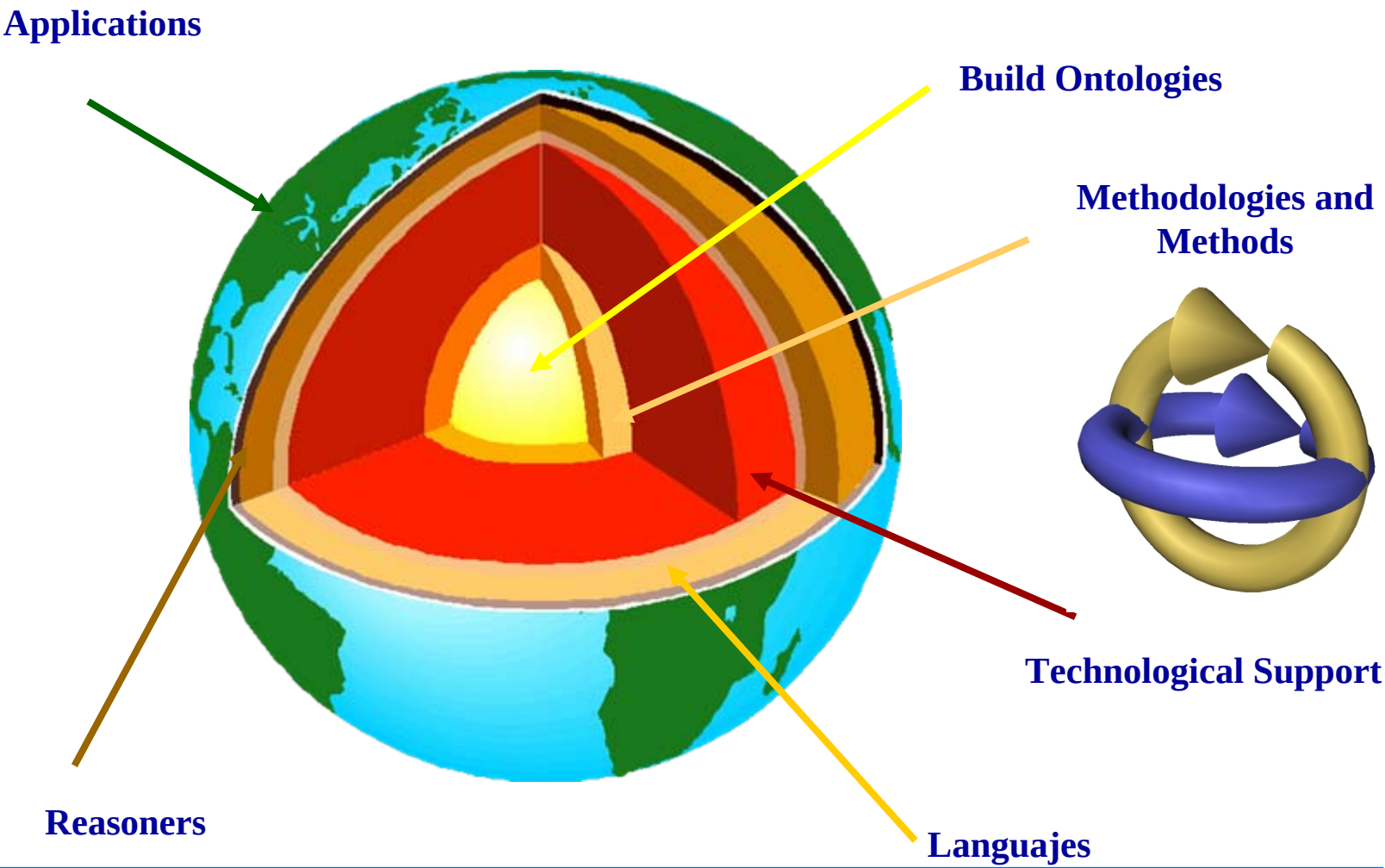
Ontological Engineering: Methodologies

Asunción Gómez-Pérez
Mariano Fernández-López
Oscar Corcho
Mari Carmen Suarez de Figueroa Baonza
{asun, mfernandez, ocorcho,mcsuarez}@fi.upm.es
Grupo de Ontologías
Laboratorio de Inteligencia Artificial
Facultad de Informática
Universidad Politécnica de Madrid
Campus de Montegancedo sn,
28660 Boadilla del Monte, Madrid, Spain

Main Components



Methodologies for building ontologies (I)

Methodologies for building ontologies from the scratch.

Cyc methodology	**URL: http://www.cyc.com**
Uschold and King	**URL: *Not available***
Grüninger and Fox	**URL: *Not available***
KACTUS methodology	**URL: *Not available***
METHONTOLOGY	**URL: *Not available***
SENSUS methodology	**URL: *Not available***
On-To-Knowledge Methodology	**URL: http://www.ontoknowledge.org/**

Methodologies for reengineering ontologies

Method for reengineering ontologies integrated in Methontology **URL: *Not available***

Methodologies for cooperative construction of ontologies

CO4 methodology	**URL: *Not available***
$(KA)^2$ methodology	**URL: *Not available***

Methodologies for building ontologies (II)

Ontology learning methodologies

Aussenac-Gille's and colleagues methodology	**URL: http://www.biomath.jussieu.fr/TIA/**
Maedche and colleagues' methodology	**URL: *Not available***

Ontology merge methodologies

FCA-merge	**URL: *Not available***
PROMPT	**URL: *Not available***
ONIONS	**URL: *Not available***

Ontology evaluation methods

OntoClean: Guarino's group methodology	**URL: *Not available***
Gómez Pérez's evaluation methodology	**URL: *Not available***

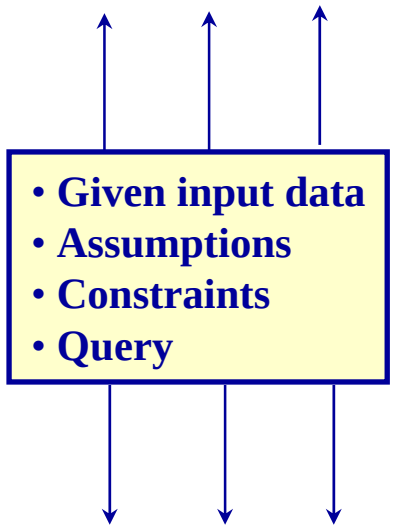
Most relevant methodologies

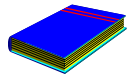
- **Cyc method**
- **Uschold and King's method**
- **Grüninger and Fox's methodology**
- **KACTUS approach**
- **METHONTOLOGY**
- **SENSUS method**
- **On-To-Knowledge**
- **DILIGENT**
- **NeOn Methodology**

TOVE Methodology

Stratification of competency questions

How do we use the solution to the question?
Composition

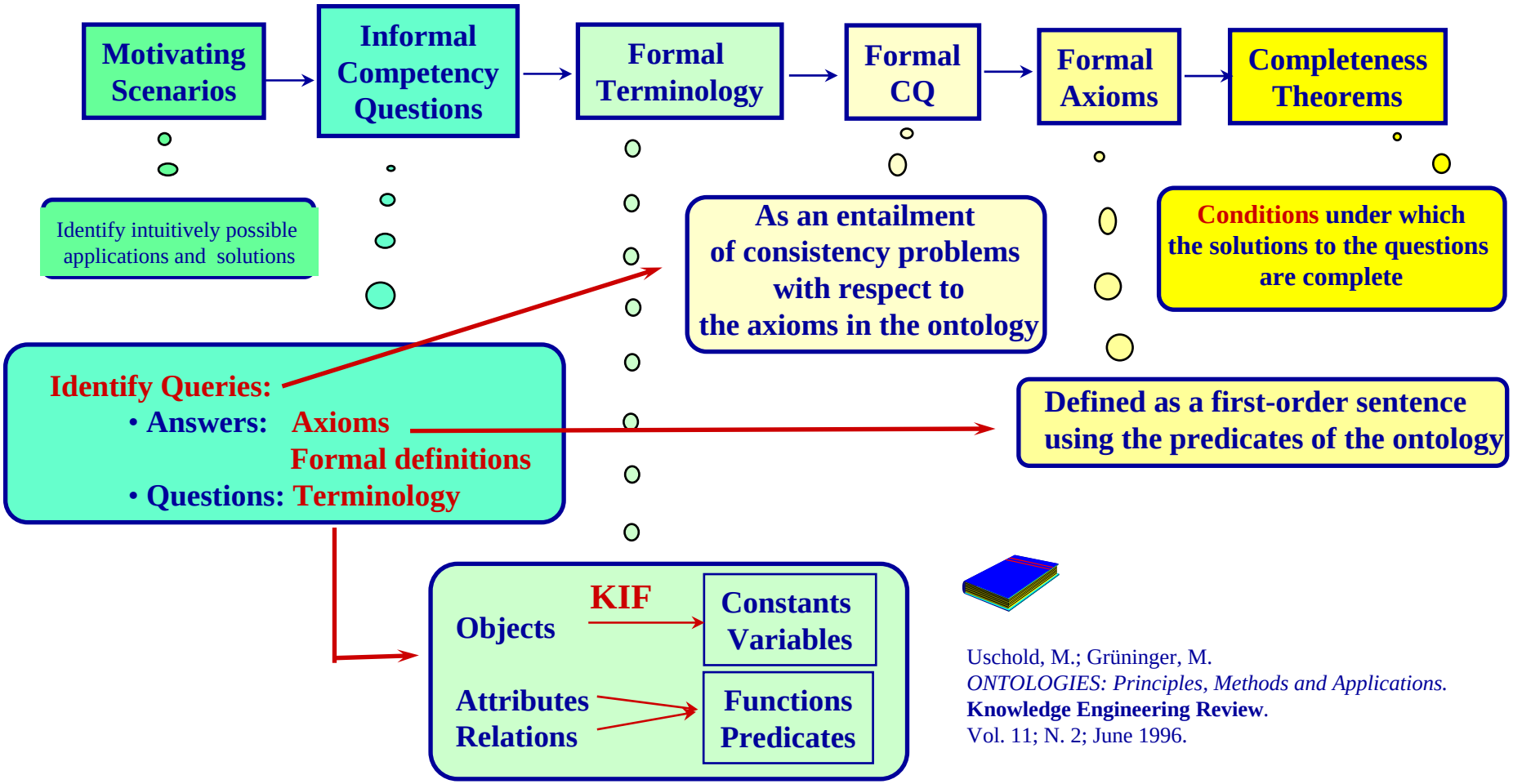




Uschold, M.; Grüninger, M.
ONTOLOGIES: Principles, Methods and Applications.
Knowledge Engineering Review. Vol. 11; N. 2; June 1996.

What do we need to know in order to answer the question?
Decomposition

TOVE Methodology



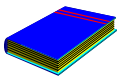
Uschold, M.; Grüninger, M.
ONTOLOGIES: Principles, Methods and Applications.
Knowledge Engineering Review.
Vol. 11; N. 2; June 1996.

Getting terminology using Competency Questions

Motivating Scenarios → **Informal Competency Questions** → **Formal Terminology**

Identify intuitively possible applications and solutions

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



Classes
Relations
Attributes
Axioms
Instances

Uschold, M.; Grüninger, M.
ONTOLOGIES: Principles, Methods and Applications.
Knowledge Engineering Review.
Vol. 11; N. 2; June 1996.

