

# Beyond RDF(S): The Ontology Perspective for the Semantic Web

*Rudi Studer*

*Institute AIFB, University of Karlsruhe,  
& ontoprise GmbH, Karlsruhe*

*Semantic Web Technologies Workshop  
Luxembourg, November 22-23, 2000*

*contributions of Michael Erdmann, Alexander Maedche &  
Steffen Staab are gratefully acknowledged*

# Agenda

- **Motivation**
- Web-based Ontologies
- Ontology Engineering and Maintenance
- Semantic Web Infrastructure
- Metadata Management
- Semantic Web Portals
- Conclusion

Motivation

Web-based  
Ontologies

Engineering/  
Maintenance

Infrastructure

Metadata  
Management

Portals

Conclusion

# 1. Motivation

## Semantic Web: The Vision

- WWW is an impressive success, cf.
  - amount of available information
  - number of human users
  
- However: WWW is currently only for human readers, but ...

*... Imagine what computers can understand when there is a vast tangle of interconnected terms and data that can automatically be followed.” (Tim Berners-Lee, Weaving the Web, 1999)*

# Web Communities

Motivation

Web-based Ontologies

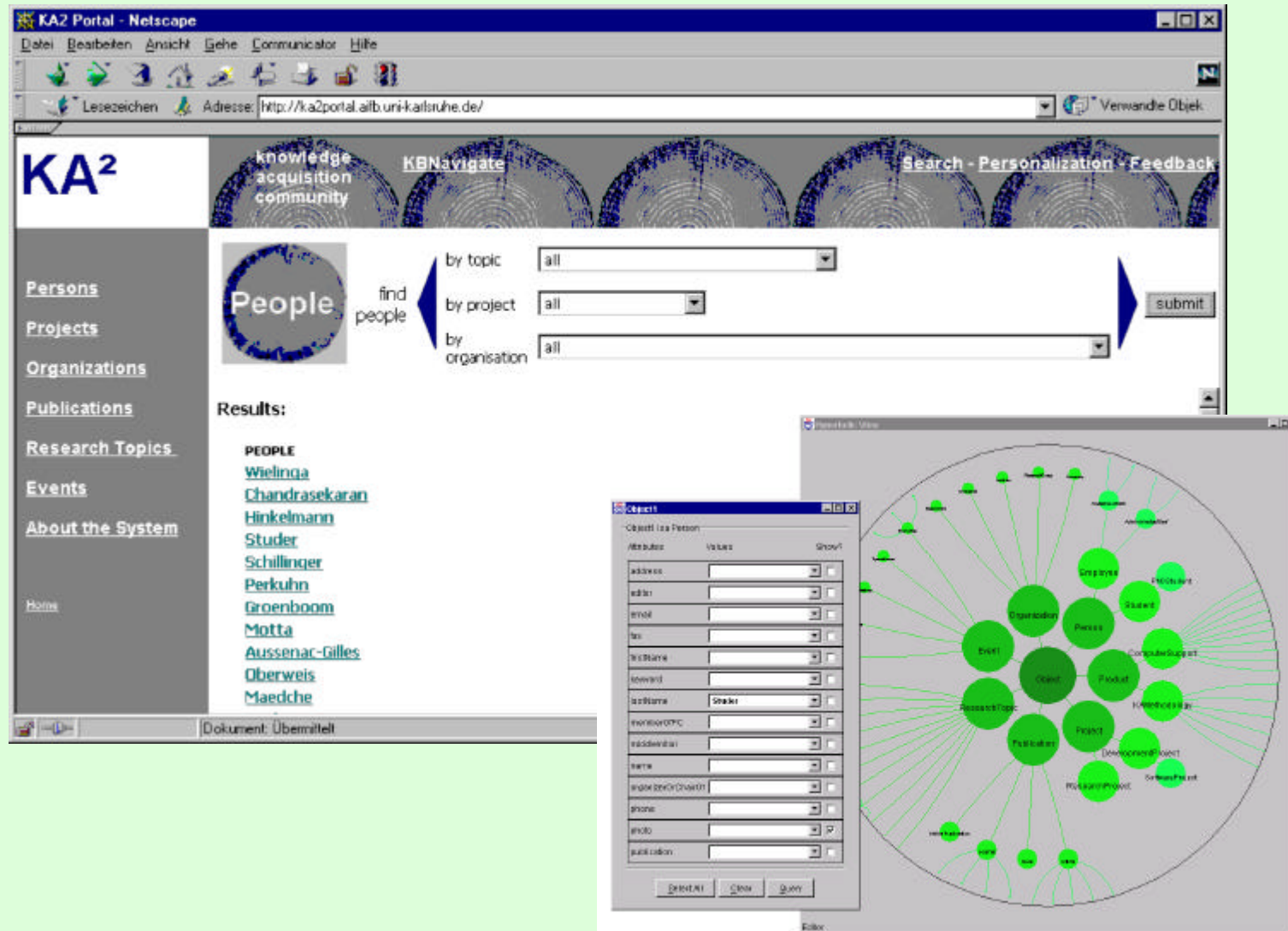
Engineering/  
Maintenance

Infrastructure

Metadata  
Management

Portals

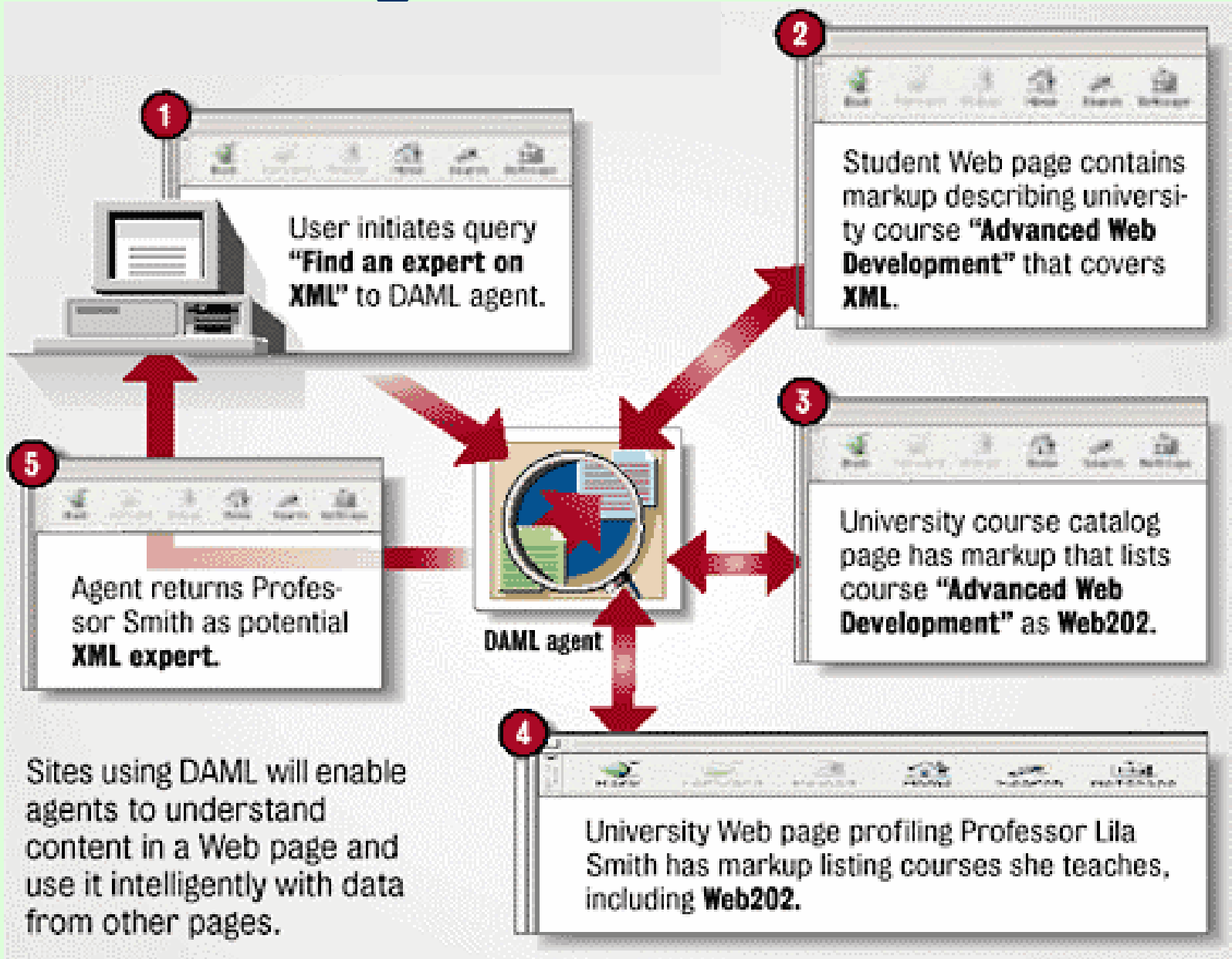
Conclusion



The screenshot displays the KA2 Portal interface in a Netscape browser window. The page features a navigation menu on the left with categories like Persons, Projects, Organizations, Publications, Research Topics, Events, and About the System. The main content area includes a search section titled 'People' with filters for 'by topic', 'by project', and 'by organisation', all set to 'all'. Below the search filters, a list of names is shown under the heading 'Results: PEOPLE', including Wielinga, Chandrasekaran, Hinkelmann, Studer, Schillinger, Perkuhn, Groenboom, Motta, Aussenar-Gilles, Oberweis, and Maedche. An inset window titled 'Object' shows a form for entering person details. In the bottom right corner, a network graph is visible, consisting of numerous green circular nodes connected by lines, representing relationships between individuals or entities within the community.

