

AIFB

Semantic Web Technologies II ss 2009

# 22.04.2009

**Engineering Ontologies and Semantic Applications** 

Dr. Sudhir Agarwal Dr. Stephan Grimm Dr. Peter Haase PD Dr. Pascal Hitzler

**Denny Vrandecic** 



**Content licensed under Creative Commons** *http://creativecommons.org/licenses/by/2.0/de/* 

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft



# Topics



- What are applications of ontologies in semantic applications?
- What are development processes and lifecycle activities for engineering
  - 1) ontologies
  - 2) semantic applications?
- What are supporting infrastructures/architectures for engineering ontologies and semantic applications



AIFBO



# **Ontologies - Definition**

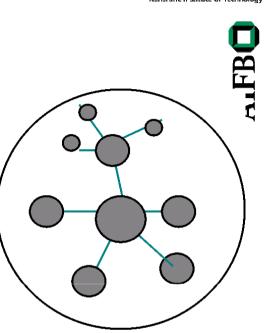


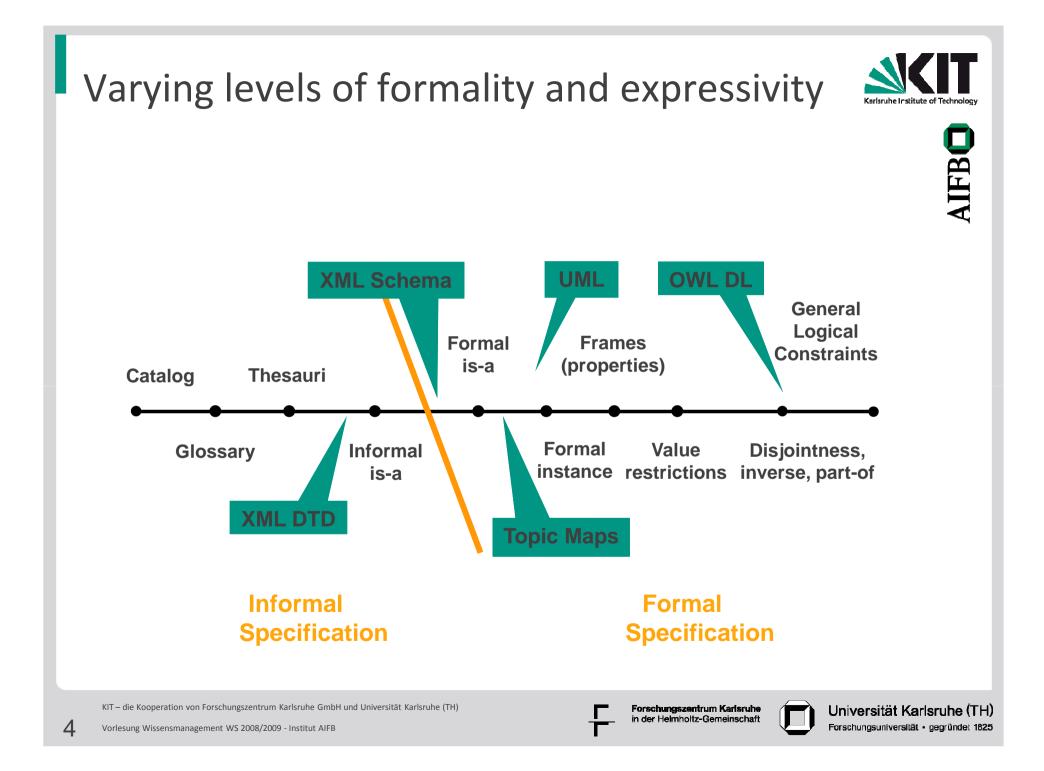
- "An ontology is a shared understanding of some domain of interest." [Uschold, Gruninger96]
- "An ontology is an explicit specification of a shared conceptualisation." [Gruber93]
- Ontologies
  - Describe a common vocabulary of terms and their relations
  - Define the meaning of this vocabulary
- Semantics based on logical languages guarantees welldefined interpretation
- Can be shared as resources

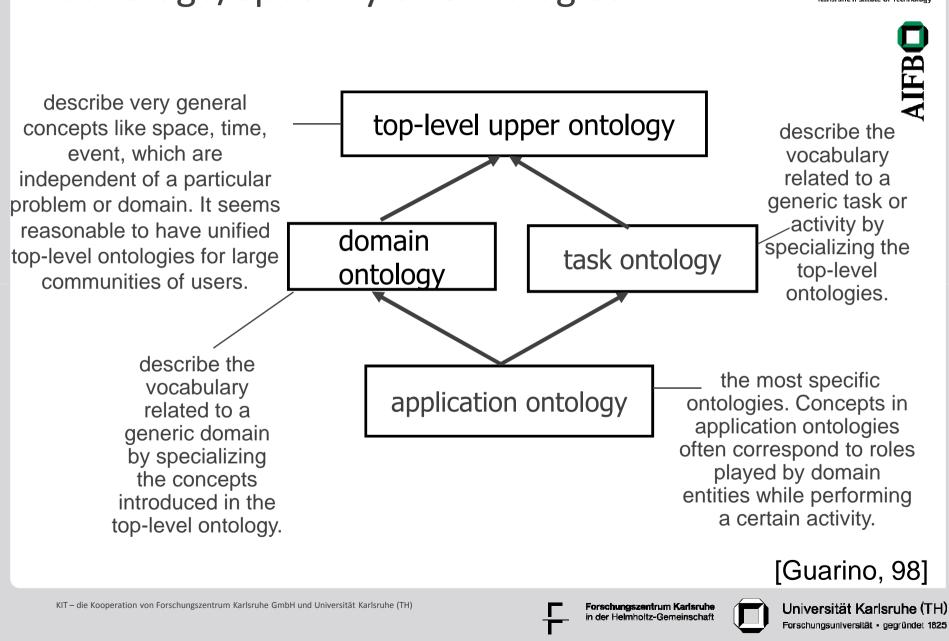
KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft









# Coverage/Specifity of Ontologies



# Specific ontologies

- Domain-oriented
  - Domain-specific
    - Medicine => cardiology => rhythm disorders
  - Domain generalizations
    - components, organs, documents, gene function
- Task-oriented
  - task specific
    - configuration design, instruction, planning, annotation analysis
  - task generalizations
    - problem solving methods





AIFBC

# **Upper Ontologies**

- A.k.a. top level, core, generic or reference
- An attempt to capture the most general and reusable terms and definitions
- "Physical", "Abstract",
  "Structure", "Substance"
- Useful for ontology re-use
- Important when generating or analysing natural language expressions