Find documents written by Person P

Identify Queries:
- **Questions: Document, Person, writes**
- **Answers: Document D1 is written by P1**

Classes: Document, Person
Relations: Writes, written by
Attributes: ---
Axioms
Instances: P1, S1

Enterprise Methodology

1. Identify Purpose and Scope

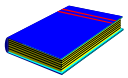
2. Building the ontology

   - Ontology Capture → • Identify key concepts and relationships • Produce unambiguous text definitions • Identify terms to refer to such concepts and relations
   - Ontology Coding → • Commit to a meta-ontology • Choose a representation language • Write the code
   - Integrating existing ontologies → How and whether to reuse ontologies that already exist

3. Evaluation

4. Documentation

5. Guideliness for each phase

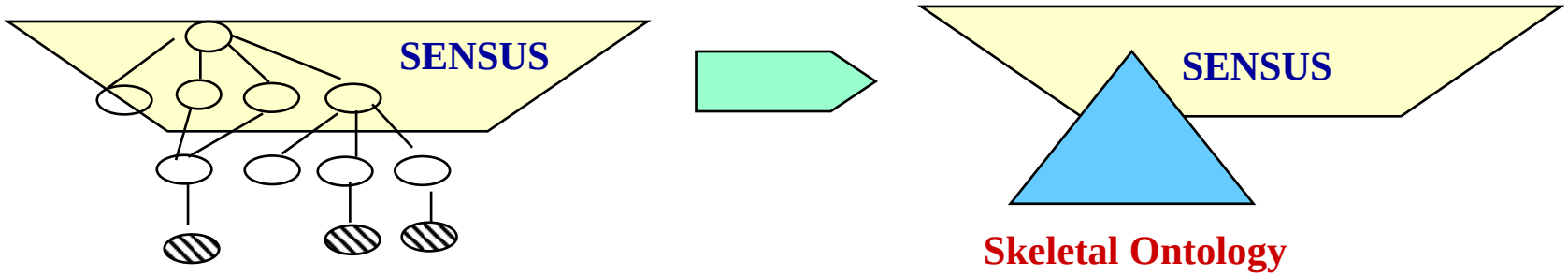


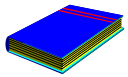
Uschold, M.; Grüninger, M. *ONTOLOGIES: Principles, Methods and Applications.* **Knowledge Engineering Review**. Vol. 11; N. 2; June 1996.

SENSUS as a basis for a domain-specific ontology (I)

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





B. Swartout; R. Patil; k. Knight; T. Russ. *Toward Distributed Use of Large-Scale Ontologies* **Ontological Engineering.** AAAI-97 Spring Symposium Series. 1997. 138-148.

SENSUS as a basis for a domain-specific ontology (II)

see example

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

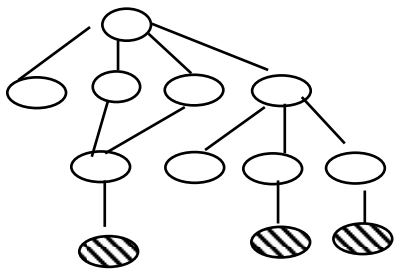

Sensus Term

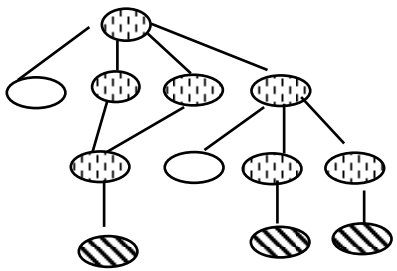

Seed

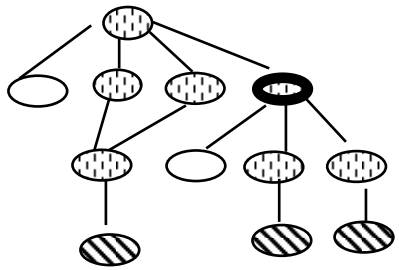

Path to root

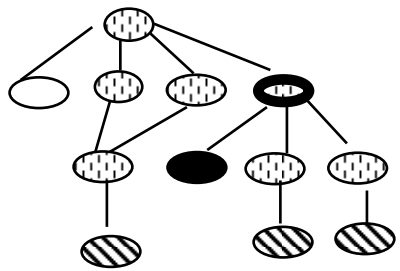

Frequent Parent

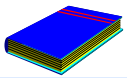

Subtree Term







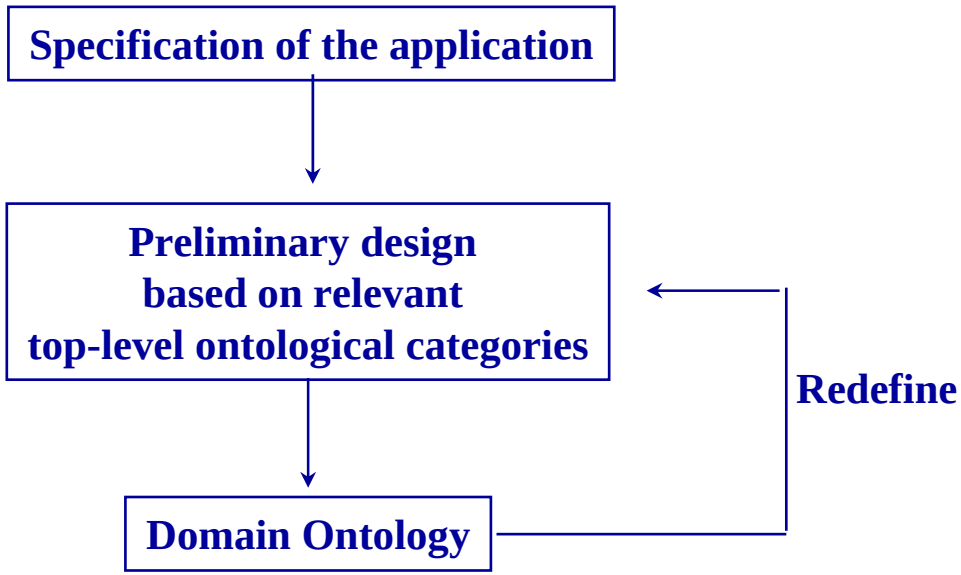


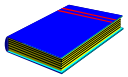


B. Swartout; R. Patil; k. Knight; T. Russ. *Toward Distributed Use of Large-Scale Ontologies* **Ontological Engineering.** AAAI-97 Spring Symposium Series. 1997. 138-148.

Bernara, Laresgoiti, Corera Methodology

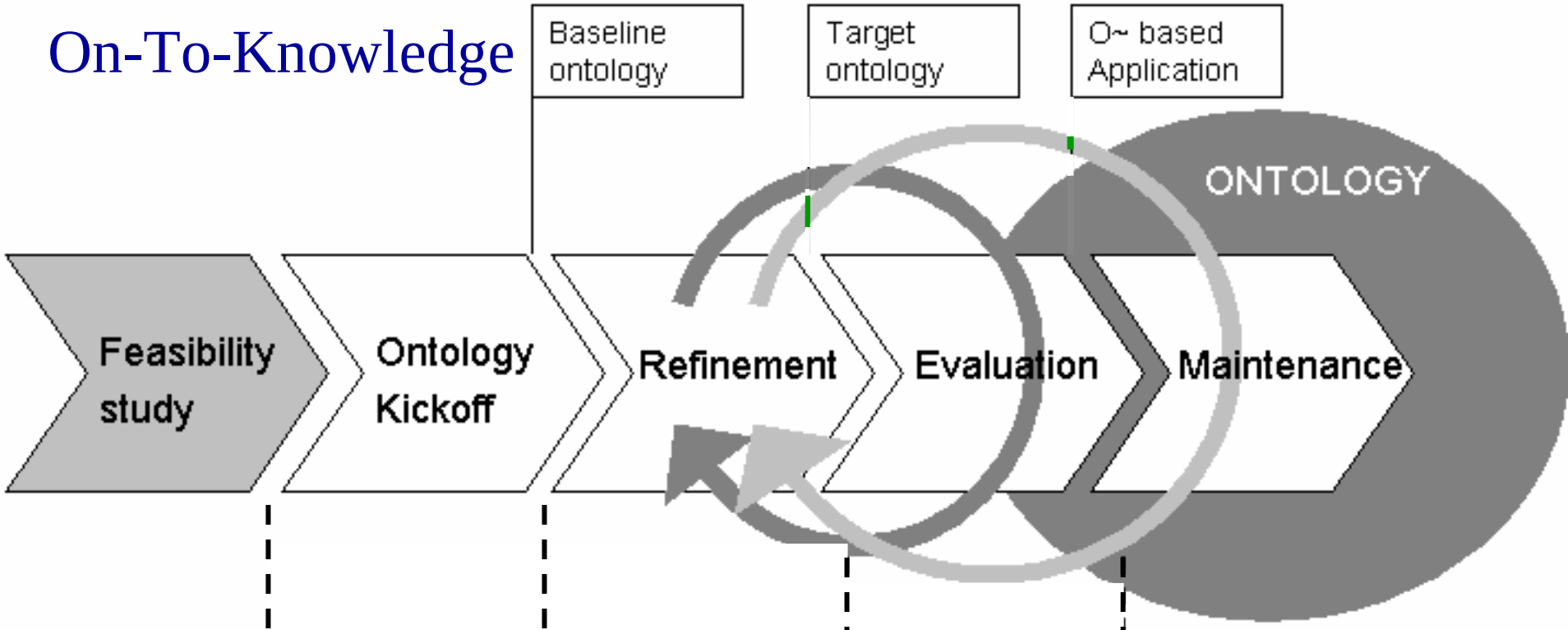
Build a preliminary ontology for refinement and augment with new definitions





A. Bernaras; I. Laresgoiti; J. Corera. *Building and reusing ontologies for electrical network applications* **ECAI96. 12th European Conference on Artificial Intelligence.** 1996. 298-302

On-To-Knowledge



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

- Requirement specification
- Analyze input sources
- Develop baseline taxonomy

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

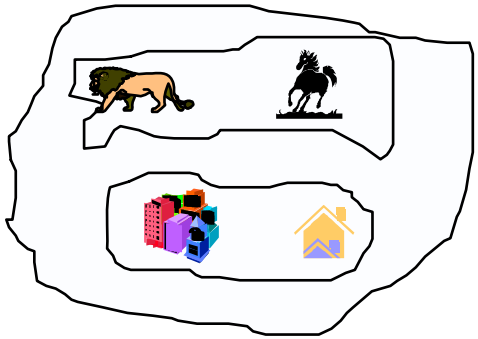
- Identify problem and opportunity areas
- Select most promising focus area and target solution

- Manage organizational maintenance process

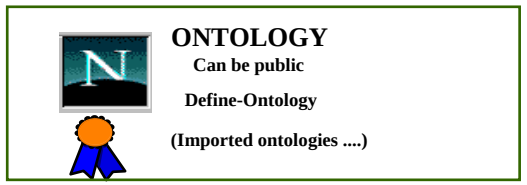
Project setting

Ontology development

METHONTOLOGY Framework

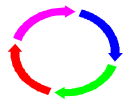


The world of ontologies

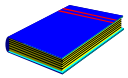


Tools

- To set up a life cycle



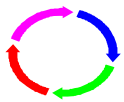
- Development process



Gómez-Pérez, A. *Knowledge Sharing and Reuse.* **In the Handbook of Applied Expert Systems.** CRC Press. 1998.