# Agents on the Web

- Motivation
- Web-based Ontologies
- Engineering/Maintenance
- Infrastructure
- Metadata Management
- Portals
- Conclusion



# XML is not the Solution

- Meaning of XML-Documents is *intuitively* clear
  - due to "*semantic*" Mark-Up
  - tags are domain-terms
- But, computers do not have intuition
  - tag-names *per se* do not provide semantics
- DTD does not distinguish between objects and relations
- XML lacks a semantic model
  - has only a "surface model", i.e. tree

# Outline for the Semantic Web Vision

## Motivation

## Web-based Ontologies

## Engineering/Maintenance

## Infrastructure

## Metadata Management

## Portals

## Conclusion

- Ontologies as key-enablers of this vision
- In this talk we outline technologies for
  - Representing,
  - Building,
  - Using and,
  - Applyingontologies on the Web

# Agenda

- Motivation
- **Web-based Ontologies**
- Ontology Engineering and Maintenance
- Semantic Web Infrastructure
- Metadata Management
- Semantic Web Portals
- Conclusion

Motivation

Web-based  
Ontologies

Engineering/  
Maintenance

Infrastructure

Metadata  
Management

Portals

Conclusion



## 2. Web-based Ontologies

- **What is a (heavy-weight) ontology?**
  - Describes a **formal shared conceptualization** of a particular domain of interest
  - Describes
    - **concepts** that are relevant for the domain
    - **relations** between concepts
    - **axioms** about these concepts and relations
  - Enforces a well-defined **semantics** on such a conceptualization

# Different Types of Ontologies

- **Light-weight Ontology**
  - concepts, atomic types
  - is-a hierarchy among concepts
  - relations between concepts
- **Heavy-weight Ontology**
  - cardinality constraints
  - taxonomy of relations
  - reified statements
  - Axioms / semantic entailments of various tastes
    - expressiveness (description logics, propositional, horn, first order logic, higher order)
    - inference systems

# Layered Languages

The W3C hierarchy of languages:



Motivation

Web-based  
Ontologies

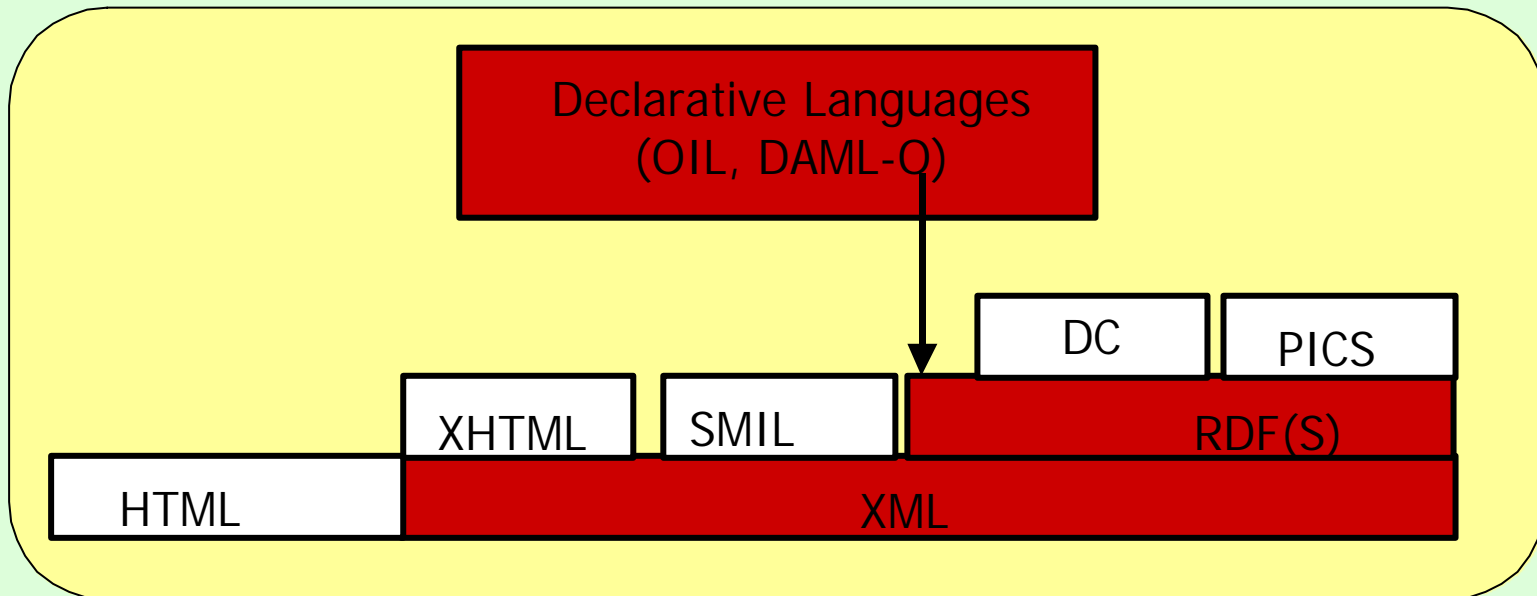
Engineering/  
Maintenance

Infrastructure

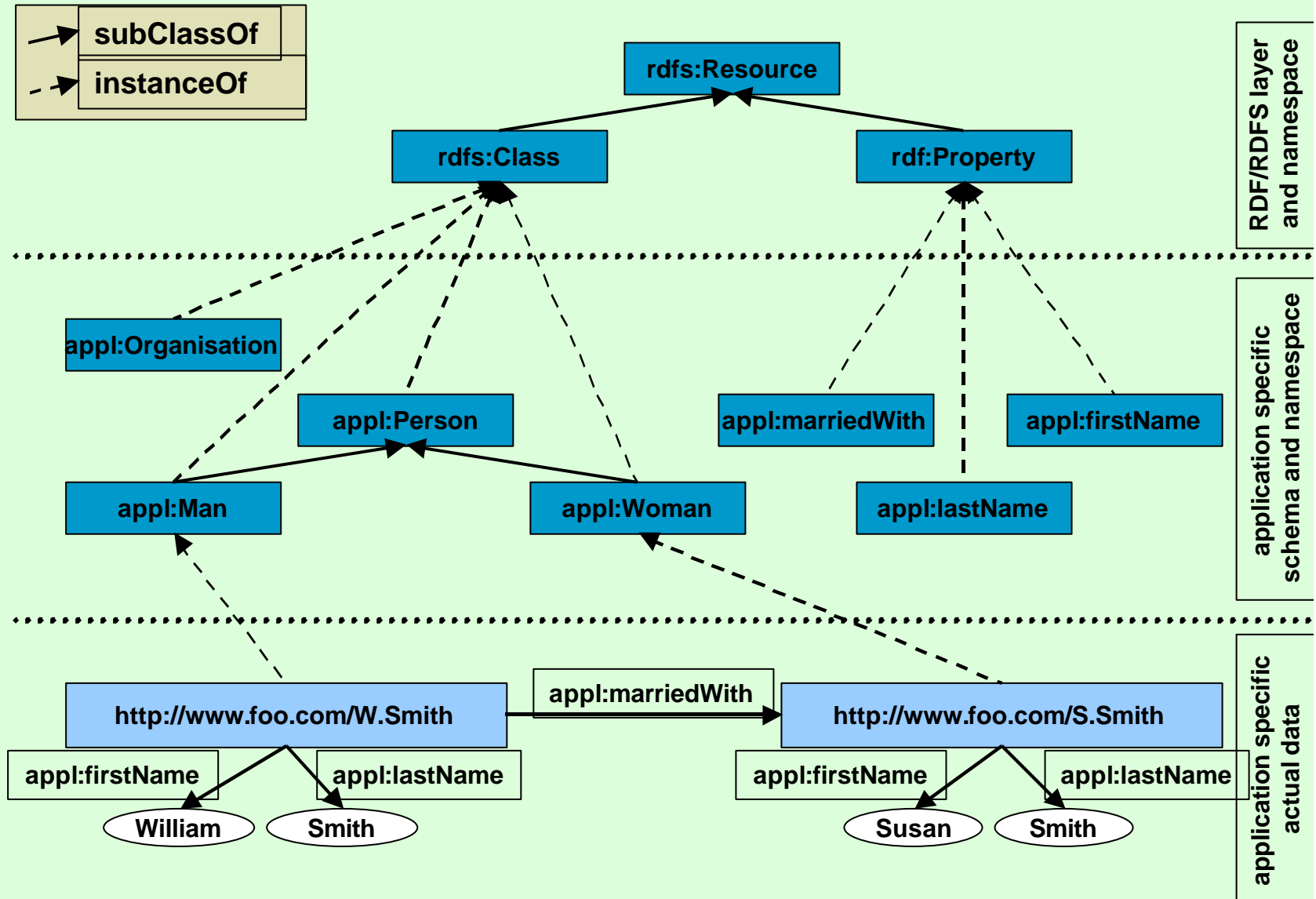
Metadata  
Management

Portals

Conclusion



# RDF Data Model



# Characteristics of RDF(S)

- RDF commits to small set of modeling primitives
- RDF does **not** commit to domain vocabulary
- RDF Schema
  - extended set of modeling primitives
    - class, subclassof, type
    - property, subpropertyof
    - domain, range
  - enables definition of a **domain vocabulary** and embedding into an is-a hierarchy
- But: RDF(S) is not expressive enough !!

# OIL: Ontology Inference Layer

- Main features of OIL
  - OIL combines modeling primitives from frame-based languages and description logic
  - OIL is an **extension of RDFS**, thus Web compatible
  - Result from the IST project On-To-Knowledge: <http://www.ontoknowledge.org>



