AIFB

# 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

AIFB



# 1. Motivation
# Semantic Web: The Vision

- WWW is an impressive success, cf.
  - amount of available information
  - number of human users

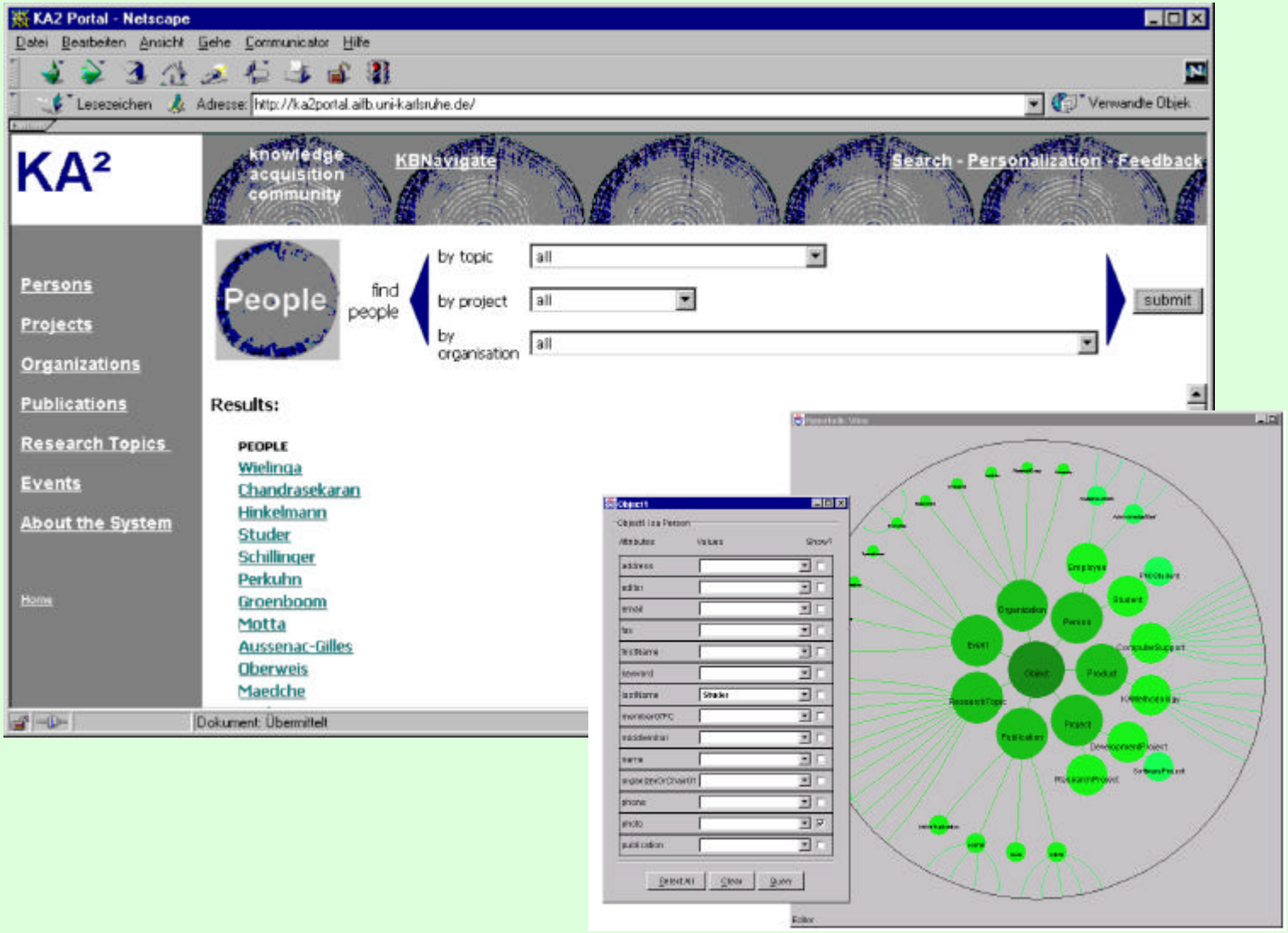
- However: WWW is currently only for human readers, but …

... *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)