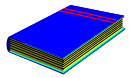
METHONTOLOGY Framework

Development Process: **Which activities**

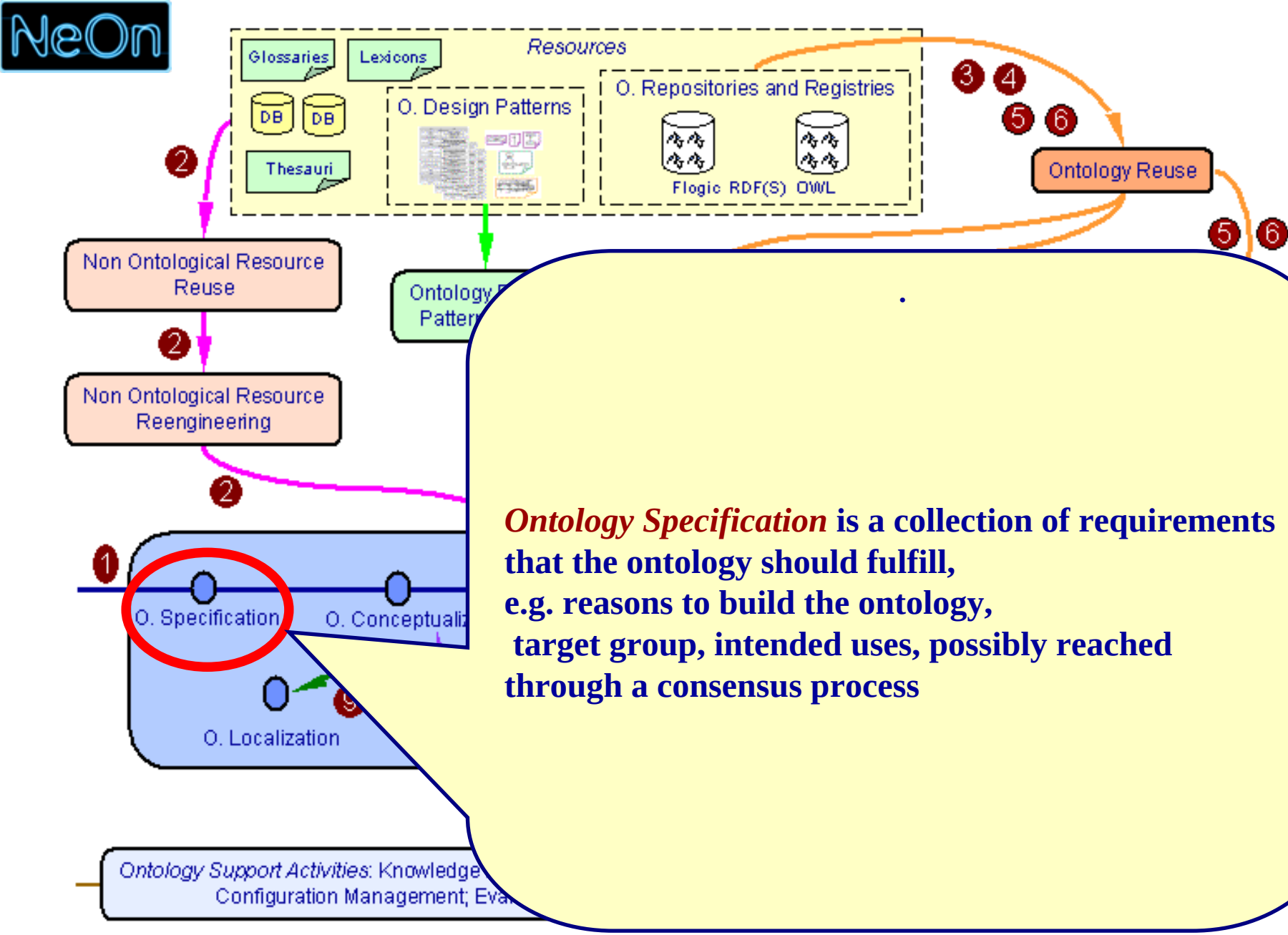


Life Cycle: **Order of activities**

Methodologies: **How to carry out the activities**



Gómez-Pérez, A. *Knowledge Sharing and Reuse.*
In the Handbook of Applied Expert Systems. CRC Press. 1998.


NeOn
Resources
Glossaries
Lexicons
DB
DB
Thesauri
O. Design Patterns
O. Repositories and Registries
Flogic RDF(S) OWL
Ontology Reuse
Non Ontological Resource Reuse
Non Ontological Resource Reengineering
O. Specification
O. Conceptualiz
O. Localization
Ontology Support Activities: Knowledge
Configuration Management; Eva
Ontology Specification is a collection of requirements that the ontology should fulfill,
e.g. reasons to build the ontology,
target group, intended uses, possibly reached through a consensus process

Definitions

Ontology (Requirements) Specification is a collection of requirements that the ontology should fulfill, e.g. reasons to build the ontology, target group, intended uses, possibly reached through a consensus process.

Requirements are those necessities/needs that the ontology to be built should represent/cover.

Competency Questions (CQs) are questions that the ontology to be built should be able to answer.

- CQs are a way to represent requirements.
- CQs can be written in natural language (NL) and can be formalized in ontology query languages (e.g. SPARQL).

Ontology Specification

Objective: To obtain an *Ontology Requirements Specification Document* (ORSD)

Description: The sequence of tasks to carry out will be presented in the following slides.

Output: The ORSD written using the ORSD template.

Recommended Techniques: Several techniques will be presented in the following slides for each task.

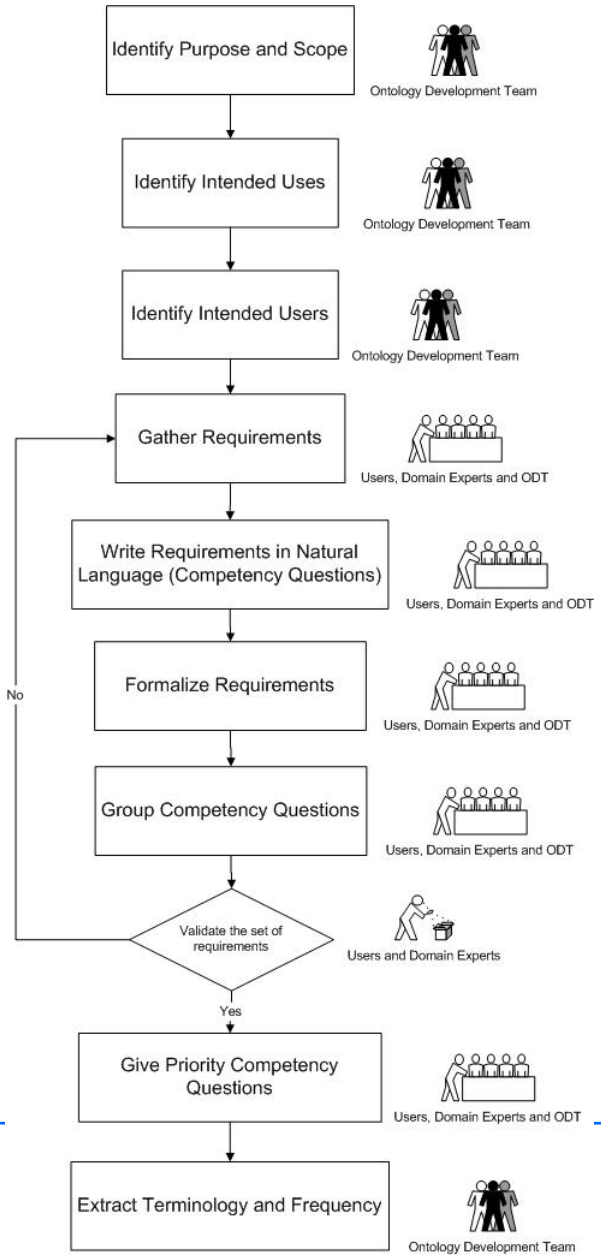
Recommended Tools: Due to no tool exists for the activity of ontology specification, in the following slides we will propose different independent tools for each task.

Ontology Specification. ORSD Template

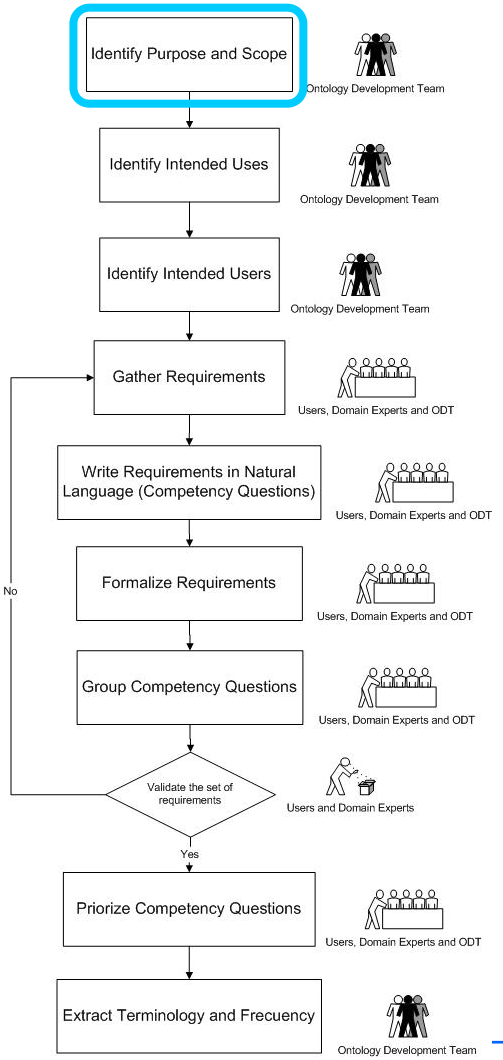
Ontology Requirements Specification Document Template

- Purpose and Scope (domain, goal, granularity, etc.)
- Intended uses (scenarios)
- Intended users
- Knowledge Resources used during this activity
- Groups of Competency Questions (CQs) with priorities
  - In Natural Language
  - In Ontology Query Language (optional)
- Pre-Glossary of terms with frequencies

Ontology Specification. Sequence of Tasks



Ontology Specification. Purpose and Scope



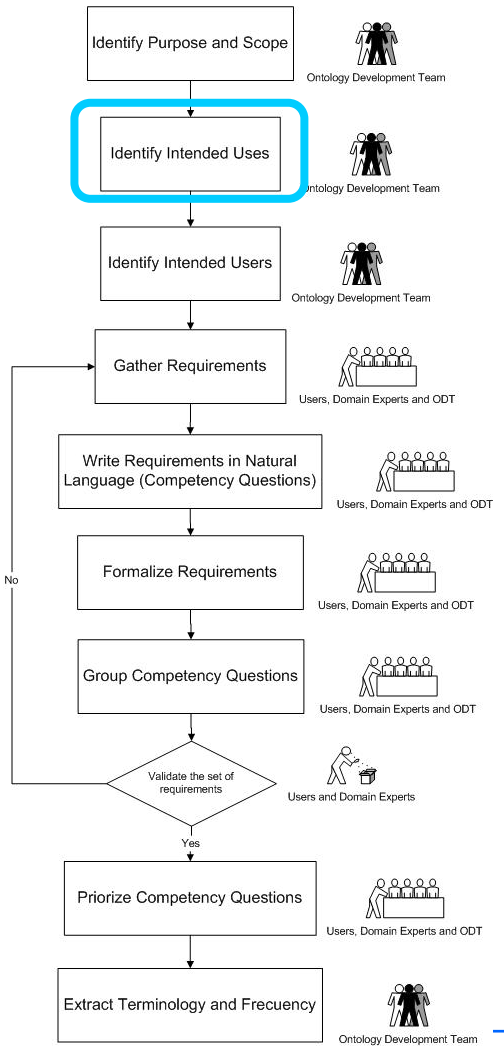
Objective: To obtain which is the main goal and domain for the ontology.