- Examples of Top Level ontologies
- SUMO
- DOLCE
- CyC
- WordNet
- EuroWordNet





**MFB** 

# Suggested Upper Merged Ontology (SUMO)



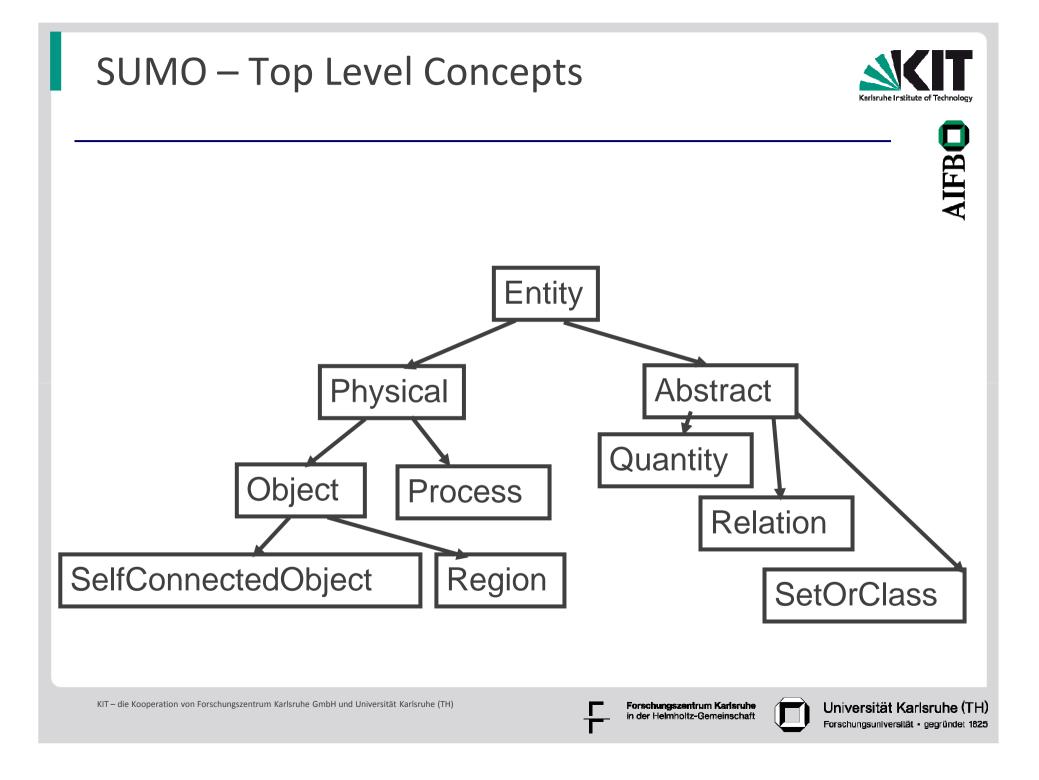
IFB

- 1000 terms, 4000 axioms, 750 rules
- Written in FOL
- Development since 2000
- Associated domain ontologies totaling 20,000 terms and 70,000 axioms

#### Free

- SUMO is owned by IEEE but basically public domain
- Domain ontologies are released under GNU
- www.ontologyportal.org





# Сус



- enCYClopedia
- Douglas Lenat at Cycorp
- Development since 1984
- general knowledge and common-sense reasoning
- Ontology 100,000's of terms
- Millions of assertions
  - "Water is wet"
  - "Everyone has a mother"
  - "When you let go of things they usually fall."
- Open version available opencyc.com





# WordNet

- "Lexical ontology"
- 100,000 word senses synsets
- Created by George Miller's group at Princeton

#### news item IS A KIND OF ...

1 sense of news item

#### Sense 1

#### news item -- (an item in a newspaper)

- => item, point -- (a distinct part that can be specified separately in a group of things that could be enumerated on a list; "he noticed an item in the New York Times"; "she had several items on her shopping list"; "the main point on the agenda was taken up first")
  - => part, portion, component part, component -- (something determined in relation to something that includes it; "he wanted to feel a part of something bigger than himself"; "I read a portion of the manuscript"; "the smaller component is hard to reach")
    - => relation -- (an abstraction belonging to or characteristic of two entities or parts together)
      - => **abstraction** -- (a general concept formed by extracting common features from specific examples)



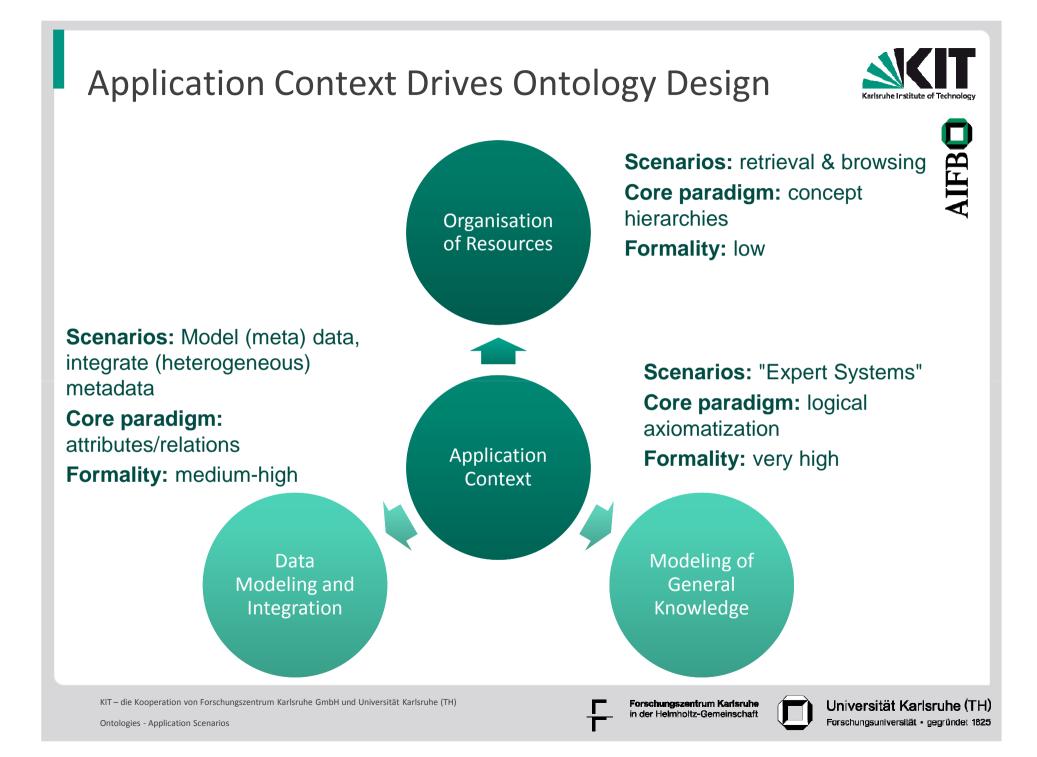
 De facto standard in the linguistics world