AIFB

# Web Communities

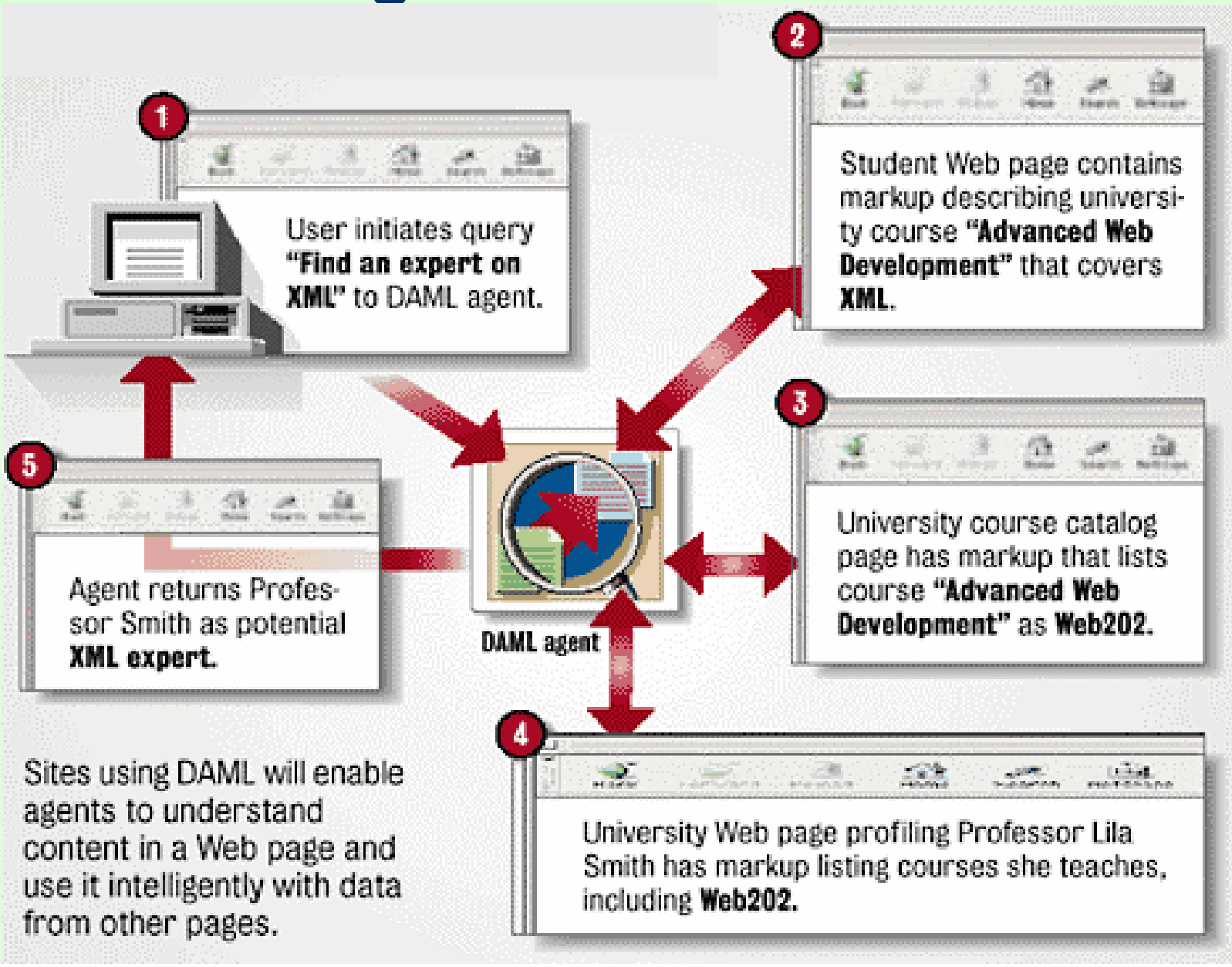




AIFB



# Agents on the Web



# XML is not the Solution

- Meaning of XML-Documents is *intuitive*ly 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

- Ontologies as key-enablers of this vision

- In this talk we outline technologies for
  - Representing,
  - Building,
  - Using and,
  - Applying