Output: The "Purpose and Scope" slot in the ORSD template.

Recommended Techniques: Interviewers (in the same place or virtual using internet/phone facilities) with users and domain experts.

Recommended Tools: ¿?

Ontology Specification. Intended Uses



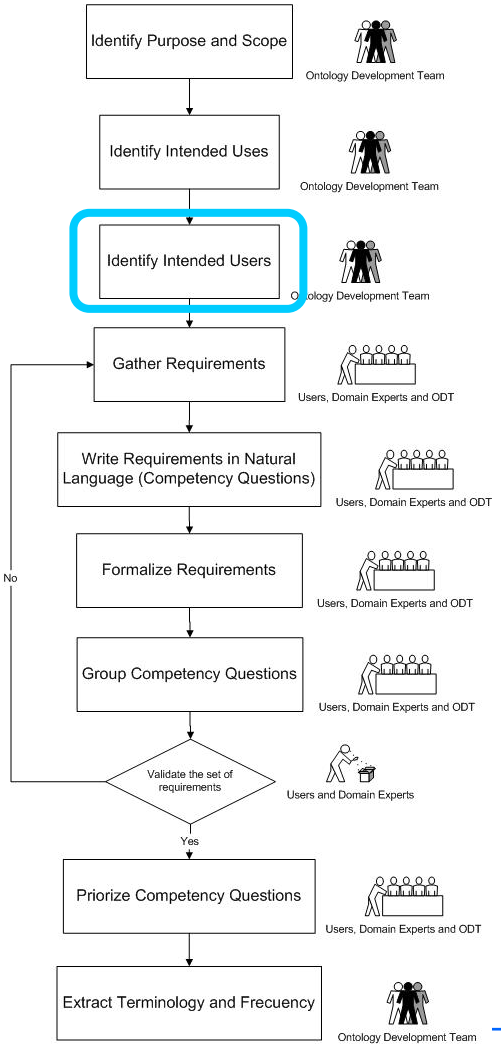
Objective: To obtain which are the main intented uses for the ontology, that is, in which kind of scenarios the ontology will be used.

Output: The “Intended Uses” slot in the ORSD template.

Recommended Techniques: Interviewers (in the same place or virtual using internet/phone facilities) with users and domain experts.

Recommended Tools: ¿?

Ontology Specification. Intended Users



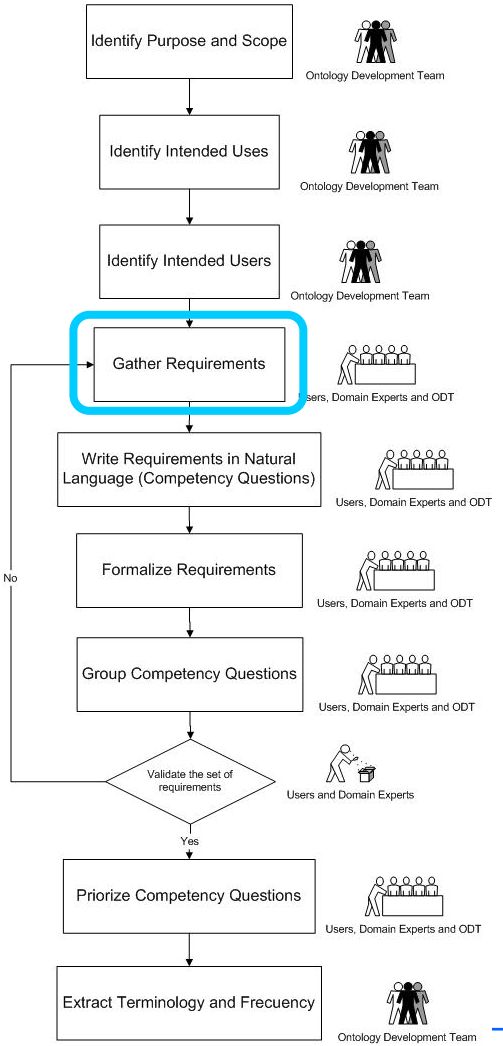
Objective: To obtain which are the main users of the ontology.

Output: The "Intended Users" slot in the ORSD template.

Recommended Techniques: Interviewers (in the same place or virtual using internet/phone facilities) with users and domain experts.

Recommended Tools: ¿?

Ontology Specification. Gather Requirements



Objective: To obtain the set of requirements (needs) that the ontology should fulfill.

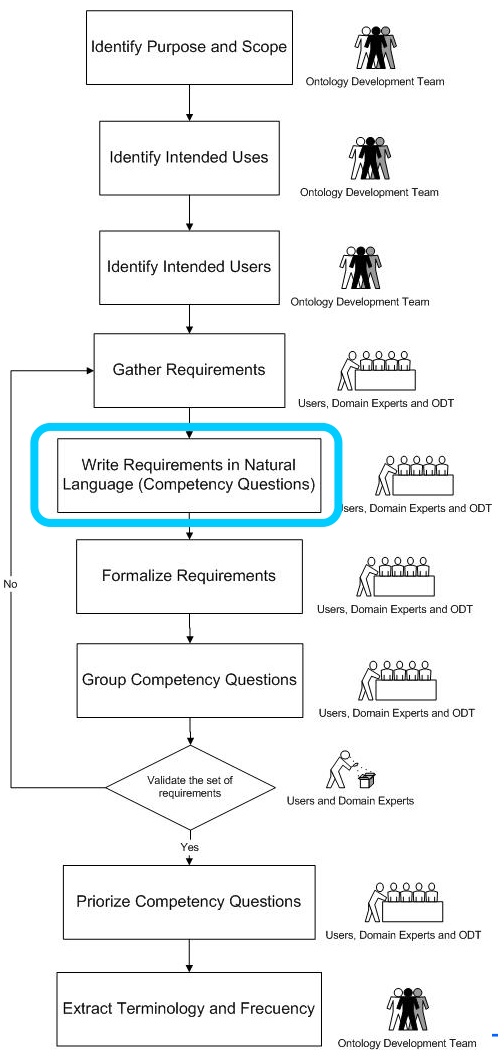
Output: A set of informal requirements to be used in the next task.

Recommended Techniques:

- Brainstorming
- Joint Application Development (JAD)
- Exploit scenarios and use cases using templates
- Interviewers with users and domain experts.

Recommended Tools: If people are geographically distributed, wiki tools can be used for obtaining the requirements.

Ontology Specification. Write CQs



Objective: To transform the informal set of requirements obtained in the previous task into Competency Questions.

Output: A set of Competency Questions written in Natural Language and a set of responses for the CQs.

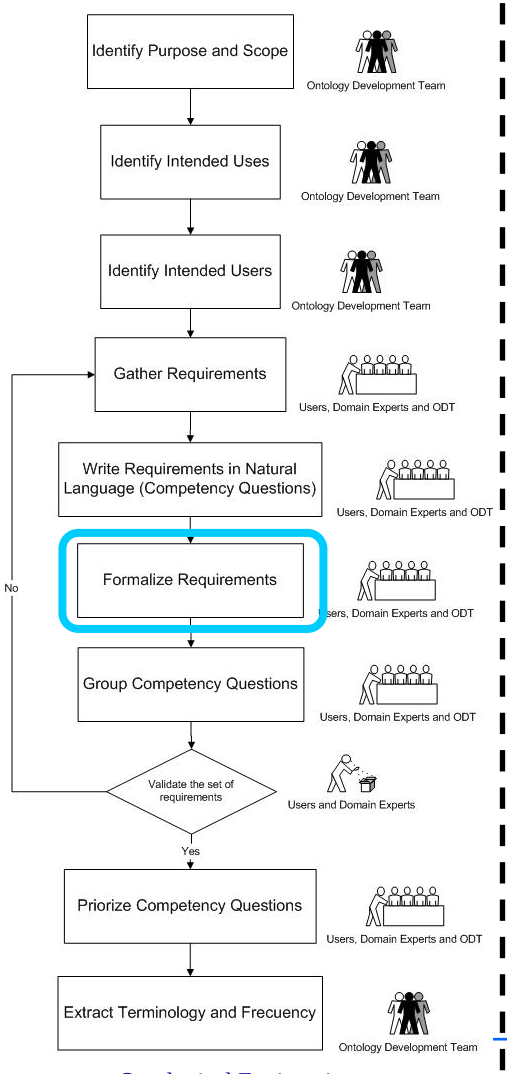
Recommended Techniques/Approaches:

- Top-Down: Complex queries decomposed in simple ones.
- Bottom-Up: Simple queries by composition derived in complex ones.
- Middle out.

Recommended Tools:

- MindMap Tools.
- Excell.
- Cicero Tool, if people are geographically distributed.

Ontology Specification. Formalize CQs



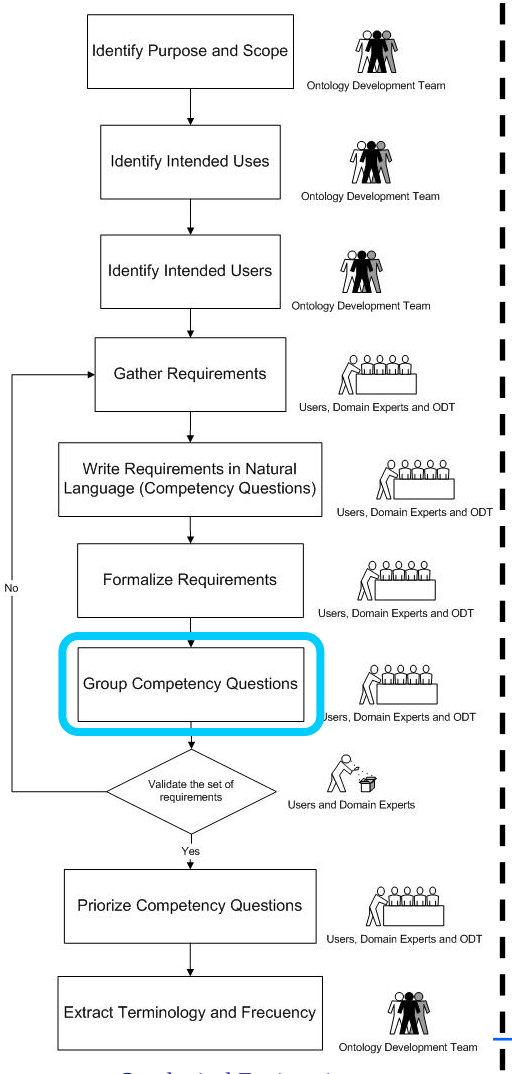
Objective: To formalize the NL Competency Questions.