AIFB

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft









Modeling of General Knowledge

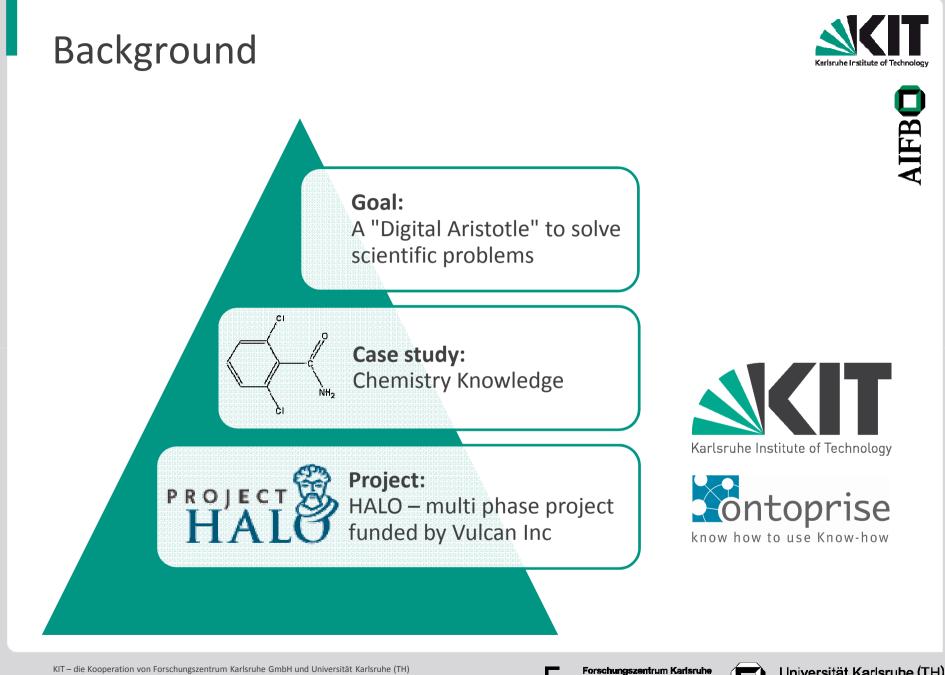
# APPLICATION SCENARIO: MODELING SCIENTIFIC KNOWLEDGE

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

**Ontologies - Application Scenarios** 

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft





**Ontologies - Application Scenarios** 

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft

# Motivation – Competency Question







"Which compounds will produce a gas when HCl is added to the solid compound? HCl is a strong acid producing a yellow-green colored gas above the acid solution?"

**Ontologies - Application Scenarios** 

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft



# Background



#### The "Digital Aristotle" – a system that:

- Encompasses much of the world's scientific knowledge
- Reasons over that knowledge
- Answers novel scientific questions
- Explains these answers
- Is quite ambitous

#### • The "Digital Aristotle" is a multi-stage effort:

- Start with a specific science (Chemistry)
- Challenge with several teams
- Answer AP-style questions
- Refine existing system



# Formalizing questions



#### Example

- Which of the following compounds will produce a gas when HCl is added to the solid compound? HCl is a strong acid producing a yellowgreen colored gas above the acid solution?
  - Ba(OH)<sub>2</sub> (s)
  - CaCO<sub>3</sub> (s)
  - CuSO<sub>4</sub> (s)
  - Na<sub>3</sub>PO<sub>4</sub>(s)
  - NaCl (s)



(every QF1 has (context ((:pair "(a) Ba(OH)2(s)" (a Reaction with (raw-material ((a HCl-Substance)) (a Ba OH 2-Substance with (state ((a State-Value with (value (\*solid))))))))) (:pair "(b) CaCO3(s)" (a Reaction with (raw-material ((a HCl-Substance) (a CaCO3-Substance with (state ((a State-Value with (value (\*solid)))))))))))))) "(c) CuSO4(s)" (a Reaction with (raw-material ((a HCl-Substance) (a CuSO4-Substance with (state ((a State-Value with (value (\*solid))))))))) (:pair "(d) Na3PO4(s)" (a Reaction with (raw-material ((a HCl-Substance) (a Ionic-Compound-Substance with (state ((a State-Value with (value (\*solid))))) (has-basic-structural-unit ((a Ionic-Compound with (nested-atomic-chemicalformula ((a Chemical-Formula with (term ((:seq (:pair 3 Na) (:pair 1 P) (:pair 4 with (raw-material ((a HCl-Substance) (a NaCl-Substance with (state ((a State-Value with (value \*solid)))))))))) (output ((forall (the context of Self) where (oneof2 (the result of (the2 of It)) where ((the value of (the state of It2)) = \*gas)) (the1 of It) (comm [QF1-output-1] Self)))))

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)



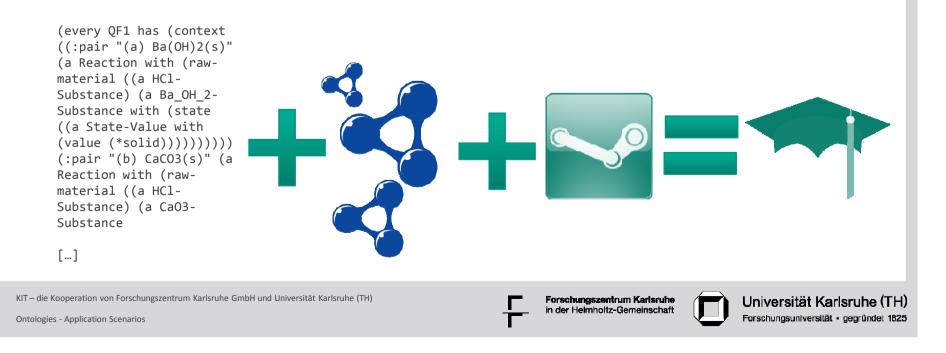
Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft

