Ontology and Application for Reusable Search Interface Design

Plans for Advanced Semantic Technologies Final Project

Eric Rozell, Tetherless World Constellation
Outline

• Project Overview
• Paper Contents
• Submission Plan
• Semantic Claims
• Project Status
Project Overview
Table of Contents

• Introduction
• Related Work
• Research Methods
• Framework Ontology
• Application Prototype
• Evaluation
• Discussion
• Conclusion
Introduction

• Outlines motivations for work
  – Challenges for data managers
    • Making user friendly interfaces
    • Managing a constant influx of data
  – Challenges for scientists
    • Data service discovery
    • Learning curve for search interfaces/services
Related Work

- Discusses related topics
  - Faceted Browsing Platforms
    - Not faceted browse in general
    - E.g., Exhibit, Longwell, Flamenco, mSpace
  - Virtual Observatories (VOs)
    - Integrate services and tools for uniform access
    - Some (SPASE, VSTO) serve as vocab. catalyst
    - S2S could benefit VO work (but not replace VOs)
  - Semantic Web Services (SWS)
    - SWS ontologies (e.g., OWL-S, WSMO, SWSO)
    - Semantic annotation of web services (e.g., SAWSDL, OpenSearch)
    - Utilize ontologies for discovery, annotation for extension
Research Methodology

Semantic Web Methodology & Technology Development Process

- Establish and improve a well-defined methodology vision for semantic technology based on application development
- Leverage controlled vocabularies, etc.

Open world: evolve, iterate redesign, redeploy
Rapid prototype
Leverage technology infrastructure
Adopt technology approach
Use tools
Science/expert reviews and iteration
Evaluation
Analysis
Develop model ontology
Small team, mixed skills
Use case
Framework Ontology

Legend
Yellow: Services
Blue: Query Interfaces
Orange: Output Formats
Green: Widgets
Pink: Parameters
Application Prototype
## Evaluation

### Table 1. Comparison of some Existing Faceted Browsing Applications

<table>
<thead>
<tr>
<th>Application</th>
<th>Desktop</th>
<th>Filtering</th>
<th>Display of Results</th>
<th>Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>iTunes [22]</td>
<td>Desktop</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>ToolBar 2.0D + Tool_facets [23]</td>
<td>Desktop</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Google Base [24]</td>
<td>Web</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Amazon [26]</td>
<td>Web</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>eBay [27]</td>
<td>Web</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>SEEK 1.0.1 [28]</td>
<td>Desktop</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Facet Map [9]</td>
<td>Desktop</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Flamenco [12]</td>
<td>Web</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Honeycomb [25]</td>
<td>Web</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>mSpace [14]</td>
<td>Web</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Exhibit [11]</td>
<td>Web</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Longwell [13]</td>
<td>Web</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

### General
1) configurable by user, 2) configurable by admin, 3) fixed facets only, 4) prices only
5) zip only

### Multiple Selection
6) but in a second step

### Range Selection
7) fixed

### Standard Classification
8) preview, only primaries, 9) missing subfolder integration

### String Search
10) all facets searched, 11) each single facet searchable

### Special Widgets
12) sliders, 13) sliders for price spans with very high values only, 14) more advanced widgets are in the source repository, but not yet published, 14a) some facets have multiple selection disabled

### Structure Explicit
15) categories as tree, 16) multi-columns represent some facet structure (e.g. date | month)

### Sorting
17) sorting by multiple facets.

### Grouping
18) displays tree relationship of exchange of letters on demand, 19) configuration follows sorting
## Evaluation

<table>
<thead>
<tr>
<th>Domain</th>
<th>Environment</th>
<th>Filtering</th>
<th>Display of Results</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Arbitrary Restriction Order</td>
<td>String Search Integrated</td>
<td>Sorting of Results Configurable</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Result Cardinalities Shown</td>
<td>Facet Values Explained</td>
<td>Grouping of Results</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multiple Selection</td>
<td>Special Widgets</td>
<td>Search for Similar Items</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Range Selection</td>
<td>Sorting of Results</td>
<td>Sorting of Results</td>
</tr>
</tbody>
</table>

| S2S    | Independent | Web | Yes | Yes | Yes | Yes | Yes | No² | No² | No³ | List, Table, Map | No⁵ | Yes |

1. Implementing standard classification integration for data services with multiple result types is left for future work.
2. Sorting and grouping, although not currently implemented, can be implemented using a specialized result widget and a sorting/grouping parameter.
3. We seek use cases and algorithms for finding “similar” items in data management systems.
4. Various kinds of result widgets can be implemented; lists and tables are currently implemented, and mapped results and others are planned.
5. We have use cases for context-dependent facets, which will be implemented in future work.
Evaluation

- **Case Study: BCO-DMO* Application Integration**
  - **BCO-DMO Technologies**
    - Web Map Service (WMS) / Web Feature Service (WFS)
    - MapServer
    - MySQL database
    - OWL ontology & RDF knowledge base
    - Vocabulary mappings to other projects
  - **Application Integration**
    - WMS/WFS search widget
    - MapServer result widget
    - OpenSearch service leverages MySQL and OWL/RDF
    - Vocab. mappings used in federated search

* Biological and Chemical Oceanography Data Management Office
Discussion

- More general than faceted browse
- Future Work
  - Support Apache Solr and SAWSDL standards
  - Usability evaluation (end-user and developer)
  - Web browser plugins
  - Community Portal
  - Discovery Infrastructure
Conclusion

- Sufficient infrastructure for faceted browse
- Platform for application integration
- Reusable user interface modules
- Extensible to various web standards
- Custom, uniform interfaces for scientists
Submission Plan

• Computers & Geosciences
  – Application Article
    • 5000 words
    • Novel software application or web service
    • Applied computer methods (OWL/RDF)
    • Comparison with other methods (Faceted Browsing)
    • Emphasis on novelty (Application Integration)
Semantic Claims

• **Claim:** S2S will use linked data practices for dereferenceable framework metadata.

• **Benefit:** Supports semi-automated discovery without the need for centralized registry.
Semantic Claims
Semantic Claims

• **Claim**: S2S leverages RDFS reasoning to infer sub-class transitivity for type checking and query.

• **Benefit**: Simplifies query development.
• User Scenario:
  – Oceanographer searching for data
  – Wishes to constrain search to a depth range (say, the within 50m of thermocline)
  – Needs to find widgets that are useful for “quantitative range”
  – Discovers class, s2s:QuantitativeRangeWidget
  – Makes request to system for instances of s2s:QuantitativeRangeWidget
  – System returns instances of class and subclasses
• **Claim:** S2S leverages OWL description logic constructs (note, not reasoning), namely universal qualifiers.

• **Benefit:** Allows for semi-automated construction of user interfaces.
User Scenario:

- Similar to last scenario, scientist finds a web service that S2S can be used for.
- Scientist selects service in S2S.
- System detects that the service uses s2s:QuantitativeRangeParameter
- System detects that s2s:QuantitativeRangeParameter can be constrained only with s2s:QuantitativeRangeWidget (universal qualifier)
- User scenario continues as before
Project Status

• First Draft Complete
• Submission Plan Developed
• Round of Friendly Reviews (starting)
• Future Work
  – Add a section between Introduction and Related Work summarizing the system
References

Grading

- Presentations: 10% ea.
- Paper: 20%
- …Or whatever is in my best interest