Output: A set of Competency Questions formalize in an ontology query language.

(This task is optional, but recommended)

Formalized CQs will be used for evaluating (semi)-automatically the ontology against the requirements.

Ontology Specification. Group CQs



Objective: To classify the CQs into different groups or categories.

Output: A classification of the set of CQs.

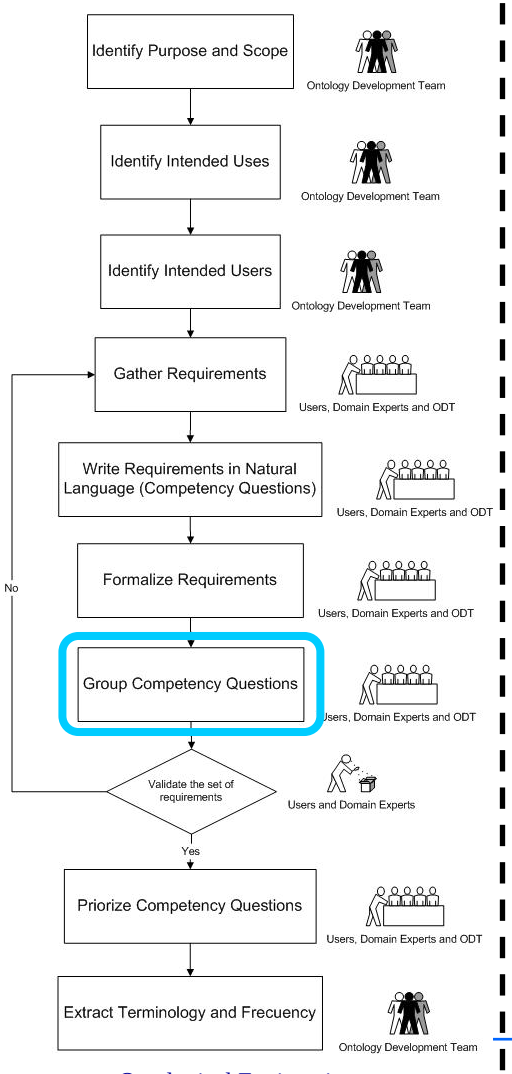
Recommended Techniques:

- Card Sorting (*manual*).
- Clustering NL sentences (*automatic*).
- Information Extraction (*automatic*).

Recommended Tools:

- If the card sorting technique is used, then MindMap Tools or Cicero Tool (for distributed teams).

Ontology Specification. Group CQs (II)



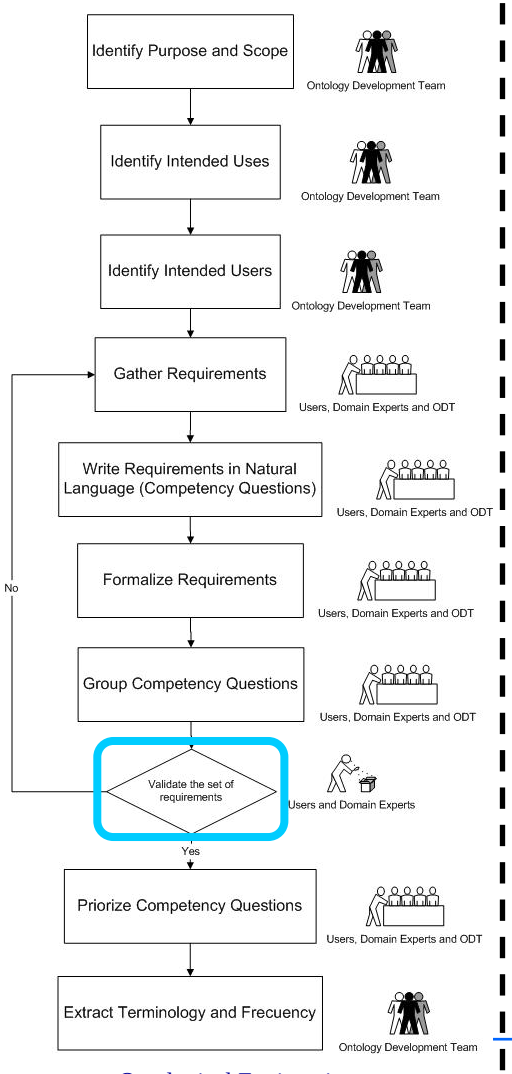
Why grouping is useful?

- *Module-based development.*
- *Prototype-based development.*

Grouping criteria:

- *Based on the main terms appearing in CQs (domain dependant criteria).*
- *Based on general dimensions or terms (domain independent criteria). E.g.: time and date, units of measure, currencies, location, languages, etc.*

Ontology Specification. Validate CQs

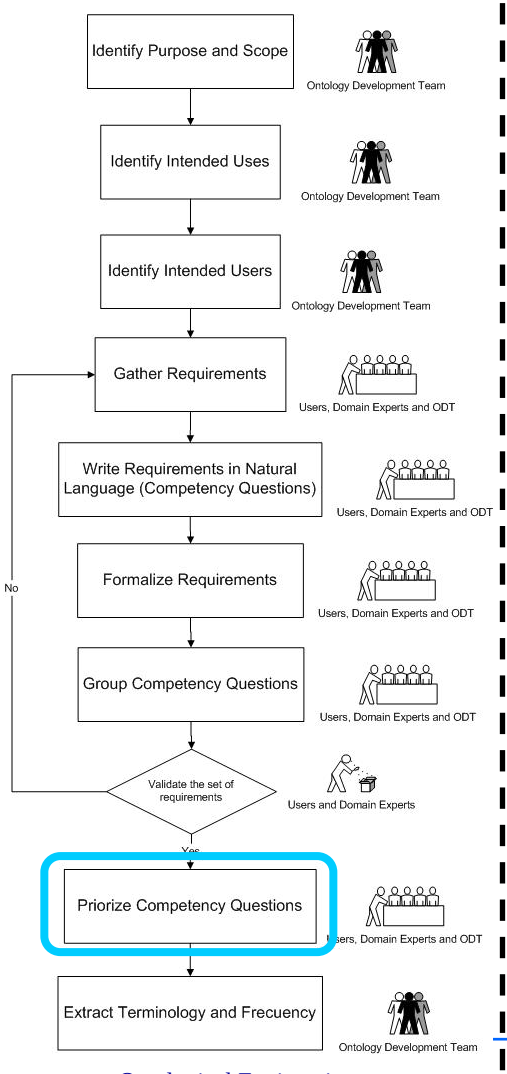


Objective: To identify possible conflicts between CQs, missing CQs, contradictions in CQs.

Criteria:

- Correctness.
- Completeness.
- Consistent.
- Verificable.
- Understandable.
- No Ambiguity.
- No Redundancy.

Ontology Specification. Give Priority to CQs



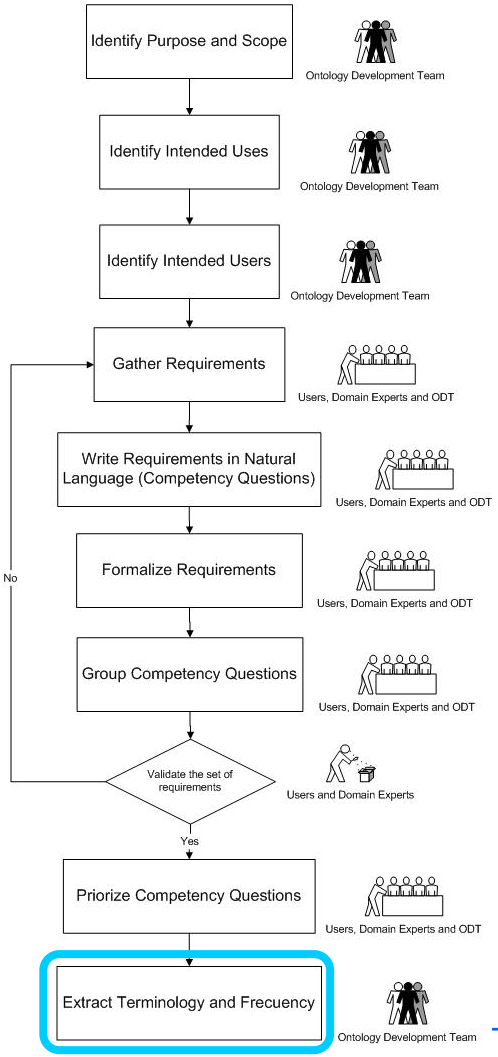
Objective: To give different levels of priority to the.

Output: A set of Competency Questions with a concrete priority.

(This task is optional, but recommended)

Priorities in CQs will be used for planning the ontology development.

Ontology Specification. Group CQs



Objective: To extract a pre-glossary to be used in the conceptualization activity.

Output: A pre-glossary with the main terms used in the CQs with their associated frecuency.

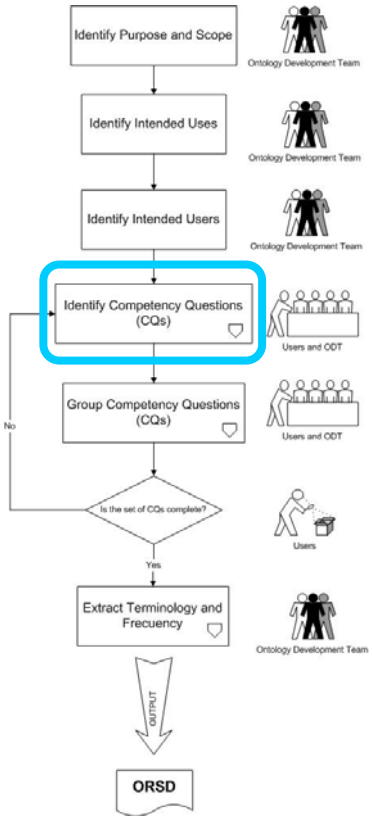
Recommended Techniques:

- Terminology Extraction (*automatic*): terms are names, adjectives and verbs.
  - From CQs we will extract the conceptual model.
  - From responses to CQs we will extract the Universe of Discurse (the knowledge base).

Recommended Tools:

- Terminology Extraction Tools.
- Terminology Frequency Tools.

Scenario 1: Ontology Specification



Identify Competency Questions (CQs)

Strategies:

1. Top down: complex queries decomposed in simple ones
2. Bottom up: simple queries by composition derived in complex ones
3. Middle out

Each CQ should have (at least) a response.