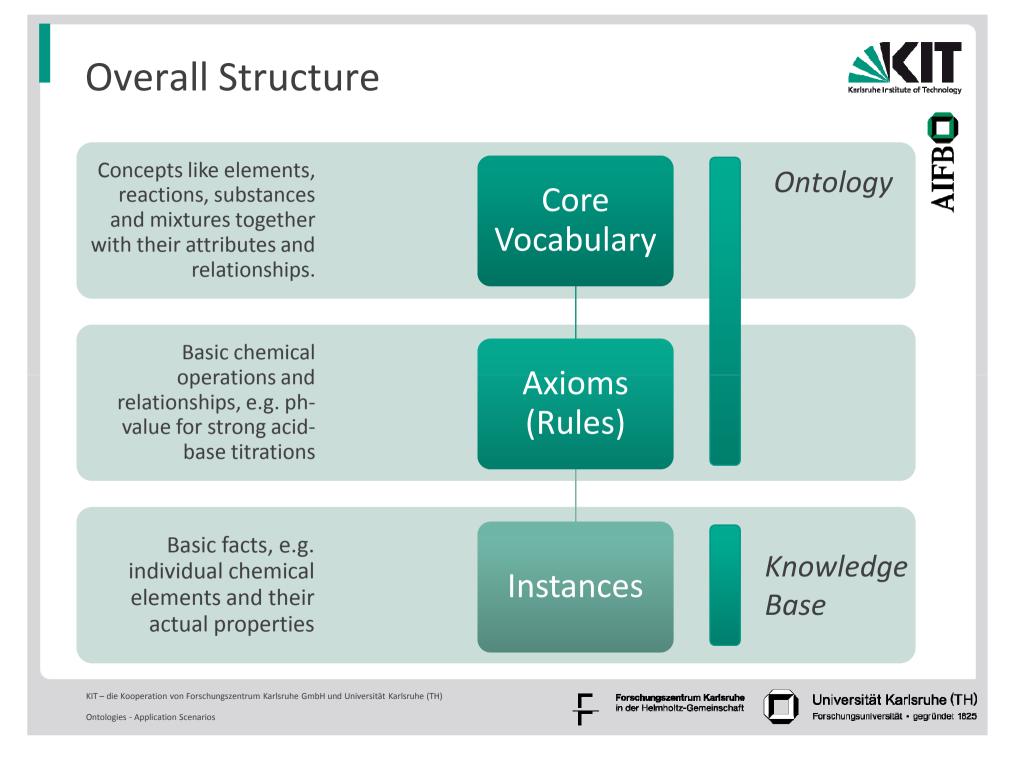


# Background Knowledge

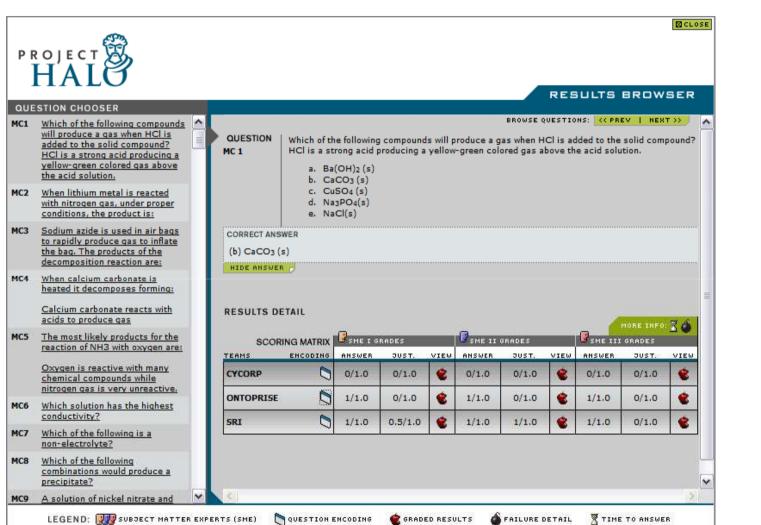


- Formalizing questions is "just" question understandin
- Needs (loads of) background knowledge (ontology)
- Needs a reasoning engine to answer the question using the ontology





### **Result browser**



#### http://www.projecthalo.com/

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

**Ontologies - Application Scenarios** 





Universität Karlsruhe (TH) Forschungsuniversität • gegründet 1825



AIFBI

# Evaluation

#### General Results

- Correctness: high
- Justification: considerably lower than correctness
- Speed: was critical, but all systems faired well

#### AP-Test Results

- Human mean average in this test is AP-2.82
- Project Halo scored an AP-3 they would have passed!





# **Refinement Phase**

#### Task

Gather knowledge faster, cheaper, better...

#### Approach

- Semantic Wiki for Knowledge Capturing
- As a base, Semantic MediaWiki was taken (open source, developed by AIFB)
- Several enhancements to usability of the system (SMW+, developed by ontoprise)



AIFB



# Knowledge Elicitation via Semantic Wikis



ae - Halowiki - Mozilla Firefox - 181 × File Edit View History Bookmarks ScrapBook Tools Help 2 http://halowiki.ontoprise.de/halowiki/index.php/Main Page Main Page - Halowik × Main Page PROIECT discussion annotate edit history page Main Page Semantic Wi This Wiki serves as a demonstration and try-out setup of the Halo Extension @ of the Semantic MediaWiki that has been developed as a part of Project Halo @. There is semantic content available in the domains Biology, Chemistry and Physics, which has been inserted by Subject Matter Experts based on some Taskliste navigation Please Note: Be free to edit and tryout whatever you like in this wiki, but beware of the nightly reset of the wiki's database. All your EDITS WILL BE LOST the next day!!! Main Page Recent changes Quick links to project resources [edit] Random page Semantic Search Please consult the following resources for more information: Properties ■ Overview of the features of the halo-extension d Help ■ Detailed list of features Halo Extension the Halo Extension Featurelist @ search A demonstration video of the extension's main features is available at http://www.ontoprise.de/smwdemo/ & A demowiki with the Halo extension installed is available at http://balowiki.ontoprise.de 🖗 Go Search The software is available in beta status as a tarball via sourcefourge g? Installation instructions are available at Extension Installation @ toolbox = Bugs can be reported in sourceforge bugzilla Mhat links here. Related changes Upload file A list of requested additions, updates, extensions, modifications can be found at [1] Special pages Printable version Permanent link How To Get Started - First Steps [edit] Browse propertie: A basic feature of the Halo Extension is autocompletion. You can use it in nearly every input field and within the textarea in edit mode by pressing Ctrl + Alt + Space after you typed the beginning of a word (e.g. "Hydro") . Edit an article to see the Semantic Toolbar and try out its various functions . Use the Ontology Browser to gain an overview of the wiki ontology. For example, try to filter for Chemical Elements and click the category to gather available information. . Go to the Query Interface and try to compose a query. For example, ask for the category Chemical Elements and their properties Atomic Number and Atomic Weight • The search function in the main toolbar to the left can also be used in order to find annotated knowledge. Use ' "Chemical Elements" "atomic weight" as an example search term (including the quotes) Another useful feature is the Gardening Special. It provides several automated scripts ("bots") that will find anomalies, search for similar entities, check the wiki consistency or materialize template annotations. If you don't want to wait for the scripts to finish, some logs are already available so you can check the output the different bots. ■ Find a complete featurelist and documentation at Ontoworld @ A short video demonstration of the main features can be found at http://www.ontoprise.de/smwdemo/@ [ Powered By MediaW This page was last modified on 25 October 2007, at 08:49. This page has been accessed 2,119 times. Privacy policy About Halowiki Disclaimers 🚵 One active download (2 minutes remaining) 🛛 🐋 Done