*Content-driven Knowledge Management Tools through Evolving Ontologies*

Motivation

Web-based Ontologies

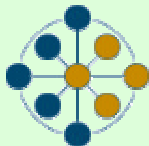
Engineering/Maintenance

Infrastructure

Metadata Management

Portals

Conclusion



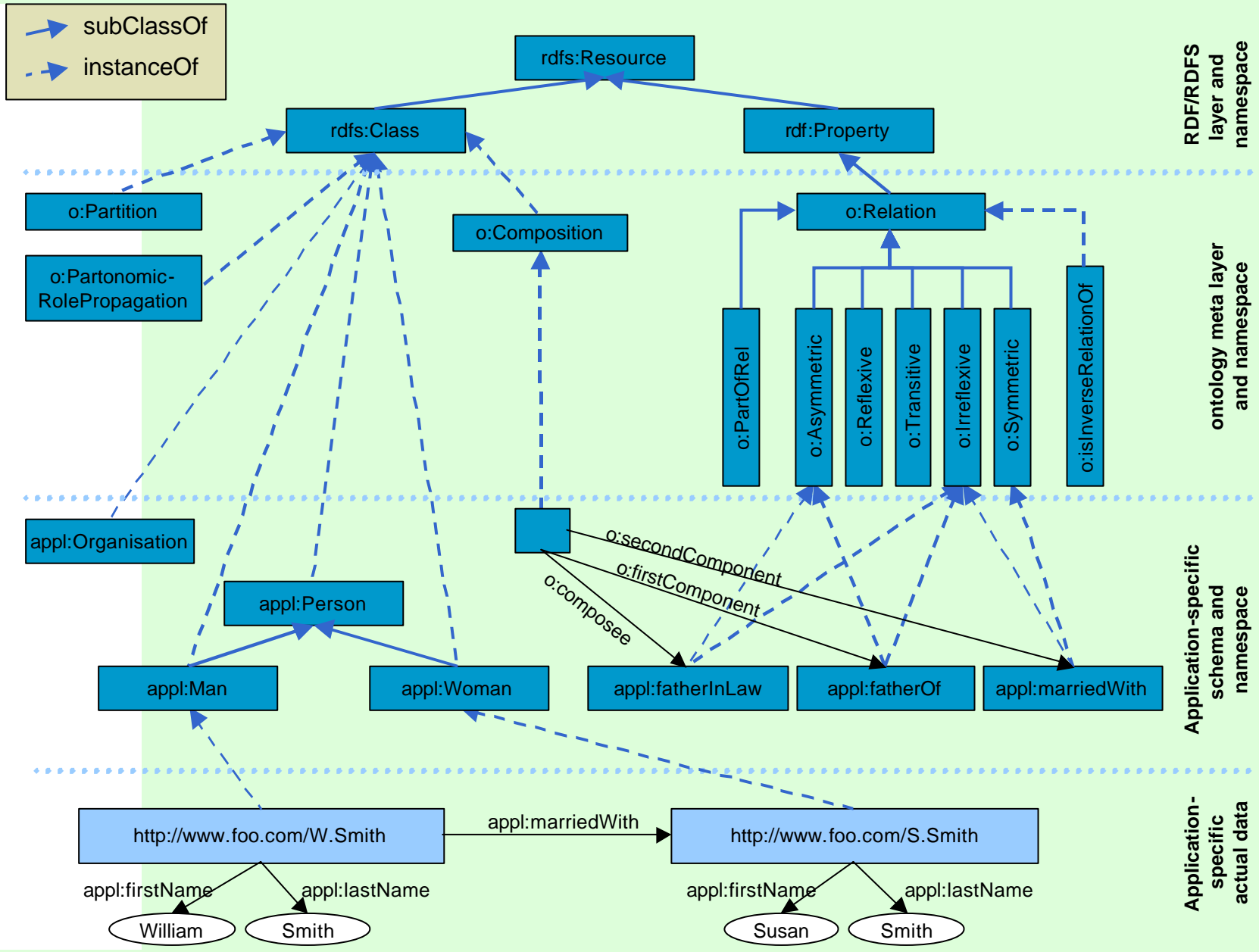
# Switching between Languages

- Different layers of languages on top of the core data model RDF(S) will be built
- We propose **semantic patterns** as
  - a means to communicate knowledge at an epistemological level
  - a means for partial execution by any particular implementation of any representation language
- Semantic patterns combine advantages of formal specification methods with design patterns

# Semantic Pattern Libraries

- Axioms for a relational algebra
  - Reflexivity of relations
  - Transitivity of relations
  - Inverse relations
- Composition of relations
- (Exhaustive) Partitions
- Axioms for subrelation relationships
- Axioms for part-whole reasoning
- Nonmonotonicity
- Axioms for temporal and modal contexts
- ...





# Agenda

- Motivation
- Web-based Ontologies
- **Ontology Engineering and Maintenance**
- Semantic Web Infrastructure
- Metadata Management
- Semantic Web Portals
- Conclusion

# 3. Ontology Engineering and Maintenance

## Tools:

- **Light-weight**
  - uncontroversial
  - all tools support light-weight ontologies
    - Protégé-2000 KA Environment, Stanford University
    - OntoEdit Web Ontology Workbench, University of Karlsruhe
    - UML-Tools
- **Heavy-weight**
  - no consensus yet
  - layering seems appropriate/necessary

Motivation

Web-based  
Ontologies

Engineering/  
Maintenance

Infrastructure

Metadata  
Management

Portals

Conclusion

# Ontology Engineering using OntoEdit

- Interaction with the user on an **epistemological level**
- Multiple, multilingual views for concepts, relations **and axioms**
- Flexible import / export of ontologies (**incl. axioms**) into several representation languages (F-Logic, OIL, DAML-ONT, SQL-2, ...)
- Access web ontology instances via RDF-crawler
- Access different inference engines for consistency checking and application debugging
- Linkable to natural language via domain lexicon

# OntoEdit Workbench

Motivation

Web-based  
Ontologies

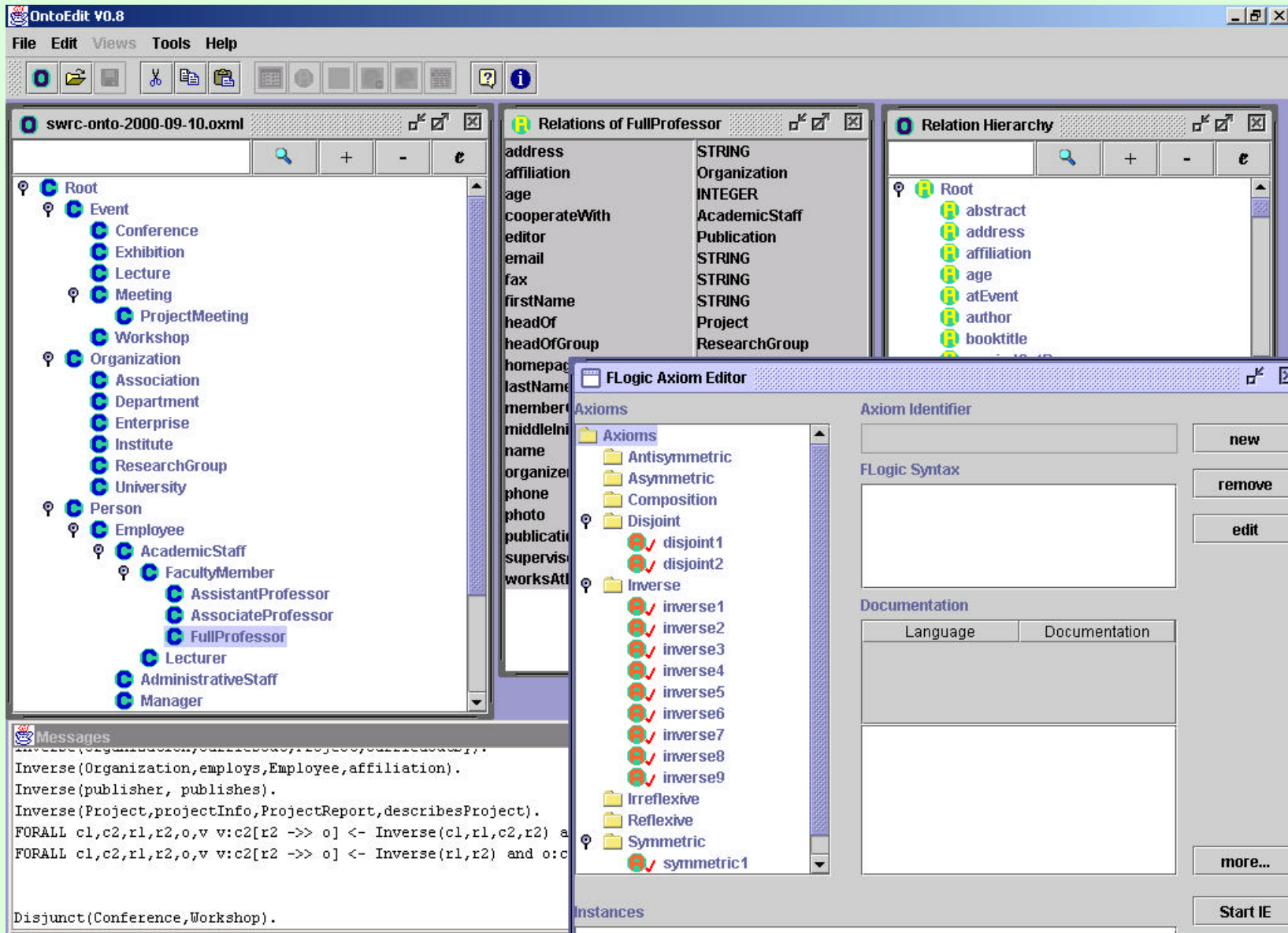
Engineering/  
Maintenance

Infrastructure

Metadata  
Management

Portals

Conclusion



The screenshot displays the OntoEdit V0.8 application window. The interface is divided into several panes:

- Left Pane (swrc-onto-2000-09-10.xml):** A hierarchical tree view of the ontology. The 'Person' class is expanded, showing subclasses like 'AcademicStaff', 'FacultyMember', 'AssistantProfessor', 'AssociateProfessor', 'FullProfessor', 'Lecturer', 'AdministrativeStaff', and 'Manager'.
- Center Pane (Relations of FullProfessor):** A table listing properties and their domains.
 

address	STRING
affiliation	Organization
age	INTEGER
cooperateWith	AcademicStaff
editor	Publication
email	STRING
fax	STRING
firstName	STRING
headOf	Project
headOfGroup	ResearchGroup
homepag	
lastName	
member	
middleIni	
name	
organize	
phone	
photo	
publicati	
supervis	
worksAtt	
- Right Pane (Relation Hierarchy):** A tree view showing the hierarchy of relations, including 'abstract', 'address', 'affiliation', 'age', 'atEvent', 'author', and 'booktitle'.
- Bottom Pane (FLogic Axiom Editor):** A workspace for defining axioms. It includes a list of axiom types (Antisymmetric, Asymmetric, Composition, Disjoint, Inverse, Irreflexive, Reflexive, Symmetric) and a text area for entering FLogic syntax. The 'Axiom Identifier' and 'Documentation' sections are also visible.
- Bottom Left (Messages):** A log window showing messages such as 'Inverse(Organization,employs,Employee,affiliation).', 'Inverse(publisher, publishes).', and 'Disjunct(Conference,Workshop).'.

# Ontology Learning

- Millions of ontologies will be built
- Ontology Engineering is difficult and time-consuming
- Idea: Apply Machine Learning to Ontology Engineering
- Build the ontology in an application-oriented way, based on existing resources  
=> “Reverse Engineering“

# Ontology Learning

Motivation

Web-based  
Ontologies

Engineering/  
Maintenance

Infrastructure

Metadata  
Management

Portals

Conclusion

## Ontology Engineering by Means of Ontology Learning

### Ontology Creation

- start from scratch
- built on top of existing structures
- integrate existing knowledge sources

### Ontology Maintenance

- Update parts of the ontology
- Ontology Pruning
- Ontology Enrichment & Refinement

# Relation Mining [Maedche, Staab 00]

Motivation

Web-based Ontologies

Engineering/Maintenance

Infrastructure

Metadata

Conclusion

TK-NEWS.de - Der Newsticker rund um die Welt der Telekommunikation - Netscape

Location: http://www.tk-news.de/cgi-local/news/nachricht.cgi?nr=565

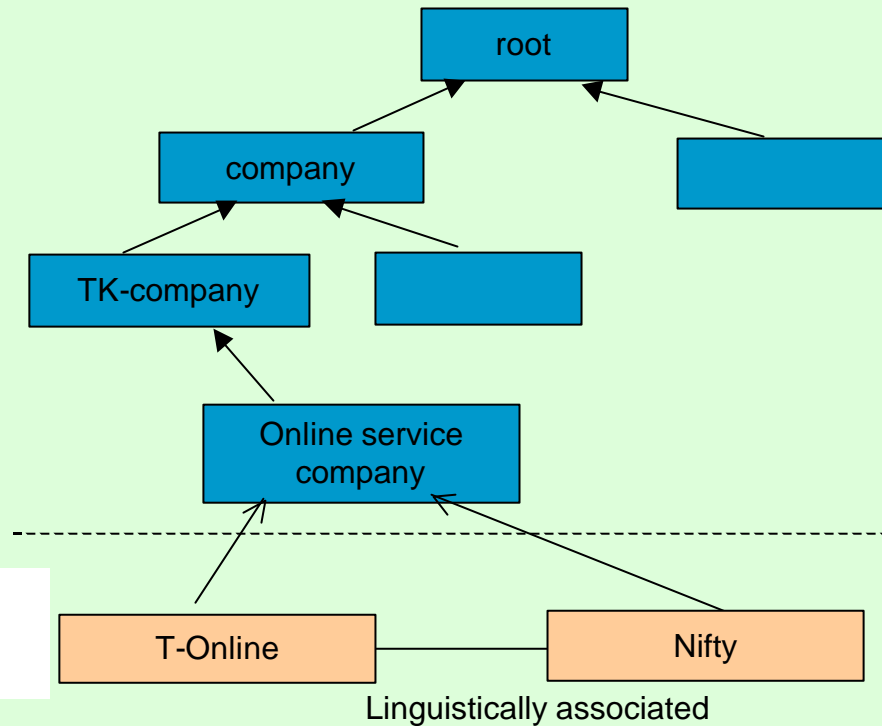
**TK-NEWS.de**

**T-Online kooperiert mit japanischem Onlinedienst Nifty**

T-Online-Abo... können künftig Angebote von Nifty nutzen und umgekehrt.

Nifty gehört zum japanischen Computerhersteller Fujitsu Ltd. und hat in Japan 3,9 Millionen Abonnenten. Die Telekom-Tochter T-Online ist mit sechs Millionen Abonnenten der größte Onlinedienst Europas und der zweitgrößte der Welt. Gemeinsam wollen beide ihre Position gegen den amerikanischen Weltmarktführer AOL stärken, der 24 Millionen Abonnenten zählt.

Nifty und T-Online bauen ihre internationalen Allianzen aus. Erst kürzlich hat Nifty mit den größten Onlinediensten Hongkongs und



**Generate suggestion:**

relation(company, company)

=> cooperateWith(company, company)



# Evolution of Ontologies

- Real world is changing all the time:
  - new businesses
  - new organizational structures in enterprises
  - new products and services
  - ...
- Ontologies have to reflect these changes
  - new concepts and relations
  - new meanings of concepts
  - concepts and relationships become unnecessary
- Maintenance of ontologies is essential
  - ontology-based applications depend on up-to-date ontologies

# Agenda

- Motivation
- Web-based Ontologies
- Ontology Engineering and Maintenance
- **Semantic Web Infrastructure**
- Metadata Management
- Semantic Web Portals
- Conclusion

# Semantic Web Infrastructure

- Infrastructural requirements:
  - Scalable RDF Repositories (all is built on top of the same data model !)
  - Scalable reasoning services for different languages
  - Resource-ID Management
  - Versioning of ontologies and corresponding metadata
  - Clients: Engineering & Maintenance

