</td>
<td>Proposed</td>
</tr>
<tr>
<td>Use</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Proposed</td>
<td>NP</td>
<td>Proposed</td>
</tr>
<tr>
<td><strong>Ontology support activities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knowledge acquisition</td>
<td>Proposed</td>
<td>Proposed</td>
<td>Proposed</td>
<td>NP</td>
<td>Described in detail</td>
<td>NP</td>
<td>Described</td>
</tr>
<tr>
<td>Evaluation</td>
<td>NP</td>
<td>Proposed</td>
<td>Described in detail</td>
<td>NP</td>
<td>Described in detail</td>
<td>NP</td>
<td>Proposed</td>
</tr>
<tr>
<td>Integration</td>
<td>Proposed</td>
<td>Proposed</td>
<td>Proposed</td>
<td>Proposed</td>
<td>Proposed</td>
<td>NP</td>
<td>Proposed</td>
</tr>
<tr>
<td>Configuration management</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>Described</td>
<td>NP</td>
<td>Proposed</td>
</tr>
<tr>
<td>Documentation</td>
<td>Proposed</td>
<td>Proposed</td>
<td>Proposed</td>
<td>NP</td>
<td>Described in detail</td>
<td>NP</td>
<td>Described</td>
</tr>
<tr>
<td>Merging and Alignment</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
</tbody>
</table>
Outline

• Ontology Development Process
• Ontology Development Lifecycle
• Methodologies for building ontologies
• An example: An ontology about human resources
Looking for an European Employment

LEGENDA

ES: Employment Service
Cand.: Job Seeker’s Candidacy
Vacan.: Employer Job Vacancy

Ontological Engineering
The Goal: Helping Job Seekers on their way

- **Requester ES**
- **Responding ES**
- **ES not involved**
- **Job Seeker’s Candidacy**
- **Employer Job Vacancy**

**European Employment Mediators Marketplace**

- **Lombard ES (It)**
- **Catalonia ES (Es)**
- **Wallonia ES (Be)**

**Local Matching algorithm**
1. Build a reference ontology
2. Build mappings between the reference ontology and the data sources

Build a reference ontology for the domain
Build local ontologies
Build mappings between the core and local ontologies
Build mappings between the local ontologies and the data sources
Build a reference ontology for the domain. This Reference Ontology is the core semantic component of the system. It acts as a common ‘language’ in the form of a set of controlled vocabularies to describe the details of a job posting and the CV of a job seeker.

Build local ontologies. Each ES uses its own Local Ontology, which describes the employment market in its own terms.

Build mappings between the local ontologies and the data sources (ES schema).
Ontological Engineering
SEEMP Ontology Network Life Cycle: Iterative model life cycle

- O. Elicitation
- O. Documentation
- O. Specification
- O. Localization
- O. Conceptualization
- O. Selection
- O. Searching
- O. Assessment
- O. prung
- O. Specialization
- O. Extension
- O. Evaluation
- O. Implementation
- O. Formalization
- O. Assessment
- Maintain
- Use

Repositories & libraries
RDF(S) OWL

ISC-88 (COM), ONET, EURES taxonomy, FOET, ISCED97, NACE, ISO 4217, ISO 3166, ISO 6392, HR-XML, …
Ontology Specification

Repositories & libraries

- RDF(S)
- OWL

O. Elicitation
O. Documentation
O. Localization
O. Pruning
O. Extension
O. Specialization
O. Conceptualization
O. Evaluation
O. Formalization
O. Implementation
O. Searching
O. Selection
O. Assessment
O. Reengineering

O. Specification

Search
Assessment
Select
Reengineering

ISCO-88 (COM), ONET, EURES taxonomy, FOET, ISCED97, NACE, ISO 4217, ISO 3166, ISO 6392, HR-XML, …
60 Competency Questions grouped into 5 categories (modular approach)

- Job Seeker (12)
  - What is his/her education level?
- Job Offer (12)
  - What are the required skills for the job offer?
- Time and date management (7)
  - When the job seeker completed his/her first degree?
- Currencies (4)
  - The offered salary is given in US dollars?
- General (25)
  - Given the employer information, economic activity of the employer and the job offer profile (job, contract type, salary, work condition, contract duration), what job seekers are the most appropriate?

Given the job offer profile (job, contract type, salary, work condition) and the required profile to seek (required education level, required work experience, required knowledge, required skills), what job seekers are the most appropriate?

Each organization has job offers for job seekers

Vocabulary:
Questions: contract type, salary, work condition, job seeker, job offer, ...
Answers: autonomous, 3000 euro, holiday job, ...

Classes: Contract Type, Compensation, Work Condition, Job Seeker, Job Offer ...