#### http://halowiki.ontoprise.de/

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

**Ontologies - Application Scenarios** 

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft



# Knowledge Elicitation via Semantic Wikis

	🖹 😭 📋 http://halowiki.ontoprise.de/halowiki/index.php?title=Bromine&action=edit	<u>ම</u> දු	🔹 👫 LEO de<->en	۵ م		
ogyBrowser - Halo	viki 🗴 Editing Bromine - Halowiki 🔀					
	page discussion annotate edit history		Beryllium > Main Page >	A Log in / c OntologyBrowser > Beryllium		
antic Wiki	Editing Bromine		E Links to Other Page	es		
	Warning: You are not logged in. Your IP address will be recorded in this page's edit history.		🖸 Tools	- Tools		
	B ∠ Ab CA = √n 00 m addeed win be recorded in this page 5 can instary.		Categories			
		•	Annotate Create			
ge changes	== Notable characteristics ==		Chemical elements			
n page	Bromine is the only liquid [[Nonmetal nonmetallic]] element at room temperature and one of five		Halogen			
ic Search es	elements on the period table that are liquid at or close to room temperature. The pure chemical		Properties			
es	element has the physical form of a [[diatomic]] molecule, '''Br <sub>2</sub> '''. It is a heavy, mobile, reddish-brown liquid, that evaporates easily at [[standard temperature and pressure]]s		Annotate Create Has part			
	in a red vapor (its color resembles [[nitrogen dioxide]]) that has a strong disagreeable odor		is less active than	chlorine		
1	resembling that of [[chlorine]]. A [[halogen]], bromine resembles [[is less active than::chlorine]] chemically but is less active. It is more active than [[is more active		is more active than	iodine		
Search	than::iodine]], however. Bromine is slightly [[solubility soluble]] in [[is slightly soluble		is slightly soluble in	water (molecule)		
	in::water (molecule) water]], and highly soluble in [[is highly soluble in::carbon disulfide]] [[aliphatic compound aliphatic]] [[alcohol]]s (such as [[is highly soluble in::methanol]]), an		is highly soluble in	carbon disulfide		
ks here	[[is highly soluble in::acetic acid]]. It [[chemical bond bonds]] easily with [[reacts		is highly soluble in	methanol		
changes	with:=many elements]] and has a strong [[Bleach (chemical) bleach]]ing action.		is highly soluble in	acetic acid		
ad file Bromine is highly reac reacts vigorously with	Bromine is highly reactive and is a powerful [[oxidizing agent]] in the presence of water. It		reacts with	many elements		
	reacts vigorously with [[reacts vigorously with::amine]]s, [[reacts vigorously with::alkene]]s and [[reacts vigorously with::phenol]]s as well as aliphatic and [[aromatic]] [[hydrocarbon]]s,	L	reacts vigorously with	amine		
	[[reacts vigorously with::ketone]]s and [[reacts vigorously with::acid]]s (these are brominated		reacts vigorously with	alkene		
	by either [[Addition reaction]addition]] or [[Substitution]substitution reactions])). With many of the metals and elements, [[anhydrous]] bromine is less reactive than hydrated bromine; however, dry bromine reacts vigorously with [[aluminium]], [[titanium]], [[mercury (element) mercury]] as well as [[alkaline earth metal]]s and [[alkaline metal]]s. Due to its contribution to [[ozone depletion]] in Earth's atmosphere, bromine has been Press Cirl+Alt+Space to use auto-completion. (Cirl+Space in IE) Please note that all contributions to Halowiki may be edited, altered, or removed by other contributors. If you do not want your writing to		reacts vigorously with	phenol		
			reacts vigorously with	ketone		
			reacts vigorously with	acid		
		-1	was discovered by	Antoine Jerome Balard		
		_	was discovered in	1826		
		to be	is found in	bromide salts		
	edited mercilessly, then do not submit it here.		Ontology Browser			
	You are also promising us that you wrote this yourself, or copied it from a public domain or similar free resource (see Halowiki:Copyri details), DO NOT SUBMIT COPYRIGHTED WORK WITHOUT PERMISSION!	Mark a word				
	Summary:		🕒 🕜 Help			
	Sommary.   Save page   Show preview   Show changes   Cancel   Editing help (opens in new window)		Annotation hints			
	Templates used on this page:					
	Template:ChemicalSources					



I

#### http://halowiki.ontoprise.de/

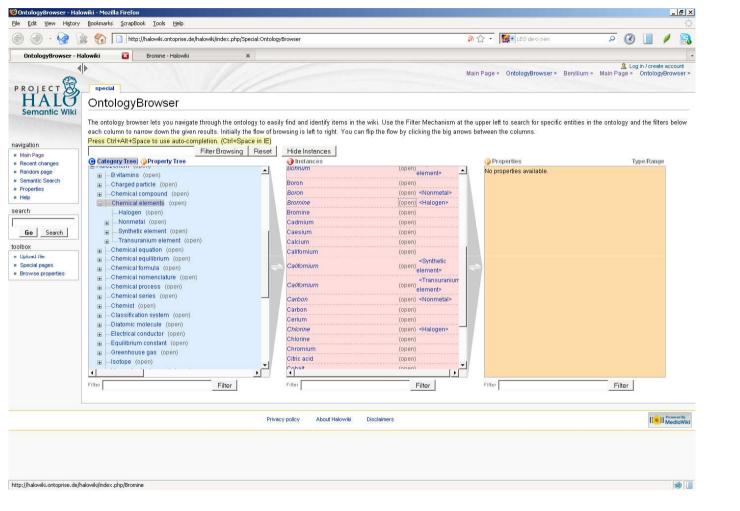
KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

**Ontologies - Application Scenarios** 

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft



# Knowledge Elicitation via Semantic Wikis





# AIFB

#### http://halowiki.ontoprise.de/

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

**Ontologies - Application Scenarios** 

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft







Data Modeling and Integration

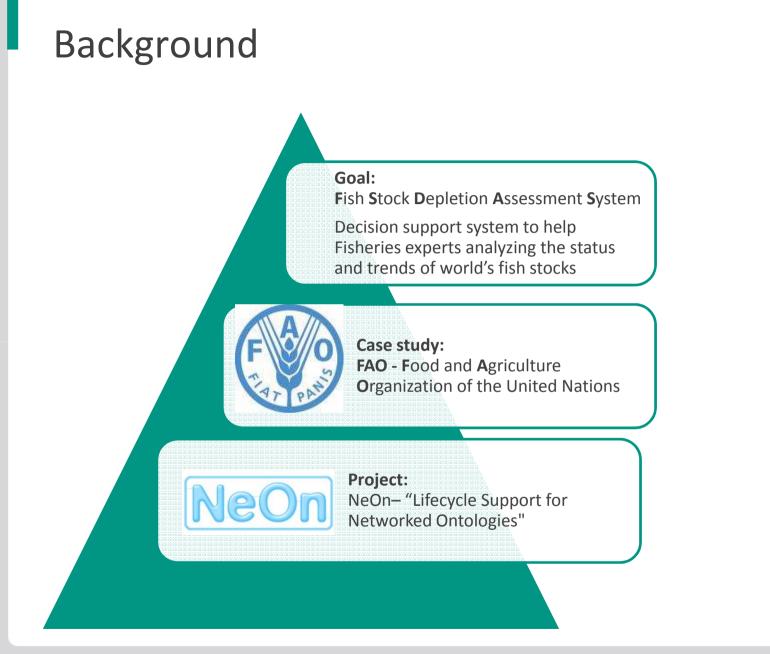
# SCENARIO: OVERFISHING ALERT SYSTEM

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

**Ontologies - Application Scenarios** 

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft





Carlsruhe Institute of Technology

AIFB

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

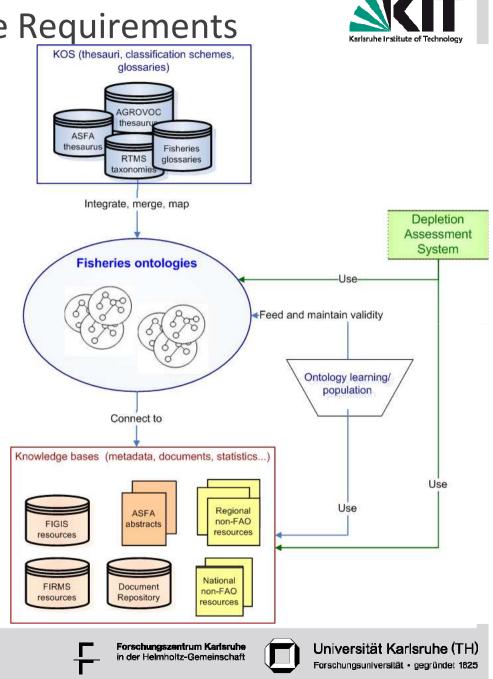
**Ontologies - Application Scenarios** 

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft



# FSDAS – Ontology Runtime Requirements

- Ubiquitous and easy access to
  - status of fish stock
  - factors affecting fish depletion
- Integration and querying of heterogeneous (non-) ontological resources through the exploitation of the Fisheries ontologies
- return relevant results to the client
- Integrate with advanced annotation and visualization tools