  ontologies on the Web


# Agenda

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

AIFB



# 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

AIFB

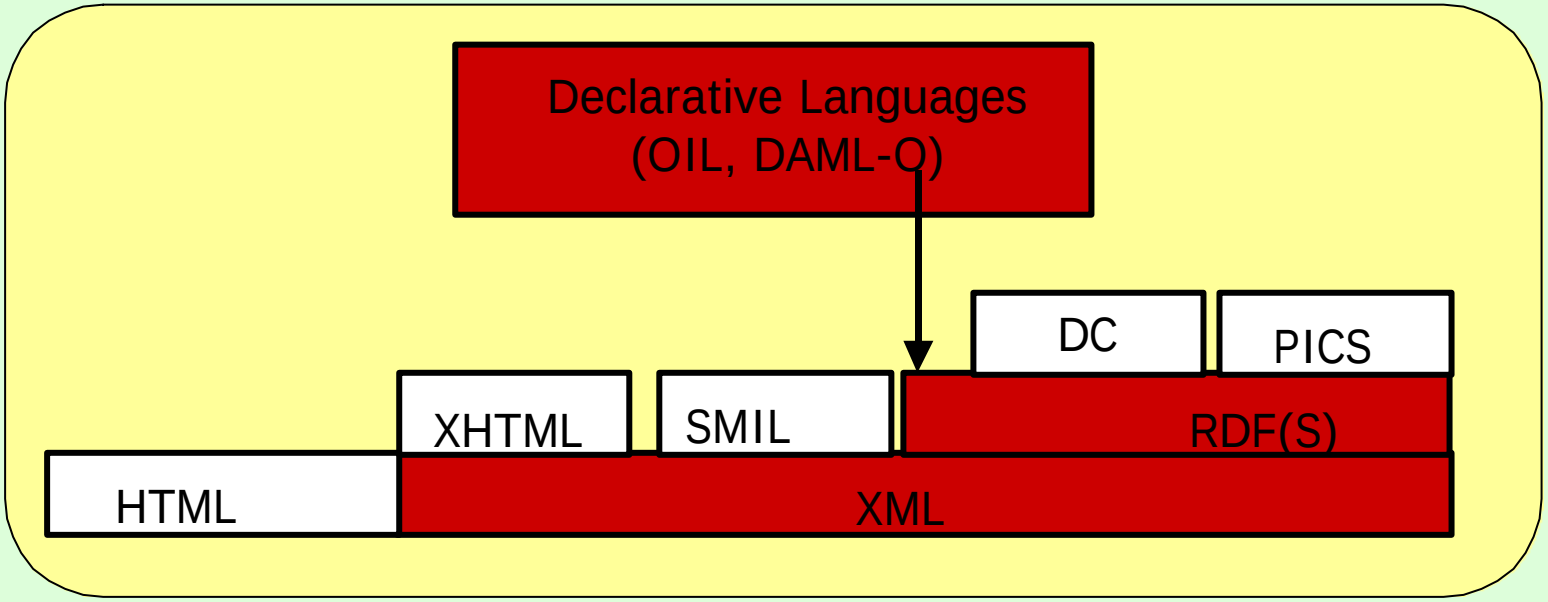


# Different Types of Ontologies

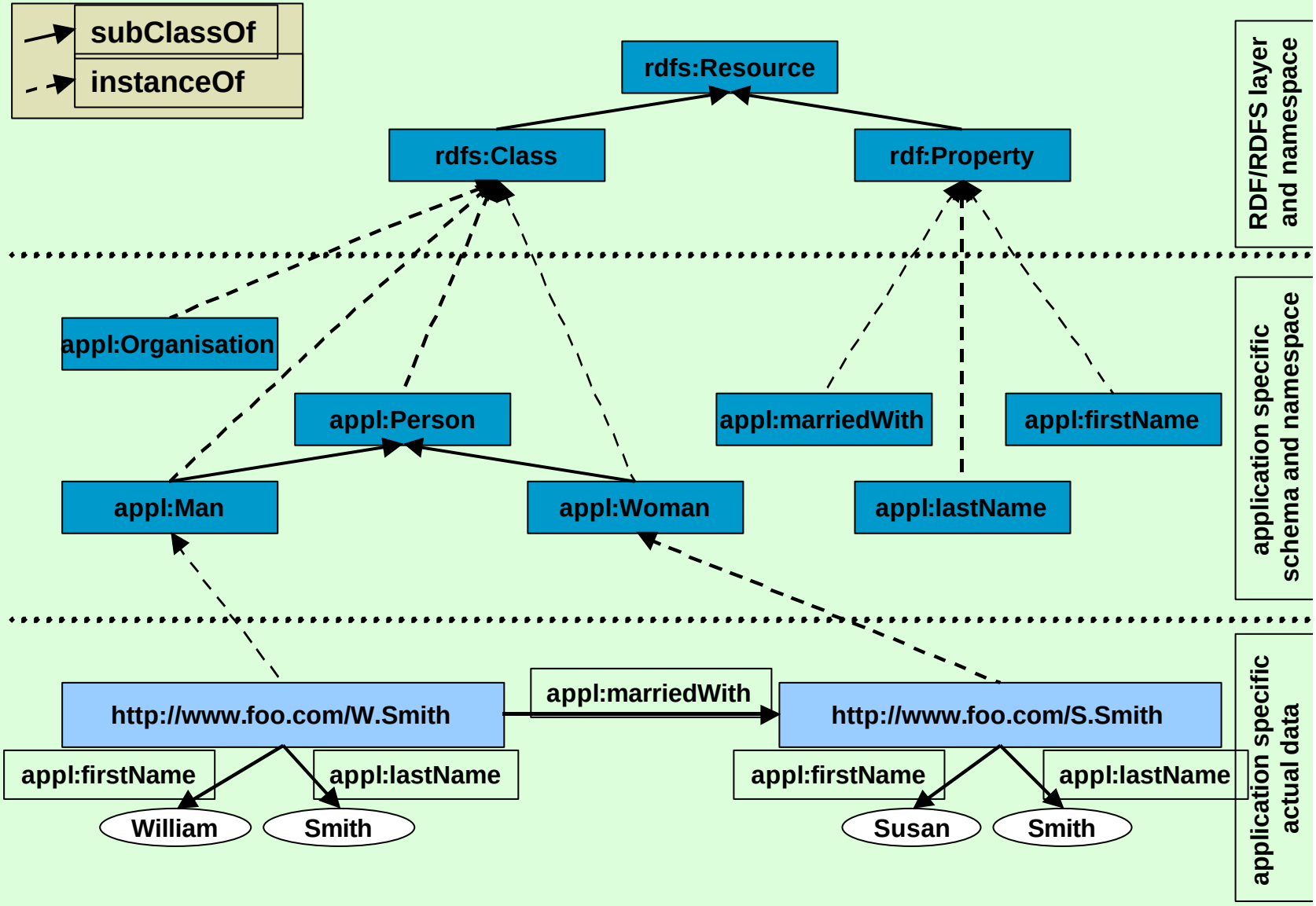
- **Light-weight Ontology**
  - concepts, atomic types
  - is-a hierarchy among concepts
  - relations between concepts
- **Heavy-weight Ontology**
  - cardinality constraints
  - taxonomy of relations
  - reified statements
  - Axioms / semantic entailments of various tastes
    - expressiveness (description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

