OntoViews – Semantic Portal Creation Tool

Eetu Mäkelä
16.11.2005

Semantic Computing Research Group
http://www.cs.helsinki.fi/group/seco/
OntoViews

- Core Functionality:
  - Semantic view-hierarchy-based searching and view projection
  - Lateral semantic collection browsing
  - Semantic text autocompletion

- Features:
  - Data-independent, has been applied to 8 vastly different datasets
  - Modifiable, has been used to create 5 differing user-interface families
  - Extensible, based on modular components
Contents

- What is OntoViews?
- Portals created with OntoViews
- A demonstration
- From ontologies to views
- Main components
OntoViews applications

- **MuseumFinland**
  - Semantic Virtual Museum combining material from three museums. Semantic Web Challenge Award, Prime Minister’s award for most technologically innovative web application

- **Orava**
  - Search index for learning video material from YLE

- **SW-Suomi.Fi**
  - Semantic version of the Suomi.Fi portal

- **Veturi**
  - Semantic Yellow Pages service discovery
OntoViews tests

- **Open Directory Project-test**
  - Directory of 2.2 million web site descriptions

- **ONKI-test**
  - General ontology browsing

- **Promoottori-test**
  - University promotion material

- **CultureSampo**
  - MuseumFinland successor under development
OntoViews demonstration

- http://www.museosuomi.fi/
View Projection from Ontologies
The Main Components of OntoViews

- Ontodella, a prolog-based logic server
- Ontogator, a java-based multi-facet search engine
- OntoViews-C, the main Apache Cocoon-based interaction and control component
The Logic Server Ontodella

- Projects view-facets from the data to be used in search
- Semantically links individual items to each other
- Utilizes defined logic rules for both functionalities
The Multi-Facet Search Engine
Ontogator

- Is a generic view-based RDF search engine
  - Defines and implements an RDF-based query interface defined as an OWL ontology
  - Replies to queries in RDF/XML that has a fixed structure
  - Extendable projection and query selector interfaces
- Has been tested with dmoz.org data to scale to up to 2.3 million items and 275,000 categories with search times of about 5 seconds.
- Future work: Does not yet scale well to accommodate multiple simultaneous users
The Interaction and Control Component OntoViews-C

- Built on top of the Apache Cocoon architecture
  - The Cocoon architecture is based upon the concept of pipelines, comprised of modular components (generators, transformers and serializers) that consume and/or produce XML
  - This forces a modular, reusable and extendable design
- In OntoViews, all components produce not only XML, but valid RDF/XML. This, along with a generator for handling HTTP requests, allows for the exposition of all parts of the system as Web Services
OntoViews
Pipeline Structure

1. User State Generator
2. Search Query XSLT
3. Ontogator Transformer
4. Logical Page Generation XSLT
5. Layout XSLT
6. XHTML Serializer
Summary – OntoViews is:

- **A search engine based on the semantics of the content:** Concept-based Multi-Facet and keyword search
- **Browsing functionality based on the semantic relations in the underlying knowledge base:** Classification tree view, explicit semantic links in item view
- **Usable for an end-user:** Developed interfaces based on UI research
- **Easily integrated and extended with additional functionality:** Seamless integration in the user interface of keyword and other searches, the search architecture allows for extensions, and the Cocoon control architecture forces a modular, reusable and extendable design. All components operate independently, consuming and/or producing RDF/XML.
- **Usable with a variety of different devices:** Different user interfaces and functionality for different devices
- **Adaptable to a wide variety of semantic data:** Facet hierarchy projection and semantic link generation are based on extendable logic rules. The projection rules have been tested with three different data sets.
- **Scalable to accommodate large amounts of data:** OntoViews has been tested with dmoz.org data to scale to up to 2.3 million items and 275,000 categories with search times of about 5 seconds.
- **Able to provide its functionality also to other programs as Semantic Web Services:** All subparts of the system are available to be used via Web Services
- **Future work:** Ideas about dynamic facet creation, further optimizations to improve scalability
View-Hierarchy-Based Browsing Interface: MuseumFinland virtual museum

- Every choice presents the user a visual view of possibly interesting targets, also giving ideas for additional search restrictions.
- Once a target is found, the interface moves on to semantic browsing.
- The goal is as fluid a browsing experience as possible, where even a user not familiar with the collection continually finds something interesting to pursue further.
- At the same time, the content of the collection and way of indexing become familiar to the user.
- Keyword search over categories provides a quick shortcut for those knowing precisely what they seek.
View-Hierarchy-Based Quick Search Interface: The Veturi Yellow Pages Service Portal

- Keyword search provides a familiar basis for the user
- The user gains a sense of security from seeing the semantic categories matched in their contexts
- At the same time again, the user gets an overview on how the content is indexed
- Semantic disambiguation can be done only as needed
- Still, the search is firmly semantically defined