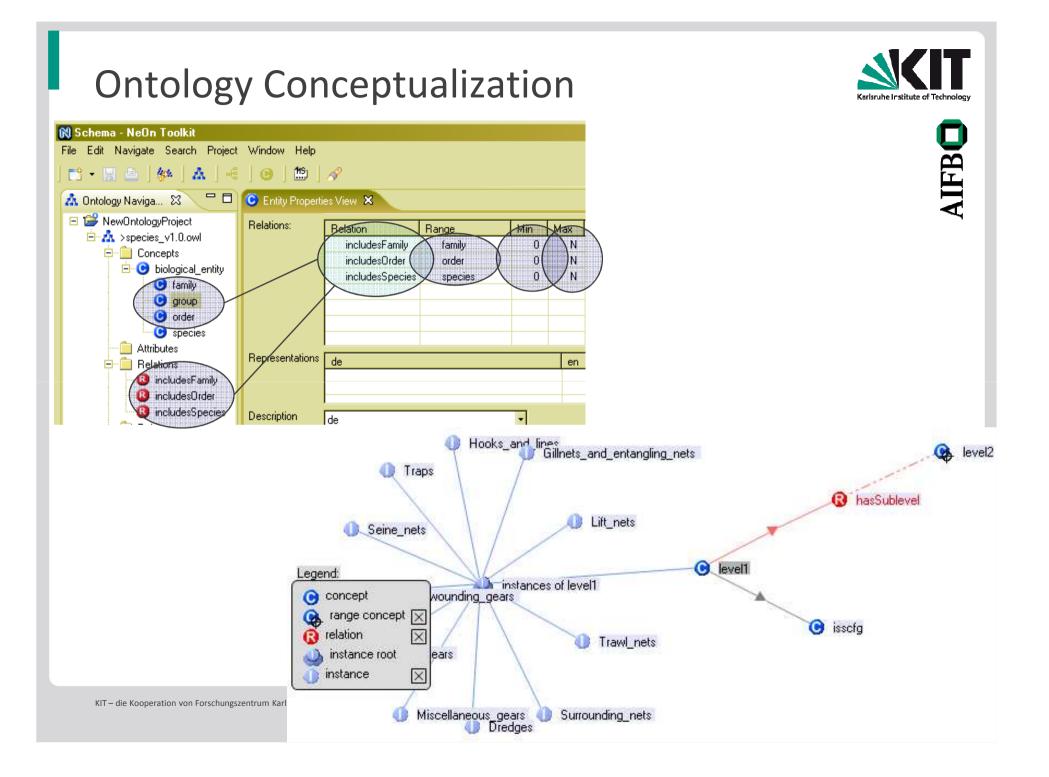
Fisheries Lifecycle Management – Ontology Engineering Requirements