Motivation

Web-based  
Ontologies

Engineering/  
Maintenance

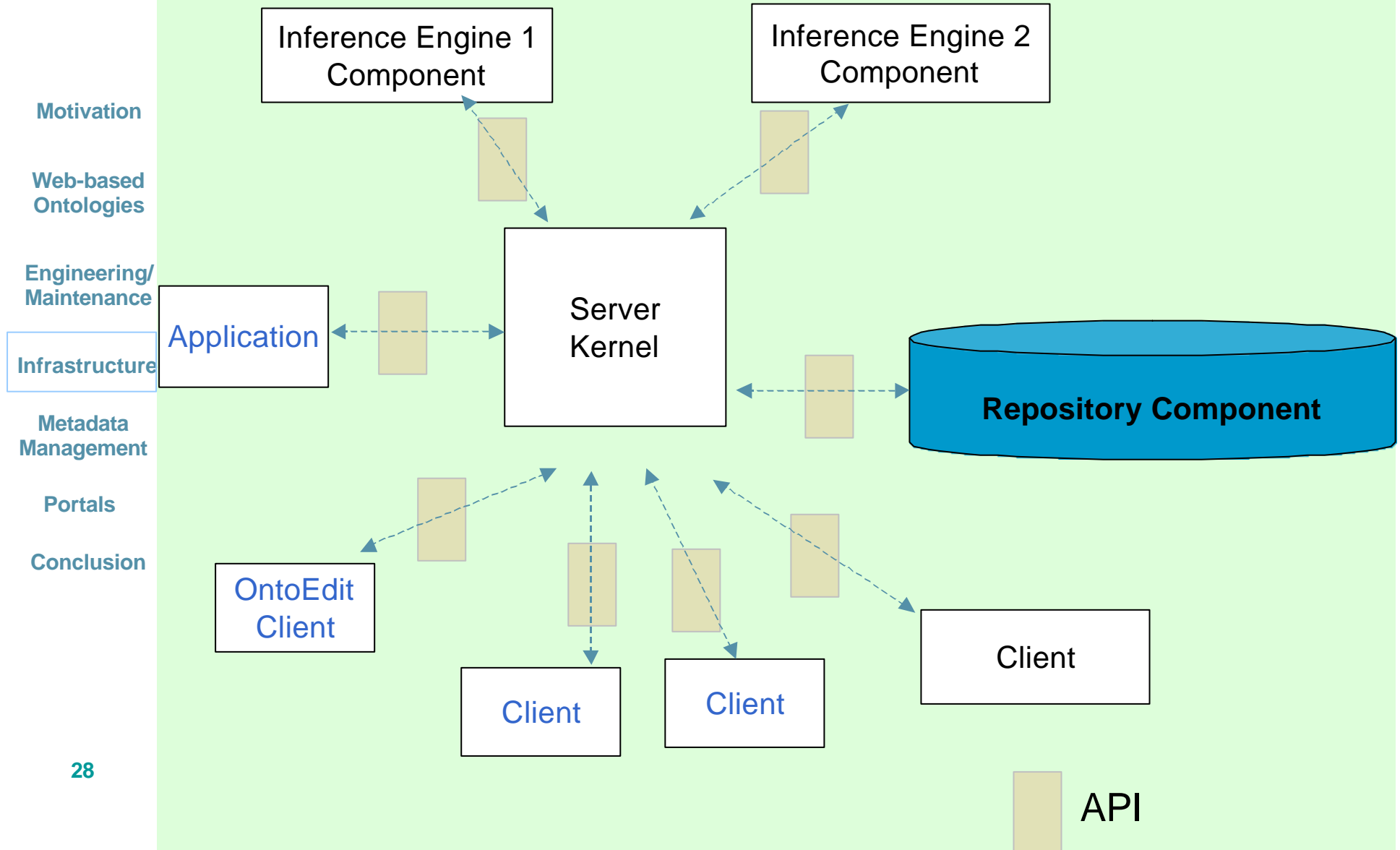
Infrastructure

Metadata  
Management

Portals

Conclusion

# OntoServer – Basic Infrastructure



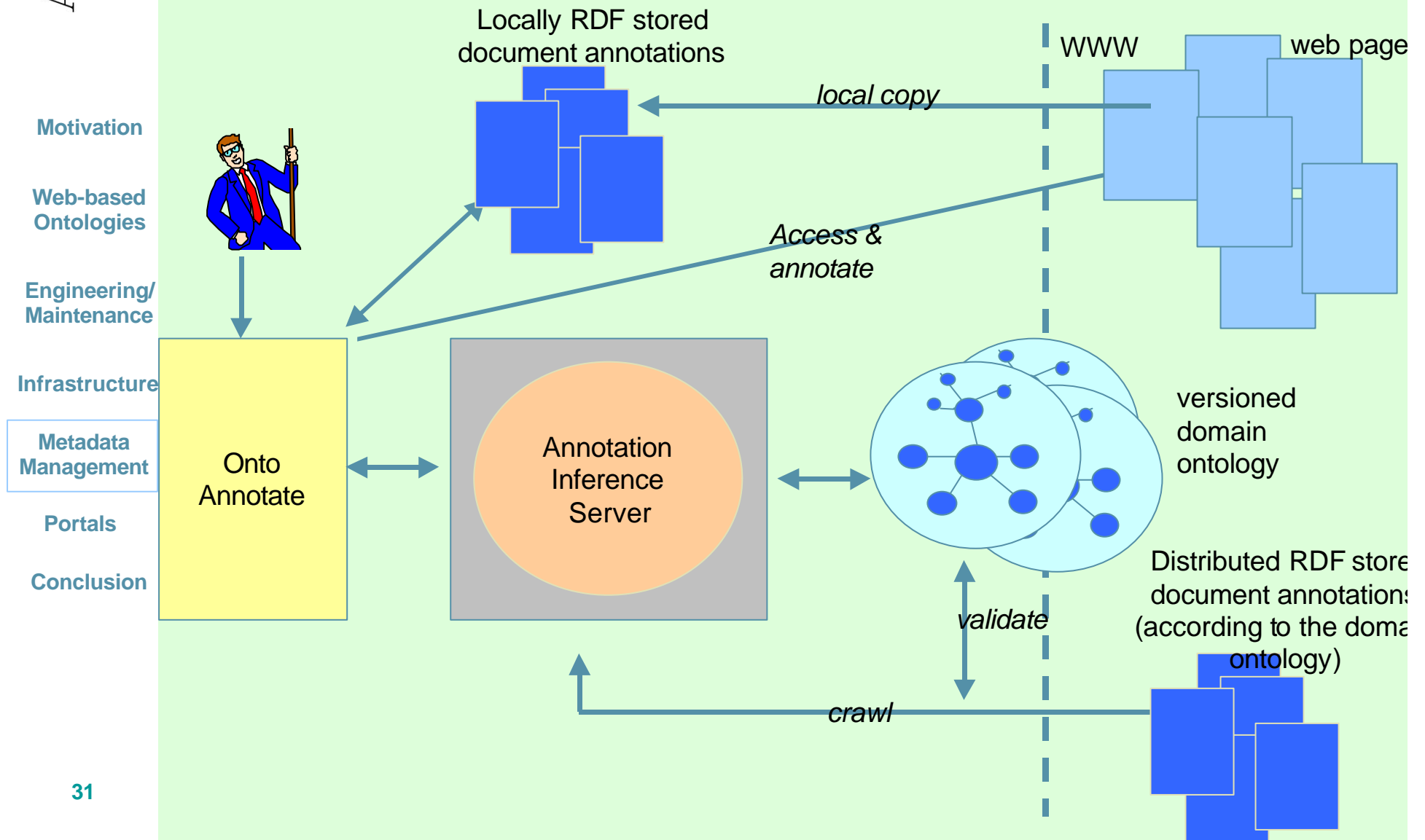
# Agenda

- Motivation
- Web-based Ontologies
- Ontology Engineering and Maintenance
- Semantic Web Infrastructure
- **Metadata Management**
- Semantic Web Portals
- Conclusion

