DAML Agent Semantic Communications Service (ASCS)

Teknowledge
John Li, Allan Terry

May 26, 2004
Semantic Search Services

Basic Facts about the Engine

- Crawls and searches the DAML/OWL web
- Returns answer to query with exact data, not just the relevant pages
- Broadens query with simple inference
  - Equivalence, inverse, generalization/specialization
- Crosses ontologies where there are mappings
- Available 7x24 for over 2.5 years
- Indexes over 7,000,000 OWL triples
- Most responses in 1-10 sec
- Java and Web Service APIs for use as a service
New Features

• Users can publish new data into repository through URL registration
• Users can subscribe queries for semantic search
  – Answers to subscribed queries arrive in email
  – Answers generated in event-based manner or on user-defined schedule
• Full Web service and Java APIs for agents
  • Basic query
  • URL registration and query subscription
  • Access to namespace vocabularies
Using the New Features

New features:
• Were developed mainly for the early adopters of semantic technologies in military and business organizations
• Help build an info-sphere for dynamic info exchange and sharing between producers and consumers
• Reduce maintenance work for hosts
• Increase situation awareness for the subscribed users
An Example
## Status

<table>
<thead>
<tr>
<th>Item</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic crawling &amp; search</td>
<td>Done</td>
</tr>
<tr>
<td>URL registration &amp; incremental indexing</td>
<td>In test</td>
</tr>
<tr>
<td>Three search GUIs</td>
<td>Done</td>
</tr>
<tr>
<td>Query subscription &amp; notification</td>
<td>In test</td>
</tr>
<tr>
<td>Java &amp; Web Service APIs</td>
<td>In test</td>
</tr>
<tr>
<td>Scalability and performance experiments</td>
<td>Preliminary data</td>
</tr>
<tr>
<td>Package and submit to SemWebCentral</td>
<td>Not started</td>
</tr>
</tbody>
</table>
Architecture

GUI or APIs

URL Registration → Crawler → URL Scrubber → Indexer

Search & Inference → OWL triples

Subscription & Notification → Ontology Mappings

Web Pages

Back End

Runtime
SemWebCentral

Open-source components
1. The back-end crawler & indexer
2. The basic and expert GUIs on top of the APIs
3. The ontology mapper utility

Government-Use licensing
1. The runtime search engine
Plan

• Quantify scalability and performance
• Extend the API test suites
• Package for adoption in enterprise intranets
  – Better documentation for software and APIs
  – Improved installation procedure
  – Existing GUIs as program samples
• Submit alpha version of components to SemWebCentral no later than August 1, 2004.