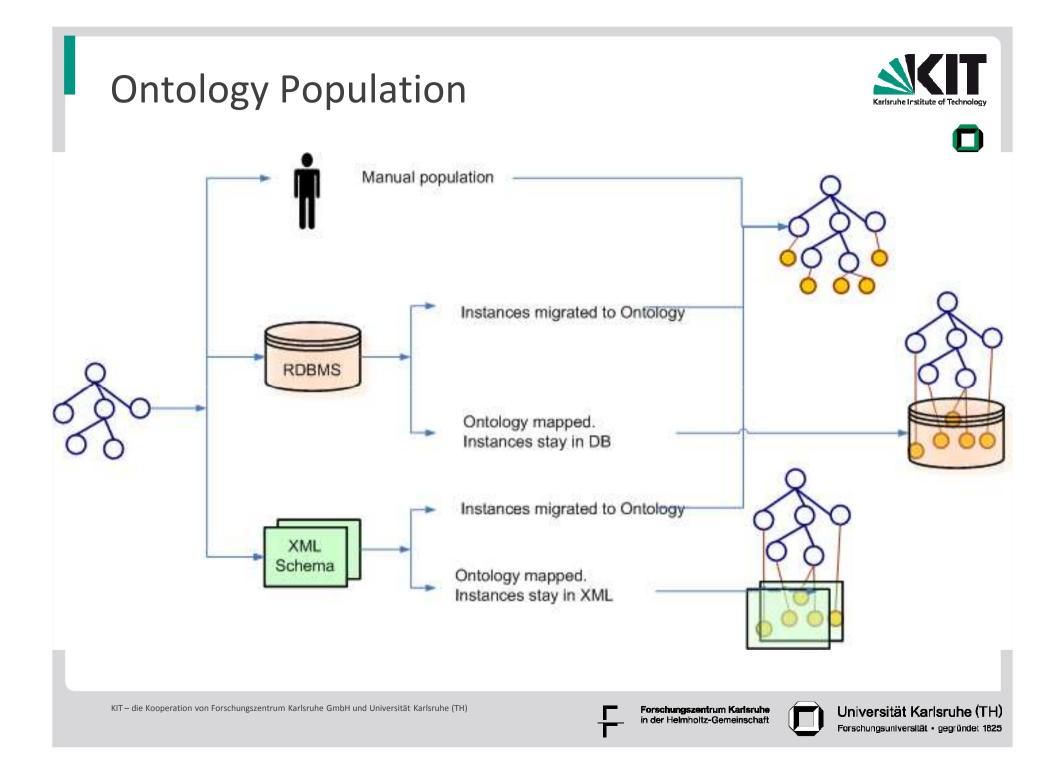


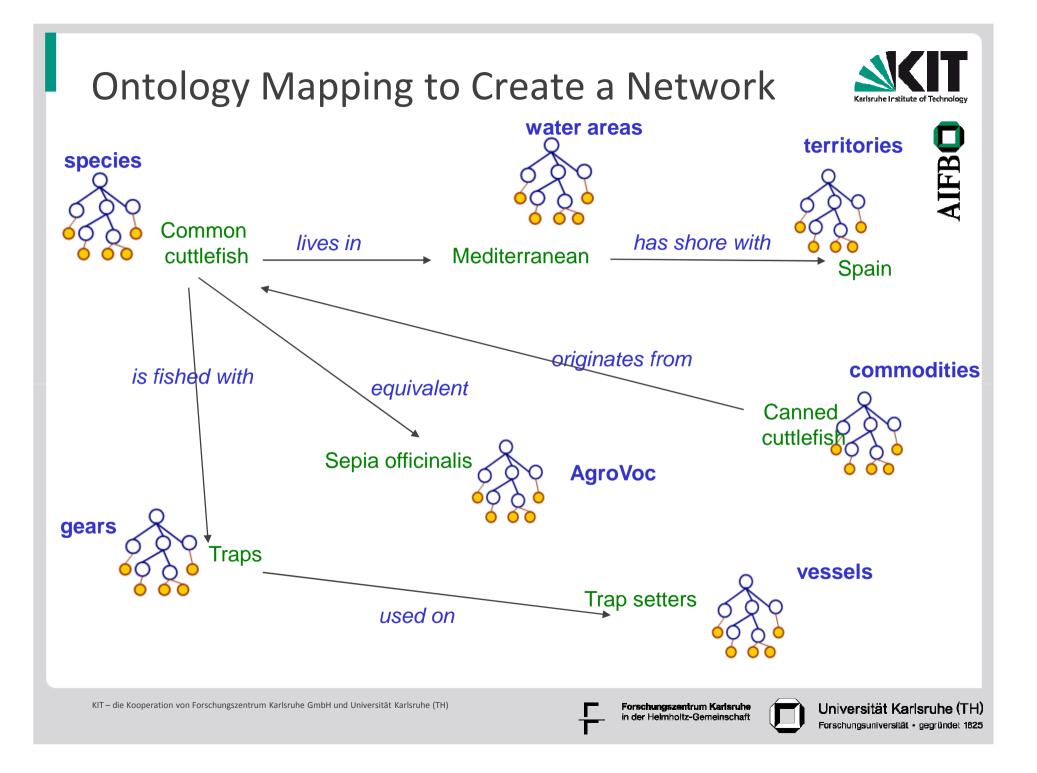
**IFBI** 

- Provide support to ontology engineers and subject experts for:
  - modeling, populating, deploying, versioning ontologies
  - keeping them updated through an editorial workflow
  - managing mappings and relations between them
- Fisheries ontologies are:
  - multilingual ontologies
  - distributed / networked









# Aspects of Engineering Ontology–based Applications

- Engineering of ontology-based applications is a complex task, which involves
  - Ontology engineering
  - Software engineering
  - Ontology management throughout the entire lifecycle

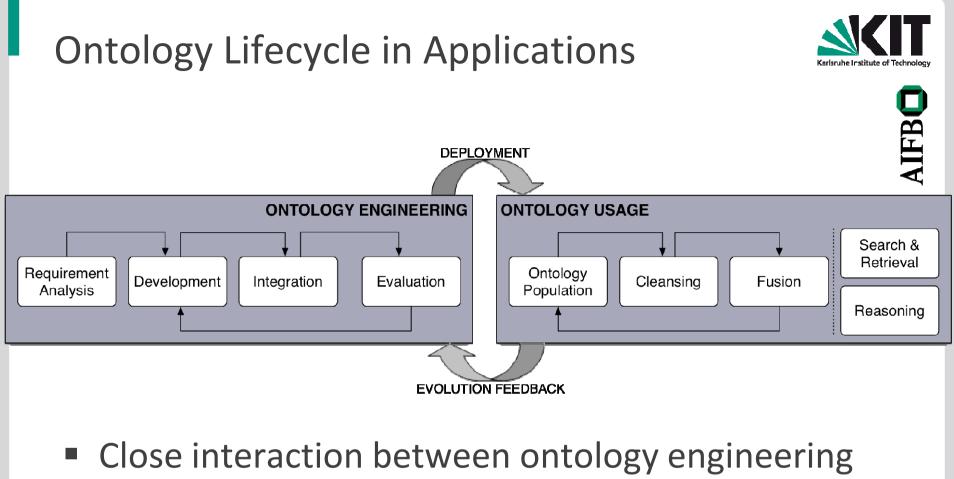
### Require systematic approach to

- the ontology and application development process,
- the ontology life cycle,
- the methods and methodologies for building ontologies and applications,
- and tool support





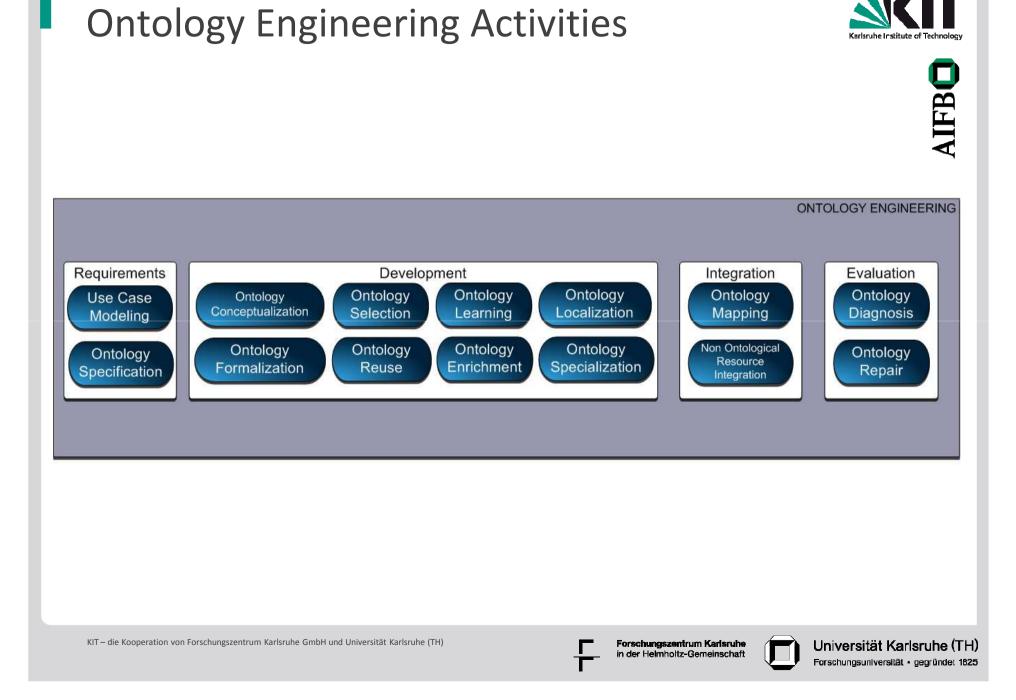




and runtime usage

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft





# How to build an ontology? Initial Steps

- Requirements specification
- Initial lexicon
- Knowledge extraction



AIFB

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

Vorlesung Wissensmanagement WS 2008/2009 - Institut AIFB

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft



# **Requirements Specification (I)**

### Domain and Goal

what is the objective

# Design Guidelines

description of domain in use

# Supported Applications

- brief characteristics of planned application
- specification of system environment

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft



Universität Karlsruhe (TH) Forschungsuniversität • gegründet 1825



AIFB

**Requirements Specification (II)** 

# Knowledge Sources

- types of knowledge sources may be very different
  - domain experts
  - (reusable) ontologies
  - documents / systems
    - dictionaries
    - thesauri
    - product descriptions
    - organisational charts

. . .