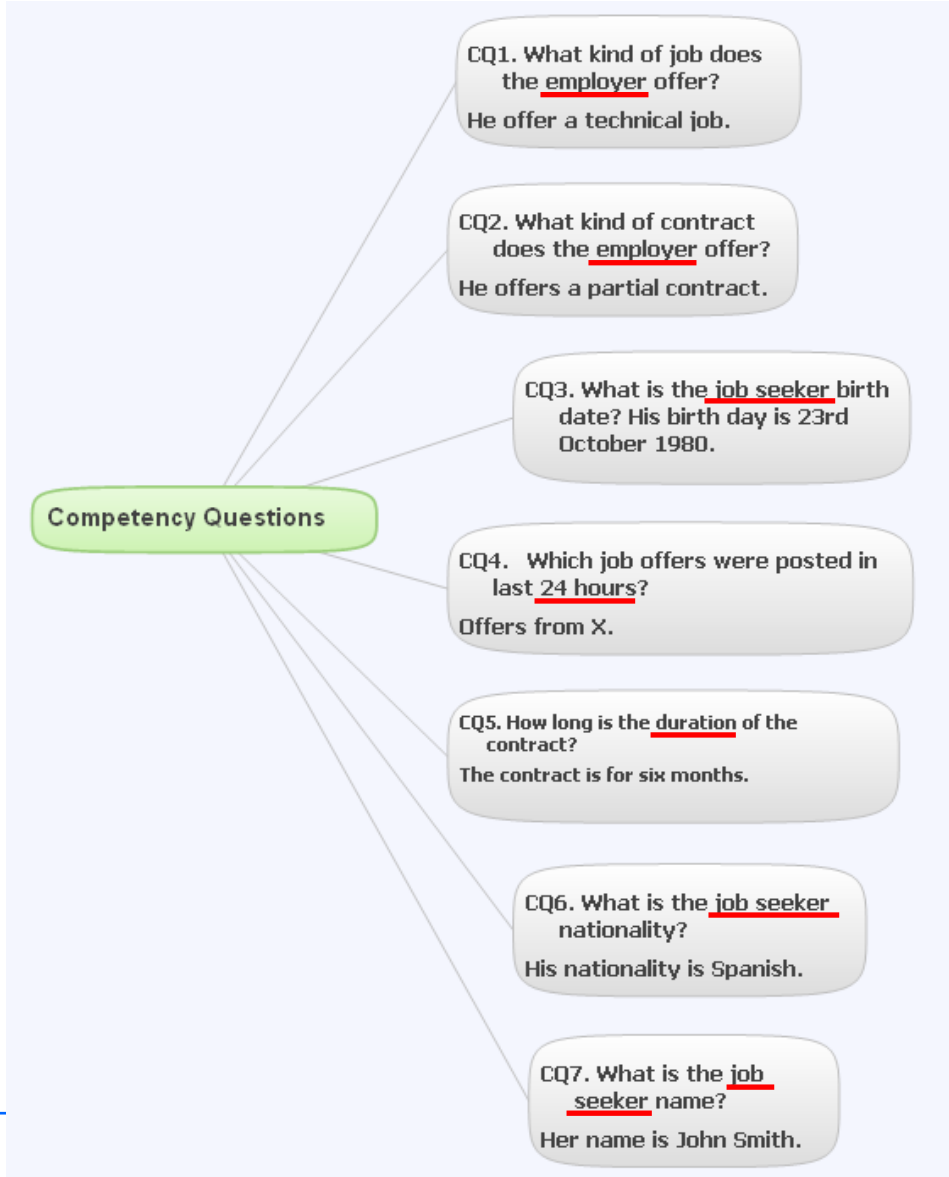
Tools:

FreeMap, MindManager, ConceptDraw MindMap

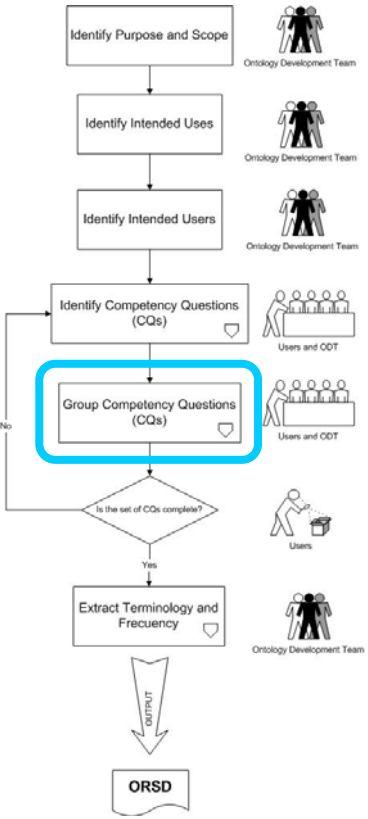
We will also provide:

- ***Guidelines for writing the CQs in a correct way****, including positive and negative examples.*
- ***Guidelines for writing a complete document****.*

Scenario 1: Ontology Specification



Scenario 1: Ontology Specification



Group Competency Questions (CQs)

Techniques:

1. Fichas *(manual)*
2. Clustering NL sentences *(authomatic)*
3. Information extraction: *(authomatic)*
   a. Tokenizar e identificar sustantivos, adjetivos y verbos
   b. Estadisticas para clustering según criterios de agrupación.

Tools:

FreeMap, MindManager, ConceptDraw MindMap

Why grouping is useful?

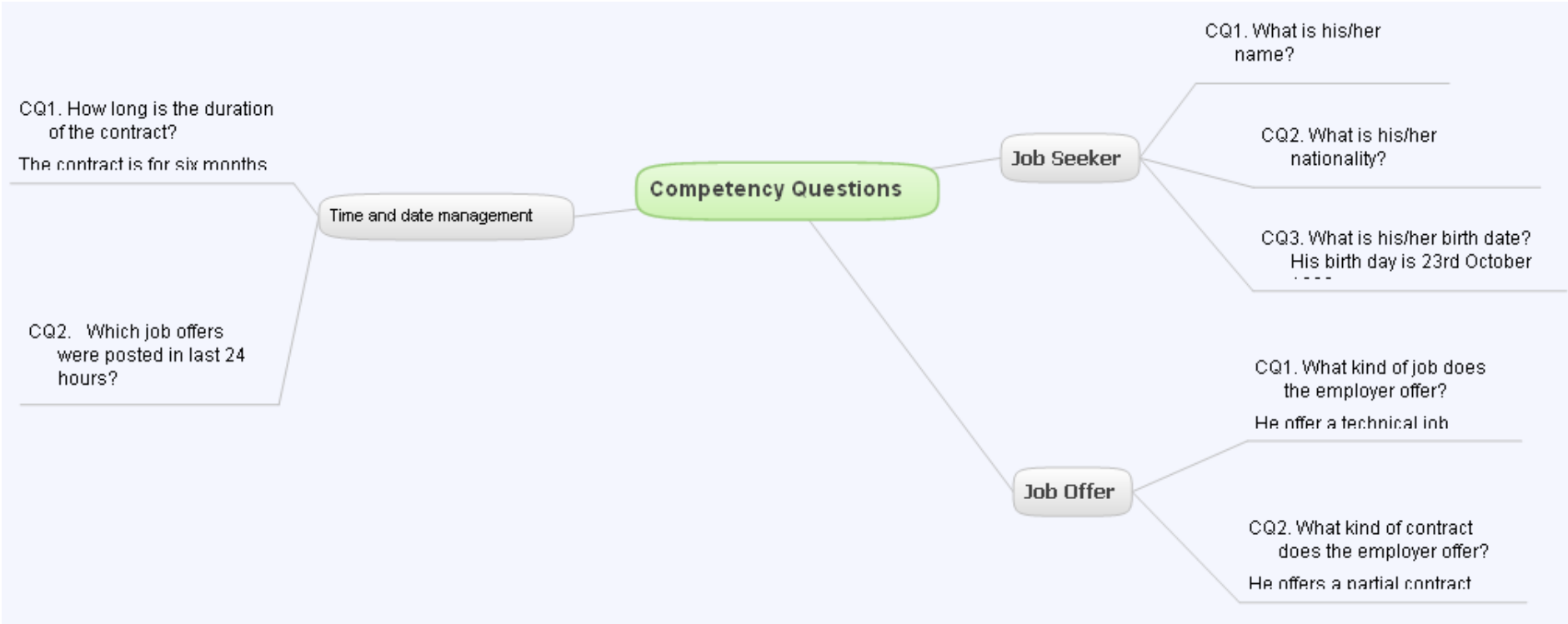
→Module-based development

→Prototype-based development

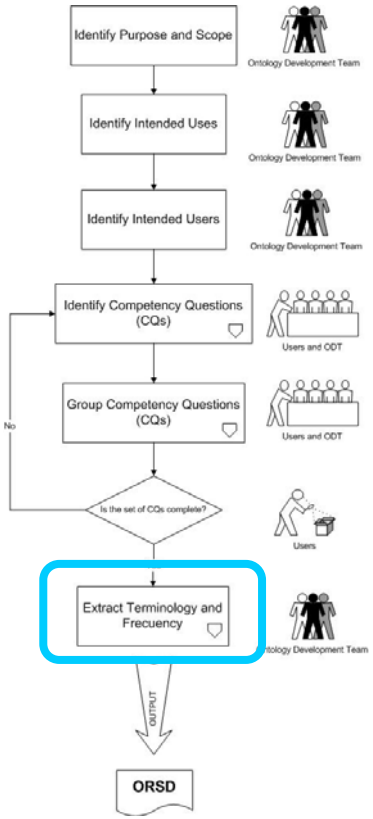
Grouping criteria:

- *Domain dependent o Principales terminos de la onto (job seeker, job offer) + domain independent o dimensiones fijas (time and date, units of measure, currencies, location, languages, …).*
- *Concept vs events modelling.*