# Layered Languages

The W3C hierarchy of languages:



# RDF Data Model



AIFB



# 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

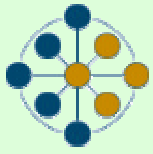
- But: RDF(S) is not expressive enough !!

# 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

ON knowledge TO

*Content-driven Knowledge Management Tools through Evolving Ontologies*



AIFB

vrije Universiteit amsterdam

BT

CognIT a.s

Rentenanstalt

Swiss Life

ENERSEARCH AB

# 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

AIFB



# 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
- ...

AIFB

Motivation

Web-based Ontologies

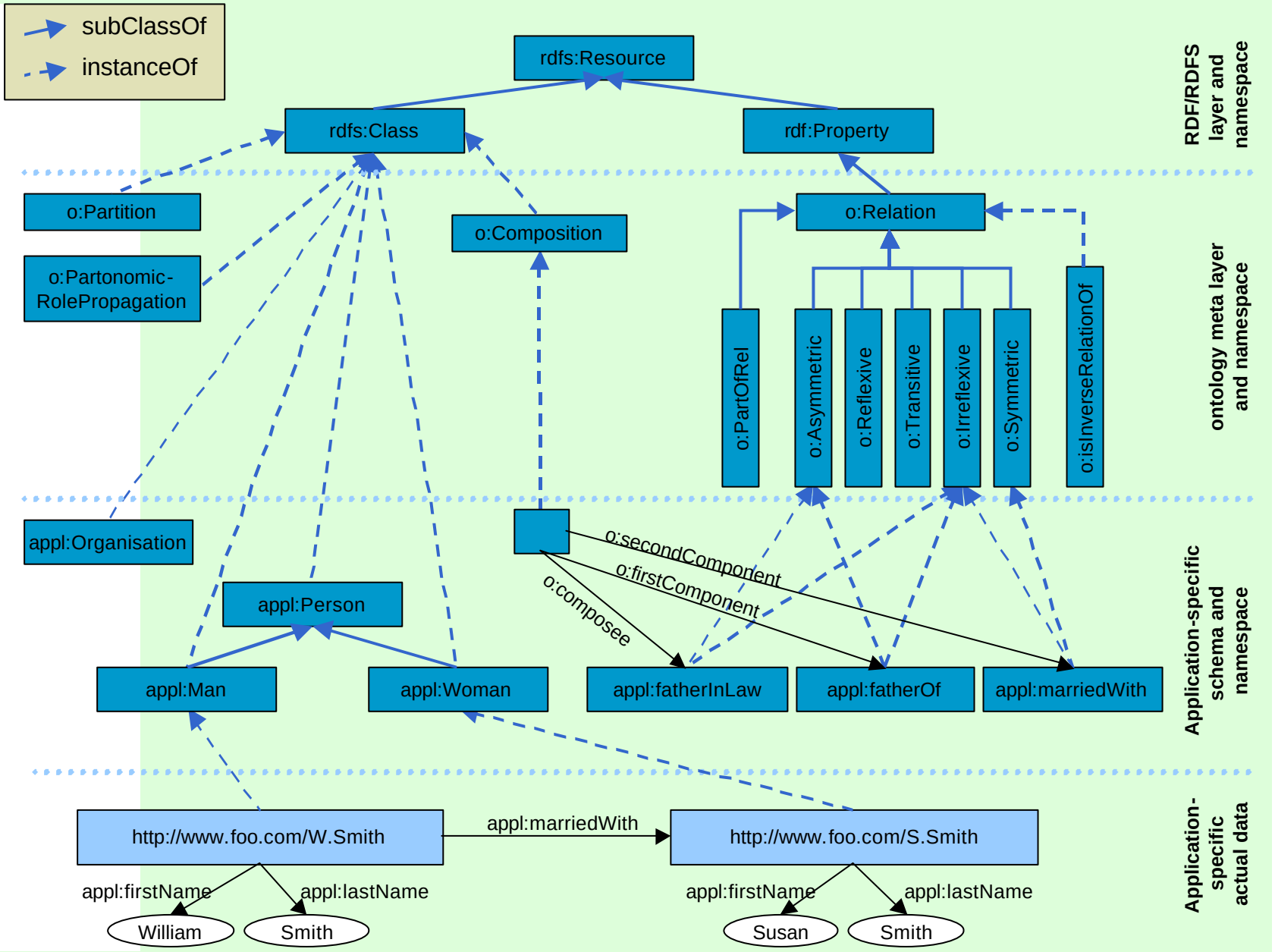
Engineering/ Maintenance

Infrastructure

Metadata Management

Portals

Conclusion



# Agenda

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

# 3. Ontology Engineering and Maintenance

**Tools:**

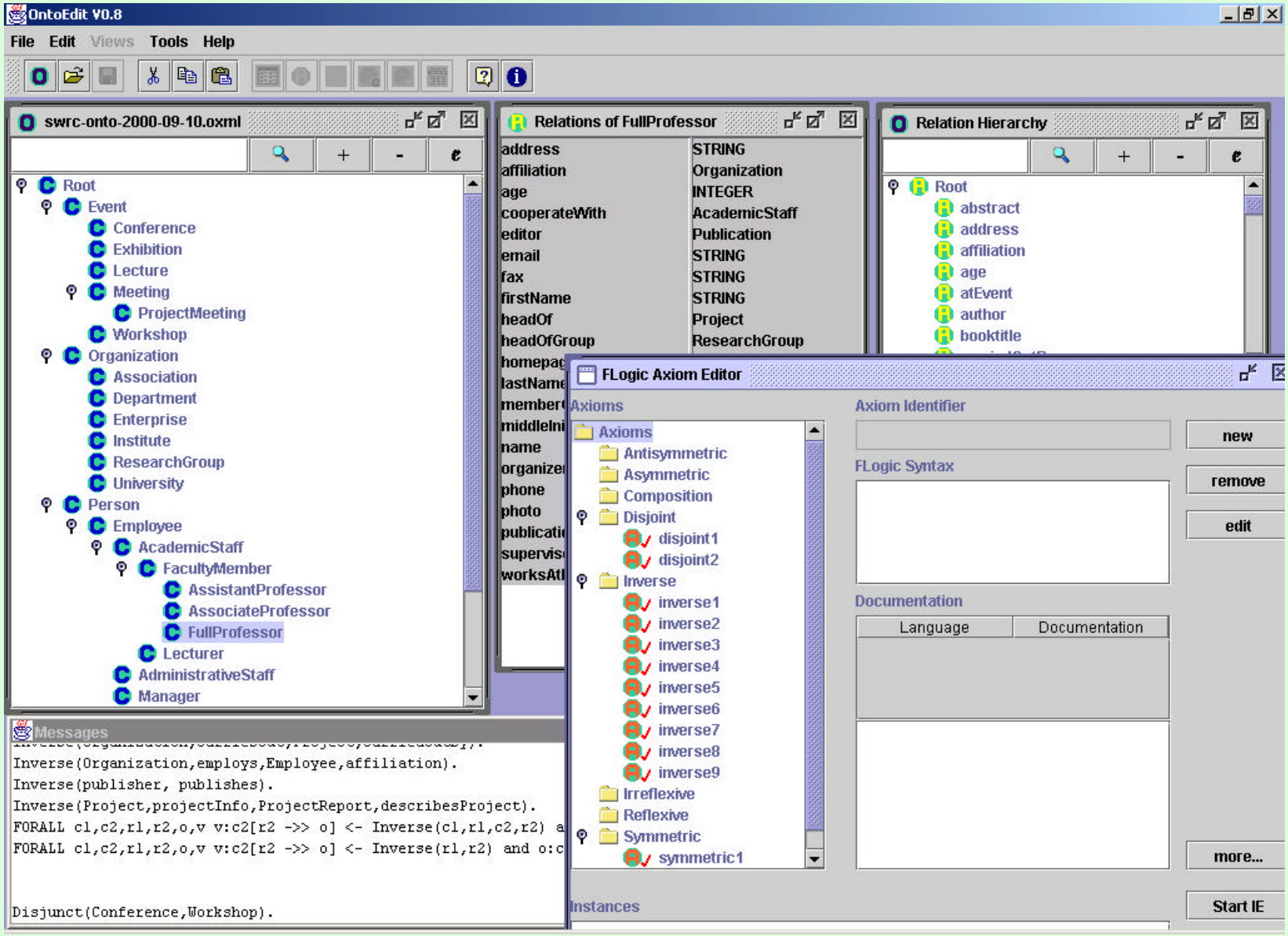
- **Light-weight**
  - uncontroversial
  - all tools support light-weight ontologies
    - Protégé-2000 KA Environment, Stanford University
    - OntoEdit Web Ontology Workbench, University of Karlsruhe
    - UML-Tools
- **Heavy-weight**
  - no consensus yet
  - layering seems appropriate/necessary