- employee role descriptions





KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)



# **Requirements Specification (III)**

- Usage Scenarios (Users and Use cases)
  - describe users/user groups
  - identify stakeholders
  - describe usage scenarios
    - how do they want to use the system?
    - what kind of support do they expect ?
    - use e.g. UML use-case diagrams

### Competency Questions

- define collection of queries that should be supported by the system
- analyze queries to find relevant lexical entries (concepts and relations)
- explore scenarios
- collect competency questionnaire





IFB

# **Competency Questionnaire**

Competency Questionnaire No. 1				
Name: Date:	skill-man-ontology 2001/03/22			
Ontology Engineer: T. Model Domain Ex			pert: X. Pert	
No.	Competency Question	Lexical Entries	Туре	
Q1	Which of our consultants has experience with JAVA programming language?	consultant	concept	
		consultant <i>is a</i> employee	isA relation	
		JAVA	concept	
		programming language	concept	
		JAVA <i>is a</i> program- ming language	isA relation	
		programming language <i>is a</i> skill	isA relation	
		employee <i>has</i> <i>experience with</i> skill	relation	
Q2	What is the salary of a senior programmer? 	salary	concept	
Q3				



AIFB

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)



Universität Karlsruhe (TH) Forschungsuniversität • gegründet 1825

Vorlesung Wissensmanagement WS 2008/2009 - Institut AIFB

T

# Initial lexicon - example



#### lexical entries

Competency Questionnaire No. 1					
Name: Date: Ontology Engineer:T. Mo		skill-man-ontolog 2001/03/22 del <b>Domain Exper</b>	-		
No.	Competency Question	Lexical Entries	Туре		
Q1	Which of our consultants has experience with JAVA programming language?	consultant	concept		
		consultant <i>is a</i> employee	isA relation		
		JAVA	concept		
		programming language	concept		
		JAVA <i>is a</i> programming language	isA relation		
		programming language <i>is a</i> skill	isA relation		
		employee has experience with skill	relation		
Q2	What is the salary of a senior programmer?	salary	concept		
Q3					

potential concepts	potential relations
Consultant Employee JAVA Programming language Experience Skill Programmer Project Customer Industry	HasExperienceWith WorksIn Contains 

AIFB

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

Vorlesung Wissensmanagement WS 2008/2009 - Institut AIFB



# **Knowledge Extraction**

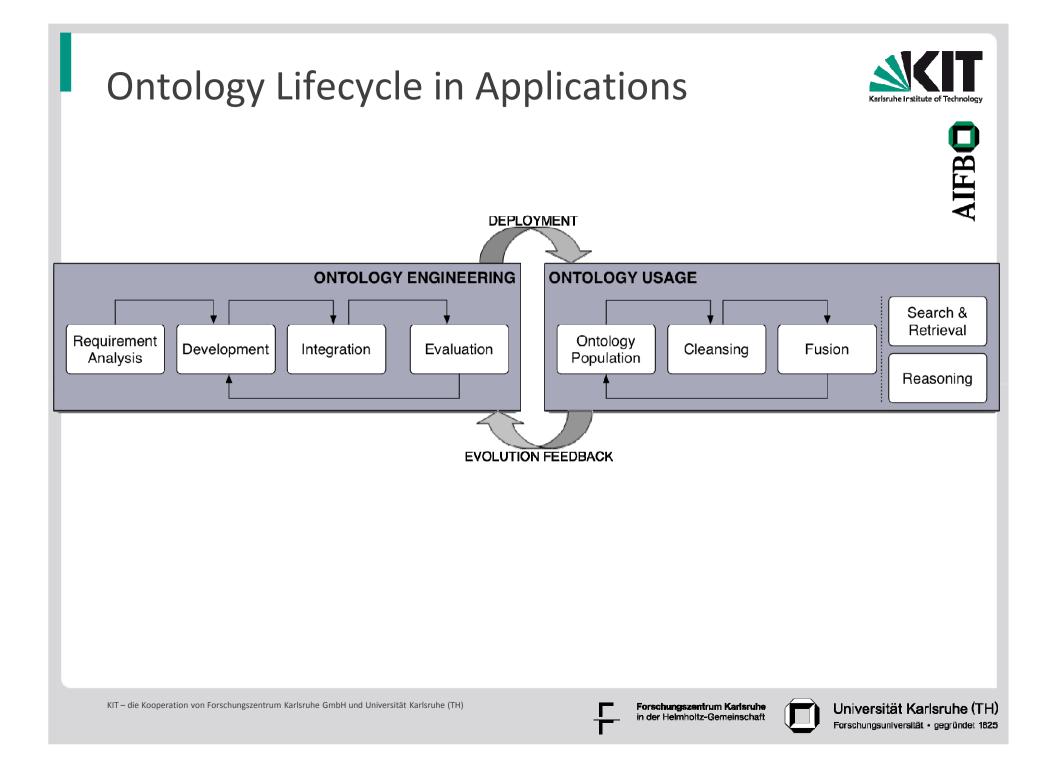


- top-down (modeling concepts and relationships on a very generic level)
- bottom-up (relevant concepts are extracted semiautomatically from available sources)
- middle-out (identify the most important concepts which will then be used to obtain the remainder of the hierarchy by generalization and specialization)

KIT – die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft

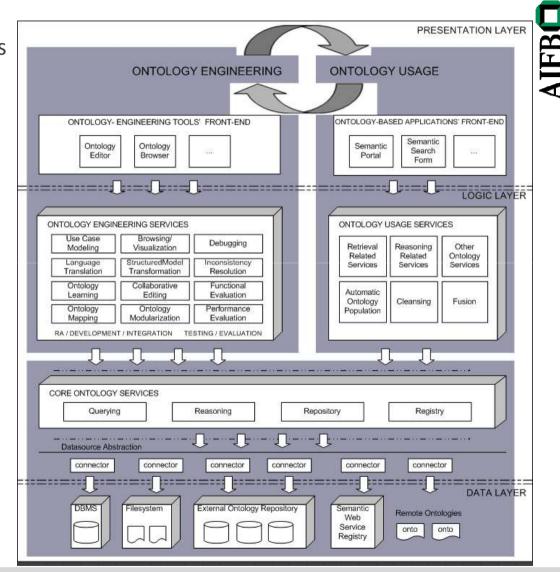




# Generic Architecture with Lifecycle Support

 Inspired by SE best practices (SOA, JEE)

- Dynamic interaction of engineering and usage activities
- Layered Organization
- Presentation Layer
  - Thin client vs. Rich client
- Logic Layer
  - Business services / objects
  - Ontology services
- Data Layer
  - Ontological sources
  - Non-ontological sources



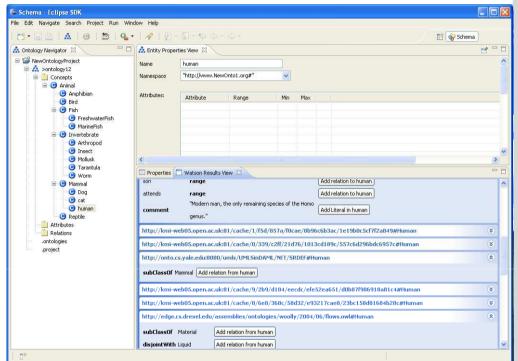


Karlsruhe Institute of Technolog

# NeOn Toolkit



- Infrastructure and reusable software components
- Based on Eclipse platform
- Based on a set of standardized APIs
- Plugins support lifecycle activities



#### http://www.neon-toolkit.org/

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft





# Summary



**MFB(** 

- Wide spectrum of types of ontologies
- Wide spectrum of ontology-based applications
  - Organisation of resources
  - Data integration
  - Modeling of generic knowledge
- Methods, methodologies and tools for engineering ontologies and ontology-based applications