## 5. Metadata Management

- There are large amounts of legacy data (HTML, XML, ...) available on the Web
- We need strategies for legacy data migration:
  - Annotation of Web documents (HTML, PDF, ...)
  - XML-Wrapper / Transformer
  - Database Converter / Exporter

# Semantic Annotation Architecture



# OntoAnnotate

Motivation

Web-based  
Ontologies

Engineering/  
Maintenance

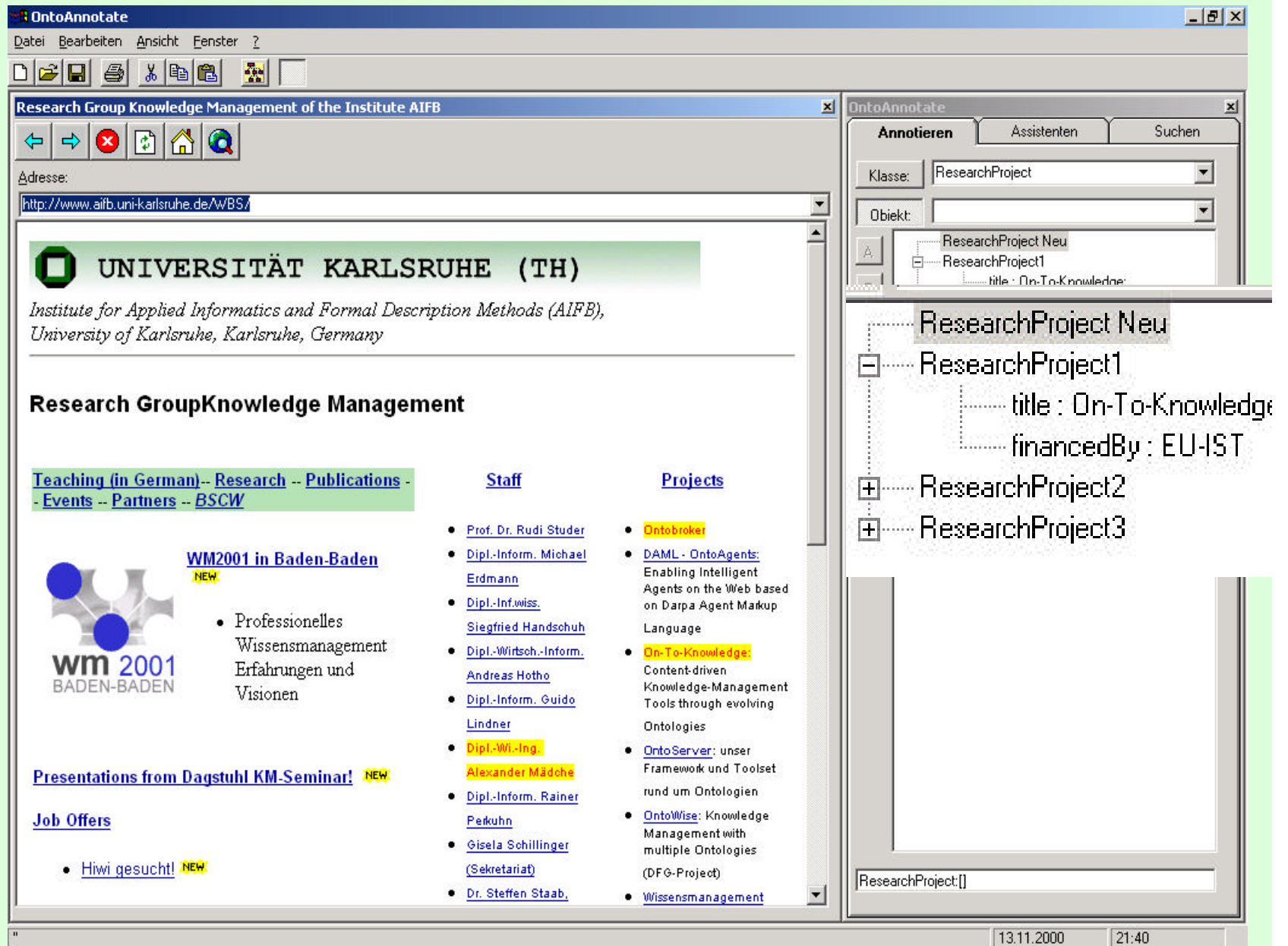
Infrastructure

Metadata  
Management

Portals

Conclusion

32



**OntoAnnotate**

File Bearbeiten Ansicht Fenster ?

Research Group Knowledge Management of the Institute AIFB

Adresse: <http://www.aifb.uni-karlsruhe.de/WBS/>

**UNIVERSITÄT KARLSRUHE (TH)**  
*Institute for Applied Informatics and Formal Description Methods (AIFB),  
 University of Karlsruhe, Karlsruhe, Germany*

**Research Group Knowledge Management**

[Teaching \(in German\)](#) - [Research](#) - [Publications](#) - [Events](#) - [Partners](#) - [BSCW](#)

**Staff**

- Prof. Dr. Rudi Studer
- Dipl.-Inform. Michael Erdmann
- Dipl.-Inf.wiss. Siegfried Handschuh
- Dipl.-Wirtsch.-Inform. Andreas Hotho
- Dipl.-Inform. Guido Lindner
- Dipl.-Wi.-Ing. Alexander Mädche
- Dipl.-Inform. Rainer Perkuhn
- Gisela Schillinger (Sekretariat)
- Dr. Steffen Staab

**Projects**

- Ontobroker**
- DAML - OntoAgents:** Enabling Intelligent Agents on the Web based on Darpa Agent Markup Language
- On-To-Knowledge:** Content-driven Knowledge-Management Tools through evolving Ontologies
- OntoServer:** unser Framework und Toolset rund um Ontologien
- OntoWise:** Knowledge Management with multiple Ontologies (DFG-Projekt)
- Wissensmanagement**

**Annotieren** Assistenten Suchen

Klasse: ResearchProject

Objekt:

- ResearchProject Neu
- ResearchProject1
  - title: On-To-Knowledge
- ResearchProject2
- ResearchProject3

ResearchProject[]

13.11.2000 21:40



# Semi-automatic Generation of Annotations

- Manual annotation of web documents is a time-intensive process
- Idea:
  - Use information extraction capabilities
  - Based on ontology learning mechanisms acquire ontology with corresponding language mapping
- Preprocess web documents linguistically and propose automatically annotations to the annotator!