# 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





# 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‘‘

AIFB

# Ontology Learning



## Ontology Engineering by Means of Ontology Learning

### Ontology Creation

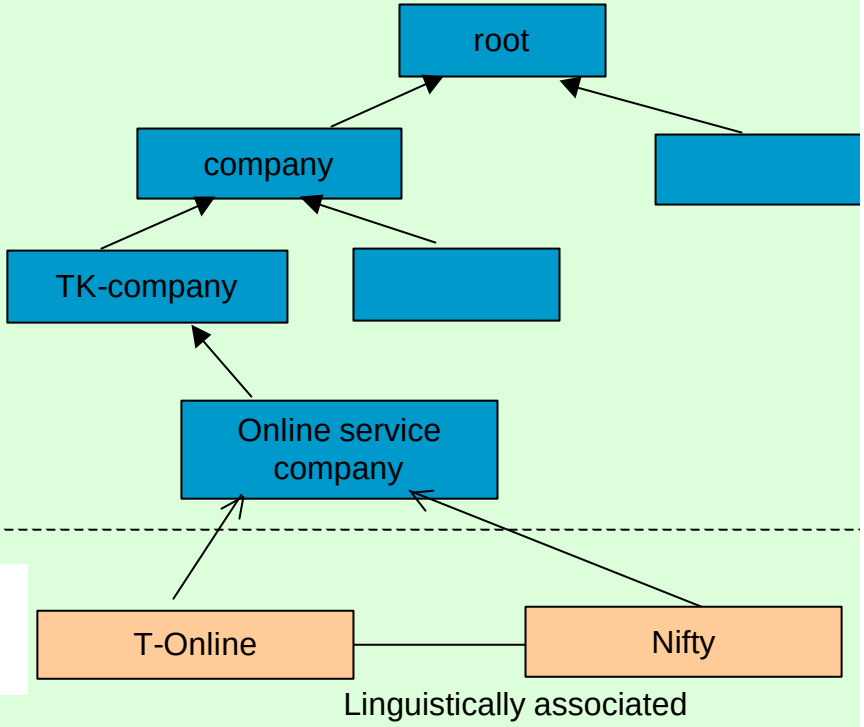
- start from scratch
- built on top of existing structures
- integrate existing knowledge sources