Relations: has job category, has compensation, requires work experience ...
Attributes: Name, date of birth, email ...
Non Ontological Resource Reuse

Repositories & libraries
RDF(S) OWL

O. Elicitation
O. Documentation
O. Locализation
O. Pruning
O. Reengineering
O. Conceptualization
O. Evaluation
O. Formalization
O. Implementation
O. Searching
O. Selection
O. Assessment
O. Use

ISC-88 (COM), ONET, EURES taxonomy, FOET, ISCED97, NACE, ISO 4217, ISO 3166, ISO 6392, HR-XML, …

Non Ontological Resource Reuse
Assessment
Select
Search
Reengineer

Ontological Engineering
Search and Assess Standards and Taxonomies

- We select the most appropriate standards and taxonomies for:

  - Occupation Classification
    ISCO-88 (COM), SOC, ISCO-88, ONET, Eures Taxonomy.
  - Classification of Economic Activities
    ISIC Rev. 3.1, NACE Rev. 1.1, NAICS
  - Apprenticeship classifications
    ISCED 97, FOET
  - Currency Classification
    ISO 4217
  - Geography Classification
    ISO 3166, Eures Taxonomy

- The IDABC\(^1\) identifies as one of the successful factors at facilitating the development of pan-European interoperable information systems:
  - “Identify, reuse and extend existing assets (taxonomies, thesauri, etc.)”

Language Classification
ISO 6392, CEF
Driving License Classification
European Legislation
Skill Classification
Eures Taxonomy
Contract Types Classification
LE FOREM, Eures and BLL Classification
Work Condition Classification
LE FOREM, Eures and BLL Classification

Assessment activity:
Matching terminology from Competency Questions against the Standards

(1) -> IDABC Content Interoperability Strategy. Working paper. Sep 2005

50
Ontological Engineering
Selection of Human Resources Management Standards

Reference Ontology shall be based on the international, European or de-facto industrial standards.

<table>
<thead>
<tr>
<th>Occupation Classification</th>
<th>Classification of Economic Activities</th>
<th>Apprenticeship Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOC</td>
<td>ISCO-88 (COM)</td>
<td>ISIC Rev. 3.1</td>
</tr>
<tr>
<td>ONET</td>
<td>ISIC Rev. 3.1</td>
<td>NACE Rev. 1.1</td>
</tr>
<tr>
<td>ISCO-88</td>
<td>NACE Rev. 1.1</td>
<td>NAICS</td>
</tr>
<tr>
<td>The degree of coverage</td>
<td></td>
<td>ISCED 97</td>
</tr>
<tr>
<td>The current European needs</td>
<td></td>
<td>FOET</td>
</tr>
</tbody>
</table>

But, we need also proprietary taxonomies …
Knowledge Resource Reengineering

- EURES Taxonomy (proprietary) - Oracle DB
- ONET - HTML
- ISCO-88 (COM) - MS Access

Integrate

Occupation Ontology

Extend
Specialize
Prune
Ad hoc wrapper
WSML exporter

Search
Assessment
Reengineer
Ontology Searching in Ontology Metadata Repositories

Ontology to describe ontology metadata information

- OMV – Ontology Metadata Vocabulary (http://ontoware.org/projects/omv)
- Knowledge Zone vocabulary (http://tinyurl.com/qfp2s)

4 Ontology Metadata Repositories

- Oyster (P2P system, http://oyster.ontoware.org)
- ONTHOLOGY.org (centralized, http://www.ontology.org/)
- Knowledge Zone (centralized, http://smiprotege.stanford.edu:8080/KnowledgeZone/)
- Swoogle (http://swoogle.umbc.edu)
Searching ontologies: Obtain the set of candidate ontologies using Oyster

Integration of Results

Entry Details
Process for Assessing Time Ontologies (I)

1. Identification of criteria for comparing the candidate set of temporal ontologies

- Time Points
- Time Interval
- Absolute and Relative Time
- Relations between time intervals
- Convex and non convex intervals

2. Assess all existing temporal ontologies against the criteria

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Cyc’s Upper Ontology</th>
<th>Unrestricted Time Ontology</th>
<th>Simple Time Ontology</th>
<th>Reusable Time Ontology</th>
<th>Kestrel Time Ontology</th>
<th>SRI’s Time Ontology</th>
<th>SUMO Time Ontology</th>
<th>DAML Time Ontology</th>
<th>AKT Time Ontology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Points</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Time Interval</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absolute and Relative Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relations between time intervals</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Convex and non convex intervals</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distinction between open and closed intervals</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Explicit modeling of proper intervals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concatenation of intervals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Different temporal granularities</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

General criteria for the time domain

Provides axioms: ✓
Process for **Assessing Time Ontologies** (II)

3. Checking which temporal properties are needed for answering the Competency Questions (identified in the Ontology Specification activity)
   
   a. When the job seeker completed his/her first degree?
   b. Is the job seeker older than 30 years?
   c. How much time did the job seeker spend completing his/her first degree?
   d. How long is the duration of the contract?
   e. Which job offers were posted in last 24 hours?
   f. Which job offers were posted in last 7 days?
   g. Which job offers were posted in last month?
   h. Was the job seeker unemployed?
   i. Was the job seeker a student between 1995 and 2000?
### The Time Ontology Selection

