SEAL - Tying Up Information Integration and Web Site Management by Ontologies: Some Open Research Issues

Rudi Studer
Alexander Maedche, Steffen Staab, Raphael Volz

Institute AIFB, University of Karlsruhe, Germany
www.aifb.uni-karlsruhe.de/WBS

FZI Research Center on Information Technologies, University of Karlsruhe

ontoprise GmbH, Karlsruhe

L3S Learning Lab, Hannover/Karlsruhe

NSF-EU Workshop DB-IS Research for Semantic Web and Enterprises
Amicalola Falls State Park, April 2-5, 2002
Agenda

- Context
  - SEAL: Semantic Portal Approach
  - KAON: Karlsruhe Ontology and Semantic Web Framework
- Selected Research Challenges
  - Access to and Integration of Databases
  - Ontology Alignment and Integration
  - View Mechanisms
SEAL: SEmantic PortAL

- Framework for
  - managing community web sites
  - providing information integration from various kinds of sources

- All functionalities are ontology-based
  - Wrapping and reverse engineering of (DB) sources
  - Integration services
  - View management
  - Semantic querying
KAON: Karlsruhe Ontology and Semantic Web Framework

- A RDF-based Software Infrastructure
  - [http://kaon.semanticweb.org](http://kaon.semanticweb.org)
- Based on RDF(S), with several extensions, e.g. for typed, multilingual lexical expressions
- Component-based, easily extendable application framework
- Open-Source Tool Suite, supporting and supported by

© R. Studer
KAON Conceptual Architecture

Applications & Services
- OntoMat App Framework
  - DB Reverse Engineering
  - Focused Crawler
- Web Service Connectors
- Web Application Framework
  - Web Portal

Middleware
- KAON-API
  - Mapping-Engine
  - RDF-API
- KAON-API

Data And Remote Services
- Reasoning Services
- P2P
- Relational Database
- XML/RDF Files
Agenda

- **Context**
  - SEAL Approach
  - KAON: Karlsruhe Ontology and Semantic Web Framework

- **Selected Research Challenges**
  - Access to and Integration of Databases
  - Ontology Alignment and Integration
  - View Mechanisms
Integration of Databases into Semantic Web Applications
(Reverse Engineering or relational Databases)

- Large amounts of databases are available that are used as backend of Web applications
  - These databases provide semantic structures (schemata) and DB instances

- Means should be provided to generate (on the fly)
  - Ontologies from databases
  - Mappings from databases to ontologies, and
  - Semantic-Web enabled content in the form of RDF statements from such databases (instances)

- Databases should keep their autonomy
  - Semantic Web applications should access the DBs via queries that exploit the mapping rules
Technical Challenges

- Creation of Concepts
  - Which DB-relations are mapped to concepts
- Creation of ontological relations between concepts
  - Exploit functional dependencies
- Generate additional axioms to guarantee semantic integrity of resulting ontology

- Build upon results from OODBMS
  - Adapt to specific characteristics of ontology languages
    - Expressive power
    - Language characteristics and primitives
Mapping Databases to Ontologies: KAON REVERSE
Ontology Alignment and Integration

- Real life applications will rely on several ontologies
  - Semantic alignment/integration is needed

- Exploit and extend results from DB research
  - Federated Databases
  - Data Warehouses
  - Mediators
The Ontology Alignment and Integration Components

Transformation Engines!

Domain Knowledge & Constraints

Evolution

Semantic Bridging

Object Identity!

Execution

Transactions!

GUI

Cooperative Consensus Building

Postprocessing

Mapping Discovery!

Reengineering!

Lift & Normalization

Similarity

Association
View Mechanisms for the Semantic Web

- Integration of heterogeneous and distributed sources
  - Provide customized ontologies for different clients
  - Part of mediation process
- Authorization
  - Restrict access to underlying sources
- Make global ontologies more manageable
  - Split into manageable pieces that are related in a semantic way

© R. Studer
Specific Challenges for View Mechanisms for the Semantic Web

- **Web-awareness**
  - Views are typically based on various sources

- **Ontology basis**
  - View definitions should be ontology specifications
  - Application of view primitives should lead to ontology-based data

- **Meet RDFS characteristics**
  - No strong typing
  - Properties as first class citizens
Views for RDFS-based Web Ontologies

- Inspired by object-oriented views
- Distinction of views on classes and views on properties
  - Due to underlying RDFS model
- Views on properties alter the definition of base properties
- Views are embedded into property / class taxonomy based on semantics of view operations

© R. Studer
Views example*

* Classes only

query = \textit{Student} \setminus \textit{Employee}

query = \textit{PhD-Student} \cup \textit{Professor}
Conclusion

- We are on the way to realize the Semantic Web

- DB techniques provide a promising starting point for handling a collection of issues
  - Have to be adapted and extended
    - Meet characteristics of RDFS-based models
    - Methods have to be 'Semantics-aware'

- Tradeoff between 'sophistication' and scalability
  - Take into account both human and machine point of view
Thank You!

www.aifb.uni-karlsruhe.de/WBS

www.fzi.de/wim

www.ontoprise.de

www.learninglab.de