Scenario 1: Ontology Specification



Scenario 1: Ontology Specification



Extract Terminology and Frecuency

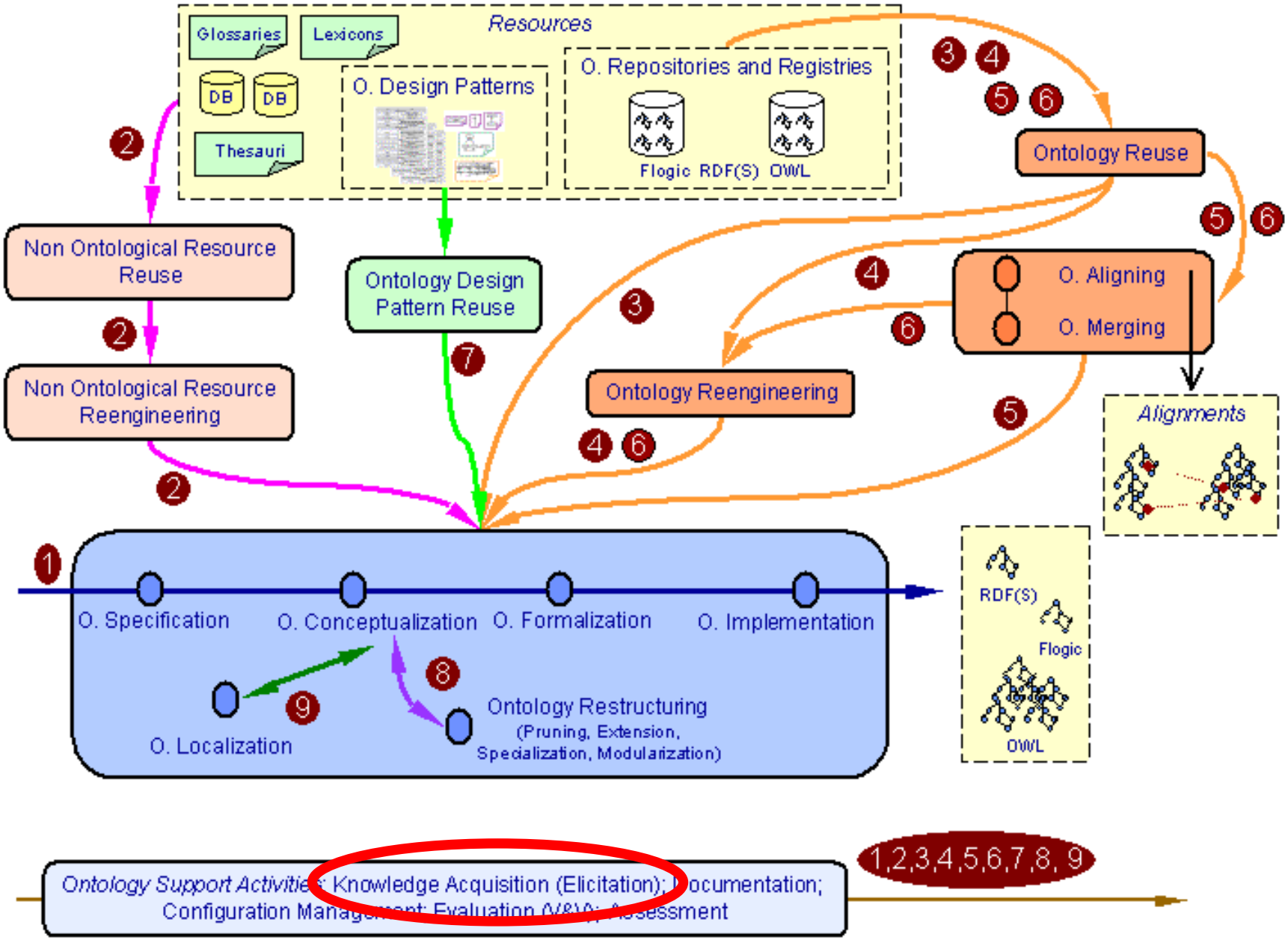
Techniques:

Terminology extraction. Terms (sustantivos, adjetivos, verbos). This should be the Pre-Glossary.

i. Preguntas: Modelo conceptual

ii. Respuestas: Objects in the Universe of Discourse (un subconjunto son las name entities)

Frequency of the terminology (to be used later on for searching)


Resources
Glossaries
Lexicons
DB
DB
Thesauri
O. Design Patterns
O. Repositories and Registries
Flogic RDF(S) OWL
Ontology Reuse
Non Ontological Resource Reuse
Ontology Design Pattern Reuse
O. Aligning
O. Merging
Non Ontological Resource Reengineering
Ontology Reengineering
Alignments
O. Specification
O. Conceptualization
O. Formalization
O. Implementation
O. Localization
Ontology Restructuring
(Pruning, Extension, Specialization, Modularization)
RDF(S)
Flogic
OWL
Ontology Support Activities: Knowledge Acquisition (Elicitation); Documentation;
Configuration Management; Evaluation (V&V); Assessment
1,2,3,4,5,6,7,8, 9

Elicitation Techniques. Types (I)

- **Protocol-generation techniques**
  - Various types of interviews: unstructured, semi-structured and structured
  - Various types of reporting techniques: self-report and shadowing
  - Various types of observational techniques
- **Rules of thumb**
  - Take experts off the job for short time periods
  - Focus on the essential knowledge
  - Collate knowledge from different experts
- **Protocol analysis techniques**
  - Used with transcripts of interviews or other text-based information
  - Useful to identify various types of knowledge (goals, decisions, relationships and attributes)
  - They act as a bridge between the use of protocol-based techniques and knowledge modelling techniques.
- **Hierarchy-generation techniques (laddering)**
  - Useful for building taxonomies or other hierarchical structures (goal trees and decision networks)

Card Sorting. Collect and organise the concepts

- Write down each concept/idea on a card
- Organise them into piles
- Link the piles together
- Do it again, and again

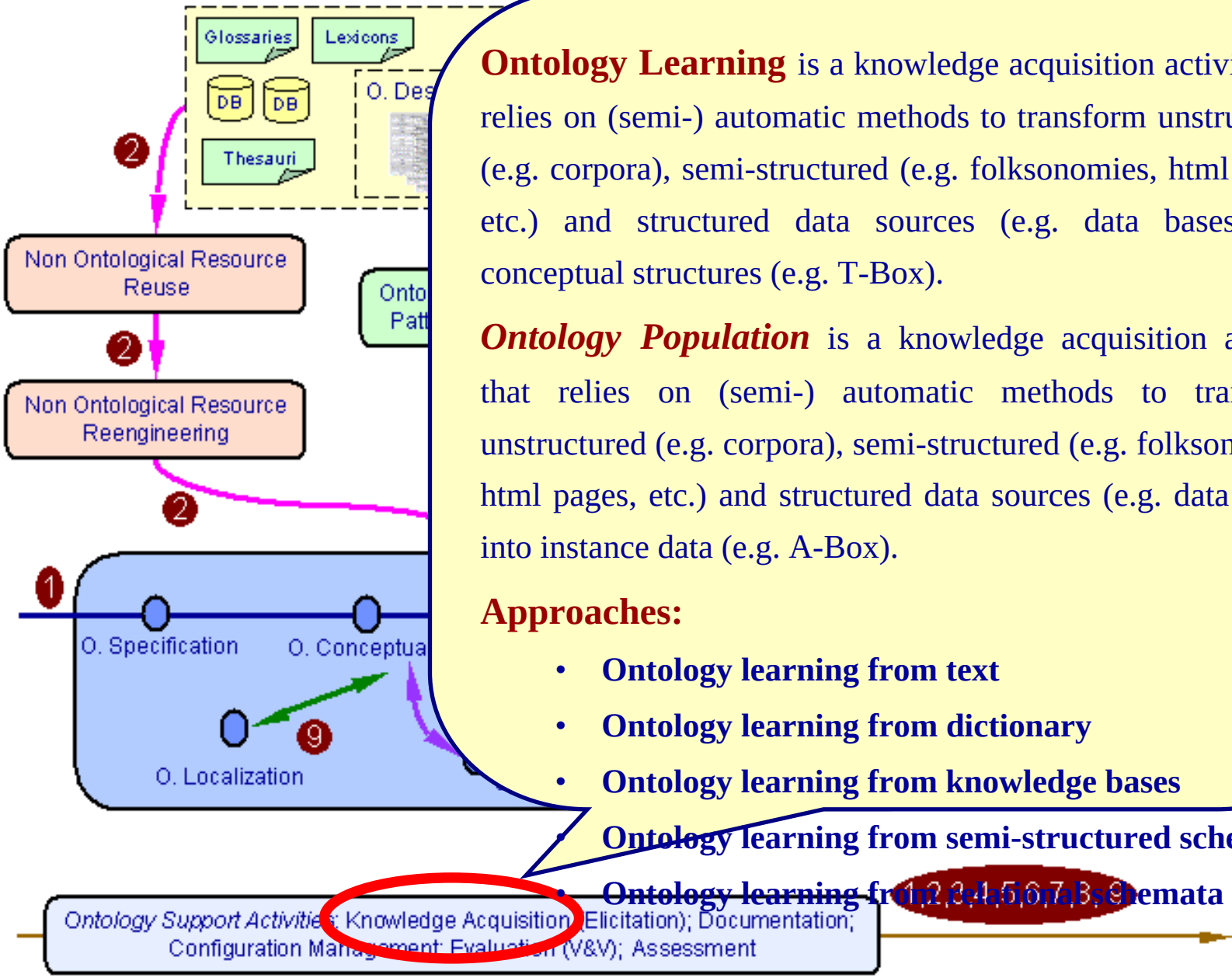
- Example. Create an ontology for an index of a children's book of animals including
  - Where they live
  - What they eat
    - Carnivores, herbivores and omnivores
  - How dangerous they are
  - How big they are
  - A bit of basic anatomy
    - numbers of legs, wings, toes, etc.

Laddering. Extend the Concepts

- Take a group of things and ask what they have in common
  - Then what other ‘siblings’ (brother or sister) there might be
- E.g.
  - Plant, Animal → Living Thing
    - Might add Bacteria and Fungi but not now
  - Cat, Dog, Cow, Person → Mammal
    - Others might be Goat, Sheep, Horse, Rabbit,…
  - Cow, Goat, Sheep, Horse → Farm animal
    - What others are there? Do they divide amongst themselves?
  - Wild, Domestic → Domestication
    - What other states – “Feral” (domestic returned to wild)

Laddering. Choose Some Main Axes

- Add abstractions where needed
  - e.g. “Living thing”
- Identify relations
  - e.g. “eats”, “owns”, “parent of”
- Identify definable things
  - e.g. “child”, “parent”, “Mother”, “Father”
    - Things where you can say clearly what it means
      - Try to define a dog precisely – very difficult
        - A “natural kind”
- Make names explicit



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).

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).

Approaches:

- **Ontology learning from text**
- **Ontology learning from dictionary**
- **Ontology learning from knowledge bases**
- **Ontology learning from semi-structured schemata**
- **Ontology learning from relational schemata**

Techniques used in different OL approaches

OL from Text

- **Natural Language Techniques**
- **Clustering techniques**
- **Machine learning**
- **Statistical aproach**

OL from Dictionary

- **Natural Language Processing**
- **Statistical aproach**

OL from Knowledge Bases

- **Rules**

OL from Semi-structured Schemata

- **Graph Theory**
- **Machine Learning**
- **Pattern Recognition**
- **Clustering**
- **Ontological Techniques**

OL from Relational Schemata

- **Mapping Techniques**
- **Reverse Engineering**

All of them described and compared in OntoWeb deliverable D1.5

OL from texts: Methods and techniques (I)

Aguirre and colleagues' method	**URL:** Not available
Alfonseca and Manandhar's method	**URL:** http://www.ii.uam.es/~ealfon
Aussenac-Gilles and colleagues' approach	**URL:** http://www-lipn.univ-paris13.fr/~szulman/TERMINAE.html
Bachimont's method	**URL:** http://opales.ina.fr/public/
Faatz and Steinmetz approach	**URL:** Not available
Gupta and colleagues' approach	**URL:** Not available
Hahn and colleagues' method	**URL:** Not available
Hearst's approach	**URL:** http://www.ii.uam.es/~ealfon
Hwang's method	**URL:** http://www.argreenhouse.com/InfoSleuth/index.shtml
Khan and Luo's method	**URL:** Not available
Kietz and colleagues' method	**URL:** http://ontoserver.aifb.uni-karlsruhe.de/texttoonto/
Lonsdale and colleagues' method	**URL:** http://www.ttt.org/salt/index.html
Missikoff and colleagues' method	**URL:** Not available
Moldovan and Girju's method	**URL:** Not available
Nobécourt approach	**URL:** Not available
Roux and colleagues' approach	**URL:** Not available
Wagner approach	**URL:** Not available
Xu and colleagues' approach	**URL:** Not available