<table>
<thead>
<tr>
<th>Feature</th>
<th>Cyc’s Upper Ontology</th>
<th>Unrestricted Time Ontology</th>
<th>Simple Time Ontology</th>
<th>Reusable Time Ontology</th>
<th>Kestrel Time Ontology</th>
<th>SRI’s Time Ontology</th>
<th>SUMO Time Ontology</th>
<th>DAML Time Ontology</th>
<th>AKT Time Ontology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Points</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Time Interval</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Absolute and Relative Time</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Relations between time intervals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convex and non convex intervals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Distinction between open and closed intervals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Explicit modeling of proper intervals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Concatenation of intervals</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Different temporal granularities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Provides axioms</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Ontological Engineering

Conceptualization

Repositories & libraries
RDF(S) OWL

O. Conceptualization
O. Elicitation
O. Documentation
O. Localization
O. Pruning
O. Extension
O. Specialization
O. Evaluation
O. Implementation
O. Formalization
O. Searching
O. Assessment
O. Selection
Reengineer
O. Specification
ISO 88 (COM), ONET, EURES taxonomy, FOET, ISCED97, NACE, ISO 4217, ISO 3166, ISO 6392, HR-XML, …

O. Assessment
Maintain
Use

Assessment
Select
Search
Reengineer

ISO 3166, ISO 6392, HR-XML, …
Conceptualization: Modular approach for ontology construction

Reusability

Application Domain O.: Job Seeker, Job Offer

Domain O.: Economic Activity, Occupation, Education, Skill, Driving License, Compensation, Labour Regulatory, Competence

General/Common Ontologies: Time, Geography, Language

Representation Ontology: WSML

Usability
ISC-O-88 (COM), ONET, EURES taxonomy, FOET, ISCED97, NACE, ISO 4217, ISO 3166, ISO 6392, HR-XML, ...

Repositories & libraries

Building Reference Ontology

Reference Ontology

Building Local Ontologies

ES Data Sources

Ontological Engineering
Local Ontologies Building Process

- **Option 1:** *Building Local Ontologies from the Reference Ontology.*
  
  ![Diagram of reference ontology to resultant ontology](image)

  - Specialize
  - Extend
  - Prune

- **Option 2:** *Building Local Ontologies as a reengineering process from ES Data Sources.*
  
  ![Diagram of ES data sources to resultant ontology](image)

  - Reengineering
### Which option is the most appropriate for the use case?

<table>
<thead>
<tr>
<th>Mappings between Local Ontologies and Reference Ontology</th>
<th>Option 1: Building Local Ontologies from the Reference Ontology</th>
<th>Option 2: Building Local Ontologies as a reverse engineering process from ES Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mappings are not complex. They use the same terms.</td>
<td>Complex mappings due to terminology heterogeneity.</td>
<td></td>
</tr>
<tr>
<td>Mappings between Local Ontologies and ES schema sources</td>
<td>Complex mappings due to terminology and structural heterogeneity.</td>
<td>Mappings are not complex. They use the same terms.</td>
</tr>
<tr>
<td>Building process</td>
<td>Structured/guided by the architecture of the Reference Ontology and scoped with applications needs.</td>
<td>Requires more sophistication of knowledge engineering and good acquaintance of all the data and their structures of the application.</td>
</tr>
</tbody>
</table>
| Changes in the Reference Ontology                      | Imply changes in  
· the mappings between local and reference ontologies.  
· the mappings between the local ontologies and the ES schema sources.  
· the Local Ontology. | Imply changes in  
· the mappings between Local Ontologies and the Reference Ontology. |
| Changes in the ES schema sources                        | Imply changes in  
· its Local Ontology (probably the part that is not a mirror of the Reference Ontology).  
· the mappings between Local Ontologies and ES schema sources.  
· in the mappings between Local Ontology and the Reference Ontology. | Imply changes in  
· in mappings between ES sources and Local Ontologies.  
· mappings between local and the Reference Ontology. |
Approach followed by SEEMP for building Local Ontologies

A hybrid approach

- Option 1 for Job Seeker and Job Offer Ontologies
- Option 2 for Occupation, Education, etc.
Repositories & libraries

ISC0-88 (COM), ONET, EURES taxonomy, FOET, ISCED97, NACE, ISO 4217, ISO 3166, ISO 6392, HR-XML, ...

Building Reference Ontology

Building Local Ontologies

ES Data Sources

Building Mappings L.O. - ES Data Sources (ODEMapster)

Local Ontologies

Building Mappings R.O. - L.O.

Mappings R.O. - L.O.

Mappings ES-LO

ODEMapster Mapping Editor

WebODE

WebODE

WSMT
Ontology Engineering: How can we build ontologies? Methods, Techniques and Methodologies

6th November 2007

Mari Carmen Suárez-Figueroa
mcsuarez@fi.upm.es
Ontology Engineering Group. Laboratorio de Inteligencia Artificial
Facultad de Informática
Universidad Politécnica de Madrid

Oscar Corcho
ocorcho@fi.upm.es
Ontology Engineering Group
Laboratorio de Inteligencia Artificial
Facultad de Informática
Universidad Politécnica de Madrid

6th November 2007