Sören Auer [1,2], Sebastian Dietzold [2], and Thomas Riechert [2]
[1] University of Pennsylvania, Department of Computer and Information [2] Universität Leipzig, Institut für Informatik

ISWC 2006

OntoWiki – A Tool for Social, Semantic Collaboration

Presentation by Patrick Roos

OntoWiki Purpose

- Tool supporting agile, distributed knowledge engineering in a (semantic) web environment
- Connect existing information resources
- Collaborative ontology and knowledge base generation

→ Similar goals, issues, and challenges to SIB-data project

Integrating Wikis and Semantic Web

Recent approaches:

- Integrate RDF triples into Wiki texts in a special syntax

Obstacles:

- Usability:
  - Adding more and more syntactic possibilities counteracts ease of use for editors.
- Redundancy:
  - To allow the answering of real-time queries to the knowledge base, statements have to be stored additionally in a triple store. This introduces a redundancy complicating the implementation.
- Scalability:
  - Knowledge base changes which involve statements with different subjects will scale very bad since all corresponding Wiki texts have to be parsed and changed.

OntoWiki Strategy

- Wiki Paradigm:
  - "making it easy to correct mistakes, rather than making it hard to make them"
  - Simplest "DB" that could possibly work
- OntoWiki:
  - Simplest Knowledge Base that could possibly work
- Simplify the presentation and acquisition of instance data from and for end users
- Regard knowledge bases as ***information maps***

Information Maps

- RDF knowledge bases
- Node (RDF resource): Web-accessible pages interlinked to related digital resources
- Divided into three parts:
  - Left sidebar: selection of content
  - Main content section
  - Right sidebar: content specific tools and complementary information

Web-accessible Page

- Selection options include the set of available knowledge bases, a class hierarchy browser and a full-text search
- Once a selection is made, the main content section will arrange matching content in a list view linking to individual views for individual instances
- Different views facilitated (e.g. map, calendar)

Editing: WYSIWYG'esque

Conference > **ISWC 2006**

title	Fifth International Semantic Web Conference
place	Athens, Georgia, USA
logo	
start	2006-11-05
end	2006-11-09
topics	⊞ SemanticWeb
URL	http://iswc2006.semanticweb.org/
rdfs:comment	The dream of the Web was to create a human communication and collaboration platform for sharing knowledge and enabling a universal space for

Concept Identification and Reuse

- Knowledge bases become increasingly advantageous, if once defined concepts (e.g. classes, properties, or instances) are as much reused and interlinked as possible
- AJAX technology to interactively propose already defined concepts
- All URI references and literals are indexed for full-text searches in the statement repository

Editing Widgets

- OntoWiki system provides a library of reusable user interface components for data editing, called widgets
- Customized for usage in specific contexts by widget configurations
- Basic Widgets:
  - Statements: allows editing of subject, predicate, and object
  - Nodes: edit literals or resources
  - Resources: select and search from/for existing resources
  - Literals: literal data in conjunction with a data type or a language identifier
  - Widgets for specific literal data types: e.g. dates, HTML fragments.
  - File widget: allows uploading of files to the OntoWiki system.

Social Collaboration

- Change tracking
- Commenting

- Rating
- Activity/Provenance

Search

- Structuring and representation of content should be employed to enhance the retrieval of information for human users
- Facet-based Browsing
  - widely deployed in the shape of tagging systems of the Web 2.0
- Semantic Search
  - full-text search for one or multiple keywords occurring in literal property values

Semantic Search

Search

Filter: All classes | All properties | Submit Query

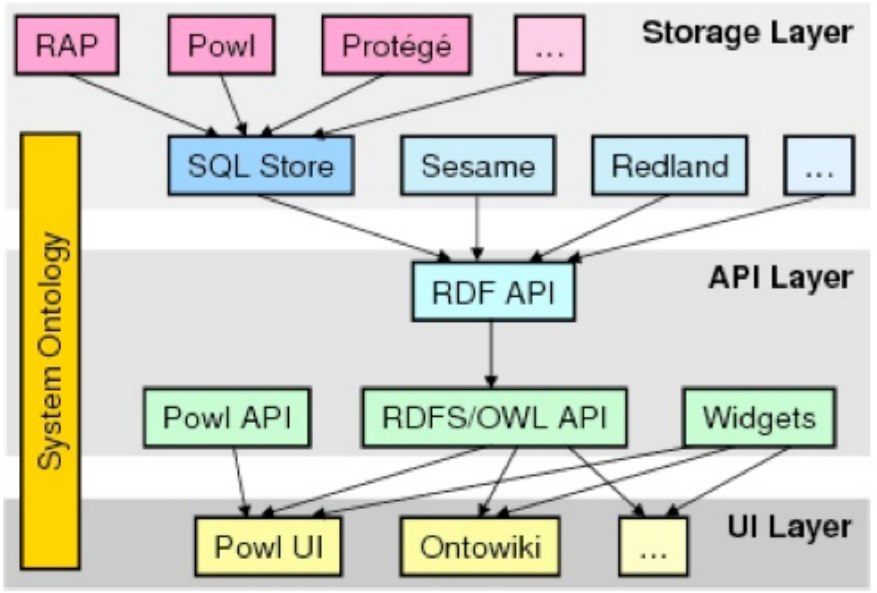
All properties
swrc:address (2)
swrc:booktitle (3)
swrc:name (1)