OL from dictionary

Hearst's method	**URL:** Not available
Rigau and colleagues' method	**URL:** http://www.lsi.upc.es/~rigau/
Jannink and Wiederhold's approach	**URL:** Not available

OL from knowledge bases

OL from semi-structured schemata

Deitel and colleagues' approach	**URL:** http://mondeca-publishing.com/s/anonymous/title11884.html
Doan and colleagues approach	**URL:** Not available
Papatheodorou and colleagues' method	**URL:** http://www.educanext.org/
Volz and colleagues' approach	**URL:** http://www.aifb.uni-karlsruhe.de/WBS/rvo/raphael-bib.html#wonderweb-D11

OL from relational schemata

Johannesson's method	**URL:** Not available
Kashyap's method	**URL:** Not available
Rubin and colleagues' approach	**URL:** http://www.nigms.nih.gov/funding/pharmacogenetics.html
Stojanovic and colleagues' approach	**URL:** http://wonderweb.semanticweb.org/publications.shtml

OL from texts: Methods and techniques(I)

Name	Main goal	Main techniques used	Reuse other ontologies	Sources used for learning	Tool associated	Evaluation	Bibliography
Aguirre and colleagues' method	To enrich concepts in existing ontologies	Statistical approach Clustering Topic signatures	Yes	Domain Text WordNet	Information not available in papers	User	Agirre et al., 2000
Alfonseca and Manandhar's method	To enrich an existing ontology with new concepts	Topic signatures Semantic distance	Yes	Domain text WordNet	Welkin	Expert	Alfonseca et al., 2002
Aussenac-Gilles and colleagues' approach	To learn concepts and relations among them	Linguistic analysis Clustering techniques	Yes	Domain Text Domain ontologies	GEDITERM TERMINAE	User	Aussenac-Gilles and colleagues, 2000a and 2000b
Bachimont's method	To build a taxonomy	NLP techniques	No	Domain text	DOE	Expert	Bachimont et al., 2002
Faatz and Steinmetz approach	To enrich an existing ontology with new concepts	Statistical approach Semantic distance	Yes	Domain corpus Domain ontology	Any ontology workbench	Expert	Faatz et al 2002
Gupta and colleagues' approach	To build sub-languages in WordNet	NLP techniques Term-extraction techniques	Yes	Domain text WordNet	SubWordNet Engineering tool	Expert	Gupta et al. 2002
Hahn and colleagues' method	To learn new concepts	Concept hypothesis based on linguistic and conceptual quality labels Statistical approach	No	Domain text	Information not available in papers	Empirical measures and by an expert	Hahn et al 1998
Hearst's approach	To enrich an existing ontology	NLP techniques Linguistic patterns	Yes	Domain Text WordNet	Welkin	Expert	Hearst 1998 and Alfonseca et al. 2002

OL from texts: Methods and techniques(II)

Name	Main goal	Main techniques used	Reuse other ontologies	Sources used for learning	Tool associated	Evaluation	Bibliography
Hwang's method	To elicit a taxonomy	NLP techniques ML techniques Statistical approach	No	Domain Text	Information not available in papers	Expert	Hwang 1999
Khan and Luo's method	To learn concepts	Clustering techniques Statistical approach	Yes	Domain text WordNet	Information not available in papers	Expert	Khan et al. 2002
Kietz and colleagues' method	To learn concepts and relations among them to enrich an existing ontology	NLP Statistical approach	Yes	Domain and non-specific domain Text Domain ontologies WordNet	Text-To-Onto	User	Kietz et al., 2000 Maedche el al. 2001
Lonsdale and colleagues' method	To discover new relationships in an existing ontology	NLP Mappings Linguistic technique	Yes	Terminological databases Domain ontology WordNet Domain text	Information not available in papers	User/Expert	Lonsdale et al., 2002
Missikoff and colleagues' method	To build taxonomies and to fuse with an existing ontology with	NLP Statistical approach ML techniques	Yes	Domain text WordNet	OntoLearn	Expert	Navigli et al., 2003 and Missikoff et al. 2002
Moldovan and Girju's method	To enrich an existing ontology	NLP techniques	Yes	Domain Text Lexical resources WordNet	Information not available in papers	Expert	Moldoban and Girju 2000 and 2001, and Harabagiu et al. 2000

OL from texts: Methods and techniques(III)

Name	Main goal	Main techniques used	Reuse other ontologies	Sources used for learning	Tool associated	Evaluation	Bibliography
Nobécourt approach	To learn concept and relations among them	Linguistic analysis	No	Domain Text	TERMINAE	User/expert	Nobécourt 2000
Roux and colleagues' approach	To enrich a taxonomy with new concepts	Verb-patterns	Yes	Domain text Domain ontology	Information not available in papers°	Expert	Roux et al., 2000
Wagner approach	To enrich an existing ontology with new relationships	Statistical approach	Yes	WordNet	Information not available in papers	Expert	Wagner 2000
Xu and colleagues' approach.	To learn concepts and relations between them	NLP techniques Statistical approach Text-mining techniques	Yes	Annotated text corpus WordNet	TFIDF	Expert	Xu et al., 2002

OL from texts: tools (I)

18 tools described	
ASIUM	**URL:** http://www.lri.fr/~faure/Demonstration/Presentation_Demo.html
Caméléon	**URL:** Not available
Corporum-Ontobuilder	**URL:** http://ontoserver.cognit.no
DOE	**URL:** http://opales.ina.fr/public/
KEA	**URL:** http://www.nzdl.org/Kea/
LTG	**URL:** http://www.ltg.ed.ac.uk/%7Emikheev/workbench.html
MO'K Workbench	**URL:** Not available
OntoLearn	**URL:** Not available
Prométhée	**URL:** http://www.sciences.univ-nantes.fr/info/perso/permanents/morin/promethee/
SOAT	**URL:** http://www.iis.sinica.edu.tw/IASL/en/index.htm
SubWordNet E.P.	**URL:** http://www.aic.nrl.navy.mil/~aha/cbr/luikm.html
SVETLAN'	**URL:** http://www.limsi.fr/Individu/gael/ManuscritThese/
TDIDF	**URL:** Not available
TERMINAE	**URL:** http://www-lipn.univ-paris13.fr/~szulman/TERMINAE.html
TextStorm and Clouds	**URL:** Not available
TextToOnto	**URL:** http://ontoserver.aifb.uni-karlsruhe.de/texttoonto/
Welkin	**URL:** http://www.ii.uam.es/~ealfon
WOLFIE	**URL:** Not available

OL from dictionary

SEID **URL:** http://www.lsi.upc.es/~rigau/

DOODLE **URL:** Not available

OL from knowledge bases

OL from semi-structured schemata

OntoBuilder **URL:** http://www.cs.msstate.edu/~gmodica/Education/OntoBuilder/

OL from relational schemata

OL from texts: tools (I)

Name	Goal and scope	Learning technique	Method followed to learn	Sources	User intervention	Interoperability	Bibliography
ASIUM	To learn taxonomic relations	Conceptual clustering techniques	Own method	Text syntactically analysed	Whole process	Can be used to perform the knowledge acquisition to any other ontology development tool	Faure et al 2000, 1999, and 1998
Caméléon	To tune generic lecixo-syntactic patterns or build new ones. To find taxonomic and non taxonomic lexical relations in texts and to enrich a conceptual model with these lexical relations	To reuse and tuning of generic patterns (mainly for taxonomic relations), hearst's proposal, and pattern identification in text help to learn lexical relations and their validation leads to conceptual relations	Own method	Texts processed by taggers Its own base of generic patterns	Validates, adapts, or defines new domain specific patterns and relations Domain expert just validates the model	Imports lists of terms from any term extractor	Aussenac-Gilles and Seguela 2000
Corporum-Ontobuilder	To extract initial taxonomy	Linguistic and semantic techniques	Own method	Text	Not necessary	OntoWrapper and OntoExtract	Engels 2001 and 2000
DOE	To help to the ontologist in the process of building an ontology	Differential Semantic	Bachimont's method	NL text	Whole process	None	Bachimont 2000

OL from texts: tools (II)

Name	Goal and scope	Learning technique	Method followed to learn	Sources	User intervention	Interoperability	Bibliography
KEA	To Summarize documents extracting keywords	Statistical approach ML techniques Lexical processing	Own method	NL text	Evaluation	WEKA ML Workbench	Jones and Paynter, 2002
LTG	To discover internal relations of texts in NL	Statistic Inference Linguistic technique	Own method	NL text	Whole process	Can be used to perform the knowledge acquisition to any other ontology development tool	Mikheev and Finch, 1997
MO'K Workbench	To learn concept taxonomy	Conceptual clustering	Own method	Tagged text	Whole process	Can be used to perform the knowledge acquisition to any other ontology development tool	Bisson et al. 2000
OntoLearn	To enrich a domain ontology	NLP techniques ML techniques	Missikoff and colleagues' method	NL text	Evaluation	None	Velardi et al., 2002 and 2001
Prométhée	Extraction and refinement of lexical-syntactic patterns	Learning from examples	Own method	Pattern bases	Whole process	Information not available in papers	Morin 1999, 1998
SOAT	Acquisition of relationships	Phrase-patterns	Own method	NL text	Information not available in papers	Information not available in papers	Wu et al 2002

OL from texts: tools (III)

Name	Goal and scope	Learning technique	Method followed to learn	Sources	User intervention	Interoperability	Bibliography
SubWordNet Engineering Process	Build a Sub WordNet	NLP techniques Statistical approaches	Gupta and colleagues' approach	NL text	Whole process	Information not available in papers	Gupta et al., 2002
SVETLAN'	Build a concept hierarchy	Conceptual clustering	Own method	NL text	Validation	Information not available in papers	Chaelandar and Grau, 2000
TFIDF	To learn concepts and relation between them	Text-mining Statistical approach	Hybrid text-mining approach to acquire domain terms	NL text	Evaluation	SPPC NLP tool	Xu et al., 2002
TERMINAE	To build an initial ontology	Conceptual clustering	Own method	NL text	Validation	Information not available in papers	Biébow and Szulman 1999
Text-To-Onto	To find taxonomic and non-taxonomic relations	Statistical approach Pruning techniques Association rules	Kietz and colleagues' method	NL text Dictionaries Ontologies	Validations	KAON tool suite	Maedche and Volz, 2001
TextStorm and Clouds	To build a taxonomy	NLP techniques Linguistic hypothesis	Own method	NL text	Whole process	Information not available in papers	Oliveira et al., 2001
Welkin	To enrich automatically existing general purpose ontologies with new terms	Semantic Similarity measures	Alfonseca and Manandhar's method Hearst's approach	Domain corpus WordNet	Not necessary	None	Alfonseca and Rodríguez, 2002
WOLFIE	To learn a semantic lexicon	NLP techniques Statistical approach	Own method	Pre-processed corpus Examples	Validation	CHILL	Thompson and Mooney, 1997