# Maintenance of Metadata

- Metadata have to be maintained in the same way as ontologies or knowledge bases
  - metadata have to reflect changes of the sources
  - metadata have to reflect changes of the ontologies
- **sources, ontologies, and metadata** have to be maintained in a consistent way
  - organizational process is needed
  - tools are needed

# Agenda

- Motivation
- Web-based Ontologies
- Ontology Engineering and Maintenance
- Semantic Web Infrastructure
- Metadata Management
- **Semantic Web Portals**
- Conclusion

## 6. Semantic Web Portals

- Put the “**Semantic Web**” into practice for communities of interest
  - present a structured **semantic** view onto the Web
  - reflect basic paradigm of the Web:  
**self-organization**
  - crucial aspect:  
“people [and machines] can’t share knowledge if they don’t speak a common language”  
(Davenport)
    - **ontologies** provide the required conceptualizations

# Semantic Web Portals

- Integrated approach has to cover several aspects:
  - **Portal access** by community members
    - navigating
    - semantic querying
      - deliver integrated answers extended by derived facts
  - **Information provisioning**
    - all community members must be able to provide information
  - **Development and maintenance**
    - methodology and associated tools have to be provided

# Web Communities

Motivation

Web-based Ontologies

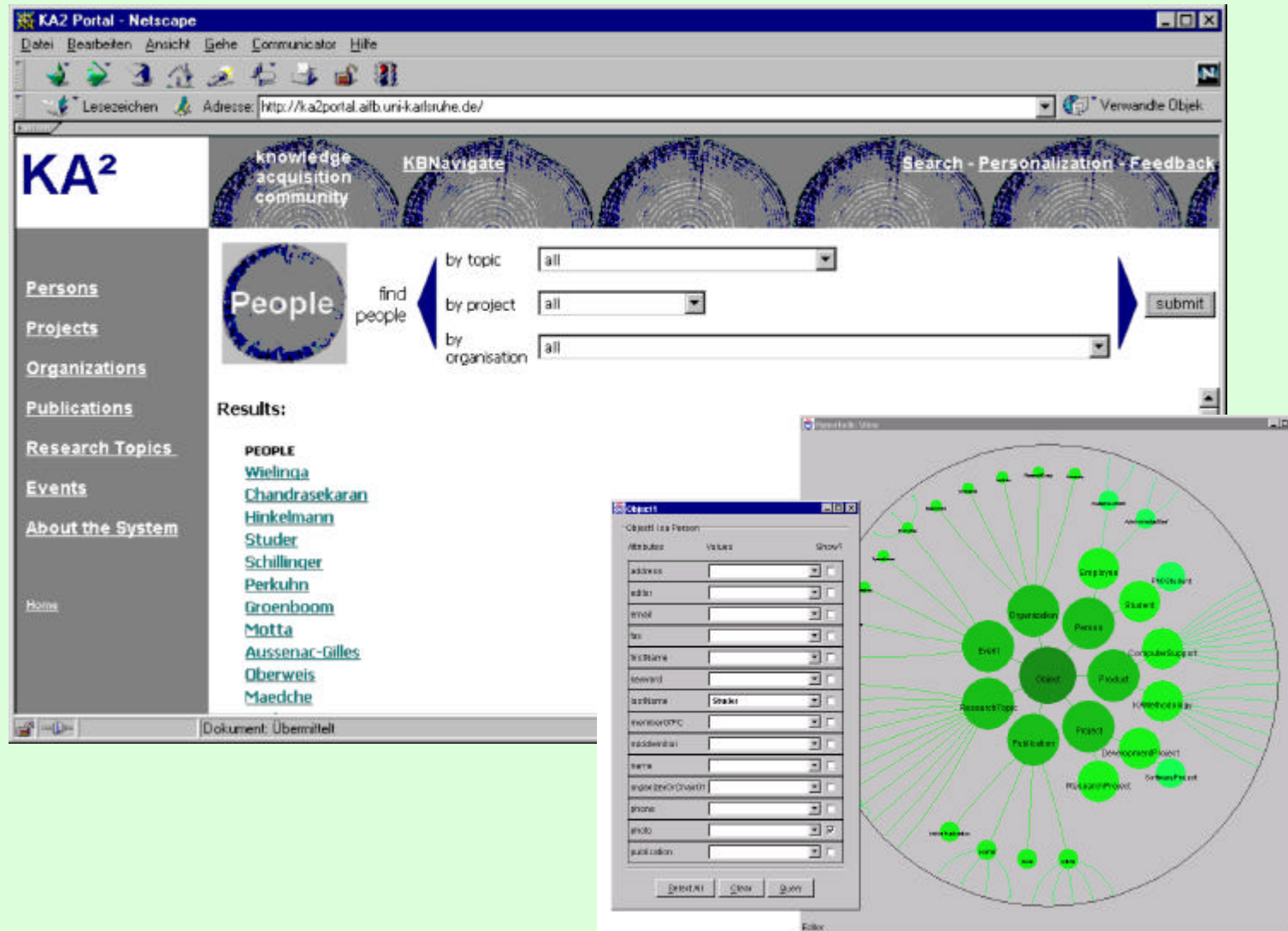
Engineering/Maintenance

Infrastructure

Metadata Management

Portals

Conclusion



The screenshot shows the KA2 Portal website in a Netscape browser window. The browser title is "KA2 Portal - Netscape" and the address bar shows "http://ka2portal.aifb.uni-karlsruhe.de/". The website header includes "KA<sup>2</sup> knowledge acquisition community", "KBNavigate", and "Search - Personalization - Feedback". A search interface is visible with a "find people" button and three dropdown menus for "by topic", "by project", and "by organisation", all set to "all". A "submit" button is to the right. Below the search interface, a "Results:" section lists several names under the heading "PEOPLE": [Wielinga](#), [Chandrasekaran](#), [Hinkelmann](#), [Studer](#), [Schillinger](#), [Perkuhn](#), [Groenboom](#), [Motta](#), [Aussenar-Gilles](#), [Oberweis](#), and [Maedche](#). In the bottom right corner, there is a "Network View" window displaying a complex network graph with green nodes and connecting lines. An "Object" window is also open, showing a form for "Object: Lisa Person" with various input fields for attributes like address, email, phone, etc.

# Conclusion

- **Ontologies** provide the **semantic underpinning** for Semantic Web applications
- **Ontology repositories** will be distributed on the Web
  - methods and tools for accessing/reusing/aligning ontologies are needed
- Ontologies will be specified in **different languages**
  - support co-existence of different languages
  - embedding in RDFS seems promising
- Reduce overhead of building up ontologies
  - exploit available resources
  - exploit text mining

**=> ontology learning**

## Conclusion (2)

- **Semantic Web Applications**
  - fast growing market
    - B2B applications
    - Agent-based services
    - Semantic Web portals
    - Knowledge Management
- **Europe has strong position** with respect to required methods and tools
  - promising starting point
  - DARPA spends US\$ 80 Mill. for the Semantic Web
  - Funding within IST programme is strongly needed