### Ontology Maintenance

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

# Relation Mining [Maedche,Staab 00]




**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

AIFB

# Agenda

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

AIFB

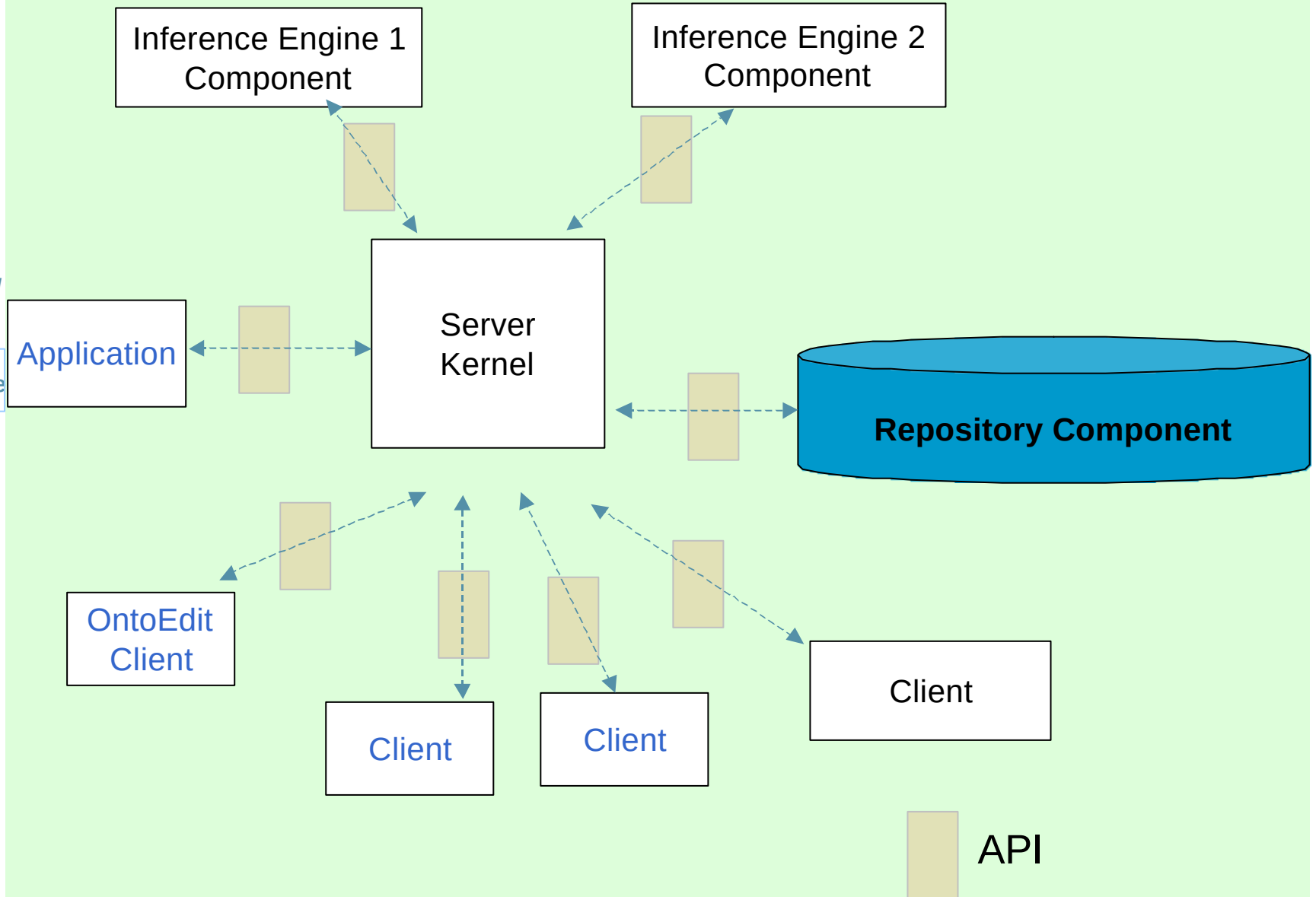
# Semantic Web Infrastructure

- 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

AIFB

# OntoServer – Basic Infrastructure





# Agenda

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

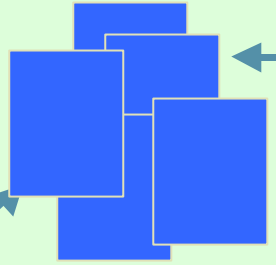
# 5. Metadata Management

- 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

AIFB

# Semantic Annotation Architecture





Locally RDF stored document annotations

*local copy*

WWW