Search returned **6** results.[c]

Relevance	Resource	Property	Value
100%	AssistantProfessor: York Sure	name	York Sure
100%	InProceedings: Studies on the Dynamics of Ant Colony Optimization Algorithms	booktitle	Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2002), New York
100%	Proceedings: Business Process Management	address	Berlin, Heidelberg, New York
100%	InProceedings: REMINDIN': Semantic Query Routing in Peer-to-Peer Networks Based on Social Metaphors	address	New York, USA
100%	InProceedings: Developing and Managing Software Components in an Ontology-based Application Server	booktitle	Proceedings of the WWW2004 Workshop on Application Design, Development and Implementation Issues in the Semantic Web, New York, NY, USA, May 18, 2004
100%	InProceedings: Reuse Of Problem-Solving Methods In Knowledge Engineering (short paper)	booktitle	Proceedings of the 6th Annual Workshop on Software Reuse (WISR-6), Owego, New York, November 1-4, 1993

Architecture: POWL

- Web-based collaborative semantic web development platform (PHP and MySQL): *powl.sourceforge.net*
- Parsing, storing, querying, manipulating, serving and serializing OWL knowledge bases
- Scalable
  - Based on RDBMS
- Supports
  - Versioning
  - Caching



Storage Layer (Powl Store)

- SQL compatible rel. DB supported by AdoDB
- Denormalized schema: resources and literals written in full in a table row representing an RDF statement (2-3 x faster than normalized)
- Tables used to store all information related to ontologies and their evolution

Table	Description
models	provides information about the models in the store
statements	contains all statements of models in the store
log_actions	holds information about editing actions on a model
log_statements	contains added and removed statements for every action

API Layer

- RDFAPI
  - Parser, serializer for different RDF serializations (XML, N3, N-Triple)
  - RDQL declarative query backend
  - Classes and methods for working with RDF models, resources and literals
  - NetAPI for publishing models on the web

The higher layered APIs, RDFSAPI and OWLAPI, extend the classes “Model”, “Resource” and “Literal” provided by RDFAPI to RDF-Schema specific classes.

RDFSAPI

- Methods exposed by the main RDFSAPI classes

RDFSModel	RDFSClass
add{Triple,Class,Property,Instance} count{Triples,Classes,Properties,Instances} find{Triples,Resources} get{Triple,Resource,Class,Property,Instance} listTop{Classes,Properties} list{Resources,Classes,Properties,Instances} list{Namespaces,Languages,Datatypes} remove{Triple,Resource,Class,Property,Instance} load	add{Instance,Property} count{Instances,Subclasses} find{Instances} get{Instance} list{SubClasses,SuperClasses, Properties,Instances} remove{Instance,Property} set{SubClasses,SuperClasses, Properties}

RDFSProperty	RDFSInstance
add{Domain,Range} add{SubProperty,SuperProperty} get{Domain,Range} list{Domain,Range} list{SubProperties,SuperProperties} remove{Domain,Range} set{Domain,Range}	addPropertyValue get{Class,Classes} get{PropertyValue} listProperties list{PropertyValues} setProperty{Value,Values}

More Architecture Info

Sören Auer. *Powl: A Web Based Platform for Collaborative Semantic Web Development*. Proceedings of the Workshop Scripting for the Semantic Web, number 135 in CEUR Workshop Proceedings, Heraklion, Greece, 05 2005.

Scalability

- Only those parts of the ontology are loaded into main memory which are required to display the information requested by the user on the screen
- Class subtrees are loaded on demand using Javascript
- Benchmark with only MySQL's query cache opimization:

Model	Triple count	Import time	Classes	Class hierarchy calulation
Wordnet[2]	473 528	624 s	6	0.30 s
NCI Cancer Ontology[3]	463 878	597 s	27 652	0.46 s
UNSPSC[4]	82 500	82 s	16 499	1.06 s

Conclusions

Application domain:

- Environment:
  - OntoWiki is a Web application and presumes all collaborators working in a Web environment, possibly spatially distributed.
- Usage Scenario:
  - OntoWiki focuses on knowledge engineering projects where a single, precise usage scenario is either initially not (yet) known or not (easily) definable
- Reasoning:
  - Application of reasoning services is (initially) not mission critical.
- Budget:
  - Only a small amount of financial and personnel resources are available

Conclusions

- Showed how tool support for agile, collaborative knowledge engineering scenarios can be provided
- Visual presentation of a knowledge base as an information map, social collaboration aspects as well as a semantic search strategy decrease the entrance barrier for domain experts to collaborate
- Which ideas would be good for SIB-data?