web page

*Access & annotate*

Onto Annotate

Annotation Inference Server

versioned domain ontology

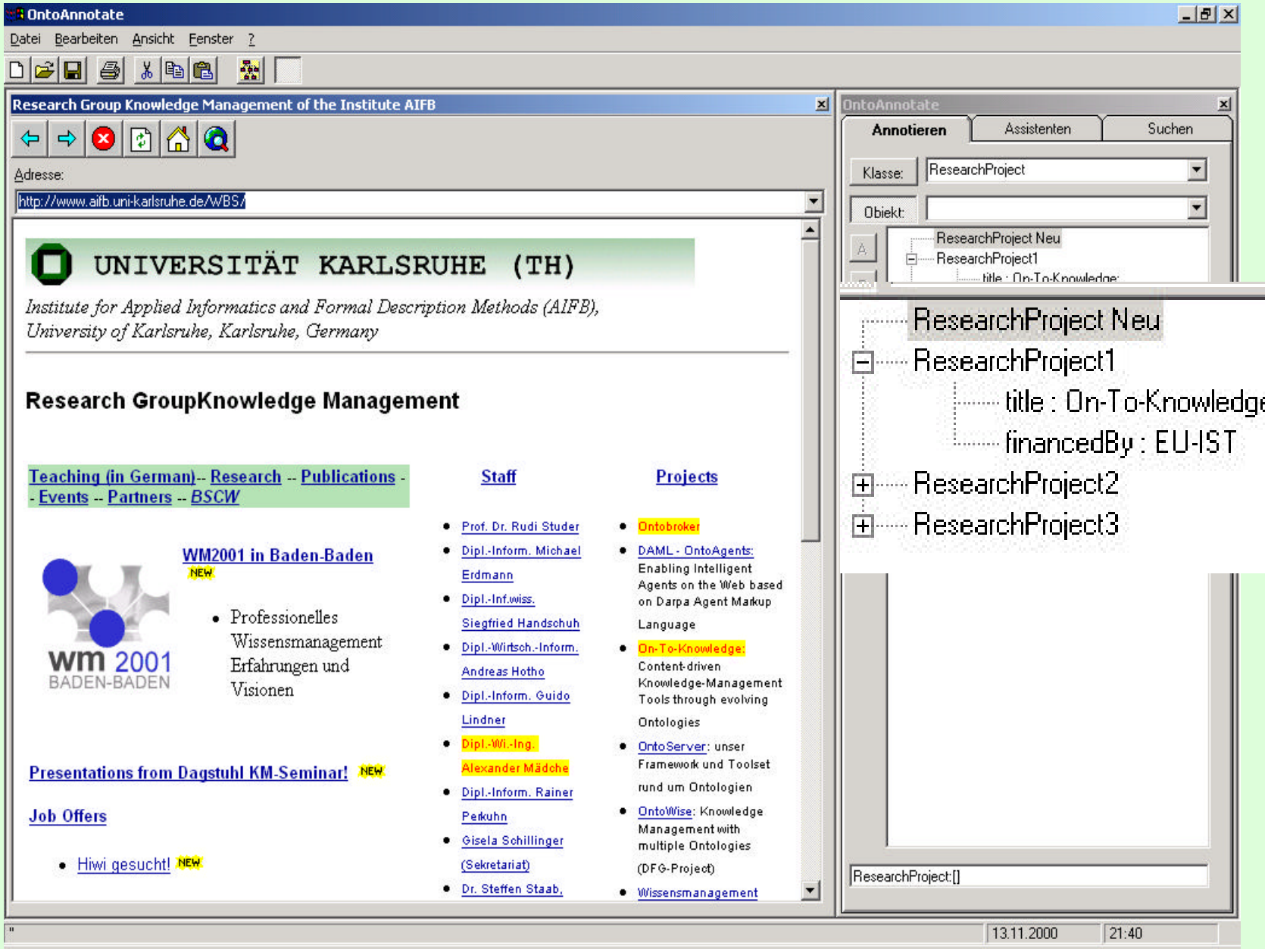
*validate*

Distributed RDF store document annotations (according to the doma ontology)

*crawl*

# OntoAnnotate





# 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!

AIFB



# 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

AIFB

# 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

AIFB

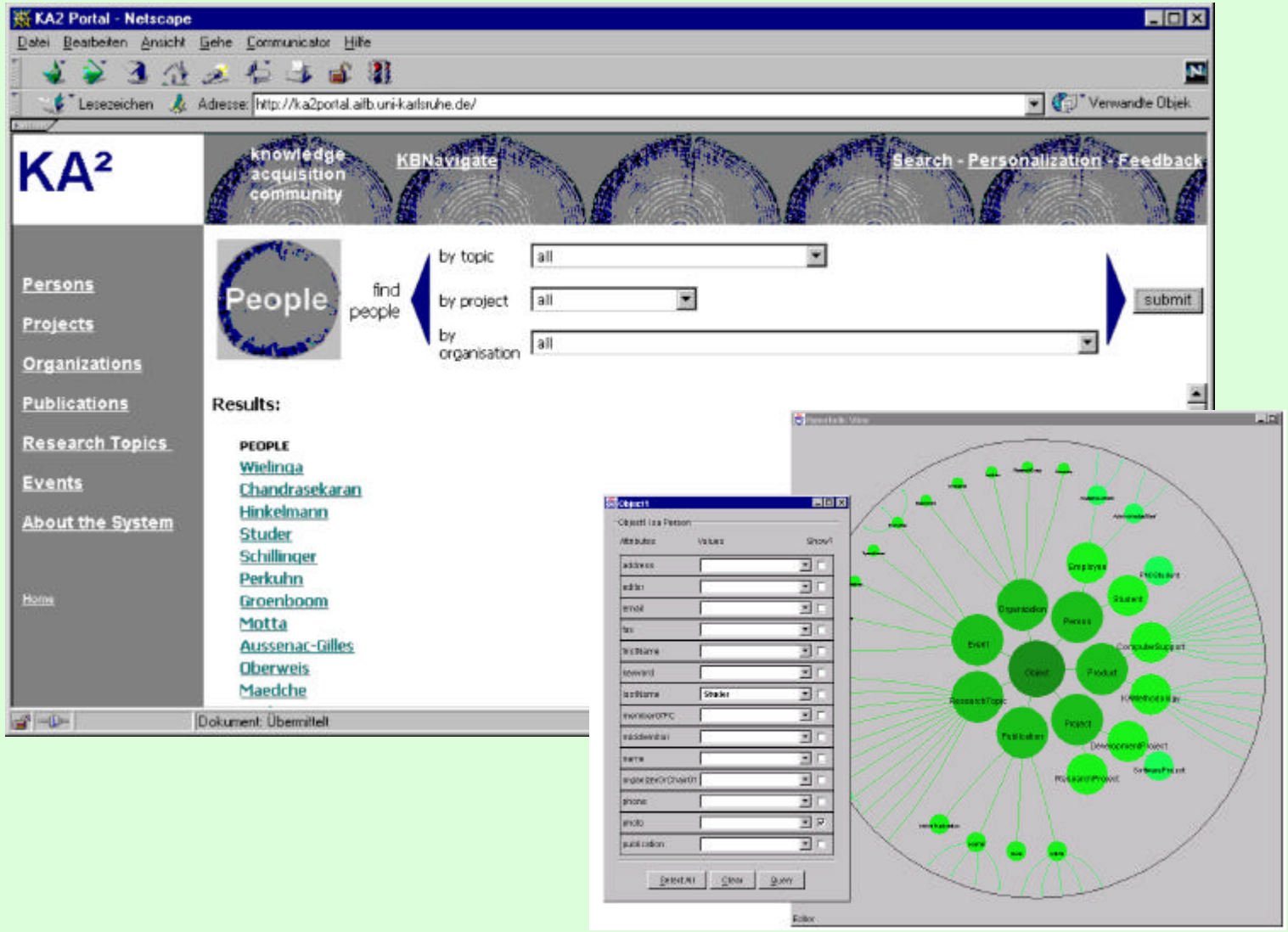
# Semantic Web Portals

- 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

AIFB

# Web Communities





AIFB



# 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**

AIFB



# 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