Semantic Data Management

Xavier Lopez, Ph.D., Director, Spatial & Semantic Technologies
Enterprise Information Challenge

Source: Oracle customer
Vision of Semantically Linked Data

The Network of Collaborative Knowledge

A web of data that can be intelligently processed by machines delivering an environment that:

• Intelligent
• Relevant
• Contextual
• Interconnected

• Solves Problems
• Answers Questions
• Easily Searchable
• Let Information follow you

“The next generation of information workers will expect a highly visual, connected, contextual information workplace they can take anywhere.”

- Forrester
Social Semantic Enterprise

Semantic Tagging
- Concept definitions with unique meaning
  - Language independent

Relational Navigation
- Recommendations
- Find related content & groups by navigating connected entities

Semantic Search
- Search for concepts
- Return semantically related results

Social Connectivite through Semantic Enablement

Island: Java

Semantc Search
- Search for concepts
- Return semantically related results
Semantic Web

• Integrate domains of knowledge through common vocabularies (ie SKOS)

• Manage relationships between collections of images and associated metadata

• RDF as flexible and extensible data model supports powerful search and end-user discovery of related content

• Rich platform for data integration, data repurposing, and better quality control and classification
The Web is about links…

The Semantic Web is about the relationships in those links

Dan Brickley
Metadata of documents

- **Author**: Auto Mechanic's Daughter (Black Goat)
- **Title**: Auto Mechanic's Daughter (Black Goat)
- **Publisher**: Paperback
- **Format**: Paperback
- **Price**: $13.15
- **List Price**: $14.95
- **You Save**: $1.79 (12%)
- **Free Shipping**: Eligible for FREE Super Saver Shipping on orders over $25

**In Stock**

Ships from and sold by Amazon.com, Gift-wrap available.

Want it delivered Tuesday, December 17? Order it in the next 15 hours and 11 minutes, and choose One-Day Shipping at checkout.

Ordering for Christmas? To ensure delivery by December 24, choose Standard Shipping at checkout. Read more about holiday shipping.

- 19 new from $2.04
- 19 used from $0.01
Linking Data Resources

User

Query & results

Data Ontologies

RDF

Structured & Unstructured Data Sources
Relational to Ontological Mapping

- Neuron
- Pathological Change
- Neuronal Property
- Channel
- Pathological Agent
- Agent
- Drug

<table>
<thead>
<tr>
<th>Compartment</th>
<th>Cell: NeuronDB</th>
<th>Receptor</th>
<th>Channel</th>
<th>Pathological Agent (PA)</th>
<th>PA Action</th>
<th>Drug</th>
<th>Drug Action</th>
<th>Stage</th>
<th>Note</th>
<th>Detail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soma</td>
<td>CA1 pyramidal neuron</td>
<td>1A</td>
<td></td>
<td>beta Amyloid</td>
<td>Inhibits</td>
<td></td>
<td></td>
<td>Early</td>
<td></td>
<td>66240</td>
</tr>
<tr>
<td></td>
<td>Olfactory bulb mitral cell</td>
<td>Gabaa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Early</td>
<td></td>
<td>66750</td>
</tr>
<tr>
<td>Dendrite</td>
<td>CA1 pyramidal neuron</td>
<td>1A</td>
<td></td>
<td>beta Amyloid</td>
<td>Inhibits</td>
<td></td>
<td></td>
<td>Early</td>
<td></td>
<td>66240</td>
</tr>
<tr>
<td></td>
<td>Olfactory bulb mitral cell</td>
<td>Gabaa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Early</td>
<td></td>
<td>66750</td>
</tr>
<tr>
<td>Unspecified</td>
<td>Oocyte</td>
<td></td>
<td>1 L high threshold</td>
<td>beta Amyloid</td>
<td>Inhibits</td>
<td></td>
<td></td>
<td>Early</td>
<td></td>
<td>66252</td>
</tr>
<tr>
<td></td>
<td>CA1 pyramidal neuron</td>
<td></td>
<td></td>
<td>beta Amyloid</td>
<td>Inhibits</td>
<td></td>
<td></td>
<td>Early</td>
<td></td>
<td>66753</td>
</tr>
<tr>
<td></td>
<td>CA1 pyramidal neuron</td>
<td></td>
<td></td>
<td>NMDA</td>
<td>beta Amyloid</td>
<td>Inhibits</td>
<td></td>
<td>Early</td>
<td></td>
<td>66258</td>
</tr>
<tr>
<td></td>
<td>CA1 pyramidal neuron</td>
<td></td>
<td></td>
<td>L Calcium</td>
<td>beta Amyloid</td>
<td>Inhibits</td>
<td></td>
<td>Early</td>
<td></td>
<td>66250</td>
</tr>
</tbody>
</table>

Courtesy: SenseLab, Yale University
Importance of SKOS

• SKOS provides a simple bridge between the “print world” and the (Semantic) Web
• Thesauri, glossaries, etc, from the library community
• can be made available
• LOC is a good example
• SKOS can also be used to organize tags, annotate
• other vocabularies, …
Ontology Integration: Bioinformatics
Oracle 11g RDF/OWL Graph Data Management

- **Storage & Loading**
  - Native W3C RDF graph data store
  - Fast Bulk, batch & Incremental load

- **Query**
  - SQL: SEM_MATCH graph pattern query
  - SPARQL: supported via Jena plug-in

- **Reasoning**
  - RDF, OWL Prime, RDF++ semantic rules
  - Forward chaining inference model
  - User defined rule base

- **Scalability**
  - Scales to billions of triples
  - Partitioning, RAC, Adv. Compression

- **Standards & Interoperability**
  - Aligned with W3C specifications
  - Supported by leading semantic tools
Semantic Technologies Adopters

Life Sciences

Defense/Intelligence

Education

Telecomm

Life Sciences

Defense/Intelligence

Education

Telecomm

Publishing

11g Reference

11g Reference

11g Reference

11g Reference

11g Reference

11g Reference

11g Reference

11g Reference

11g Reference
The Power of Semantic Connections

Delivering relevant, contextual content based on relationships, profile & activity
Cisco Knowledge Network

- RDF Store
- Social Graph
- Enterprise Vocabulary Import
  - e.g. “product catalog”, “Directory”
- User Entered Tags Import
  - e.g. "web content", "wiki topics", "expertise"
- Communities
  - e.g. “job roles”, “customer accounts”
- People
- Info
- Content Repositories
  - Structured Data
  - Transactional Applications

Entity and Property Extraction
Use Case: Integrative Data Mining and Ad Hoc Query

Tailored Therapeutic Workbench (TTW)

Integrative Data Mining/Query System

System/Semantic Integration Layer

XREP plus
Proteomics Informatics
PGI
IQ
Lipid Informatics
Info Mining

Gene Expression
Proteomics
Genotyping SNPs/Haplotypes
Imaging In-Vitro/Vivo
Lipidomics
Text

Translational Informatics

Source: WWW2007, Eli Lilly, S.Stephens
Use Case: Drug Discovery / Target Assessment Tool
Why is Oracle Spatial unique in Semantics Data Management?

- Oracle 11g is the only commercial database with native RDF/OWL semantics capability.
- Can readily scale to ultra-large repositories (+1 billion)
- Growing ecosystem of 3rd party tools partners
- Leverages Oracle Partitioning and Advanced Compression. RAC is also supported.
- Semantics customers expect to deploy very large servers
Extraction, Modeling, Reasoning & Discovery Workflow

Transform & Edit Tools
- Entity Extraction & Transform
  - OpenCalais
  - Linguamatics
  - GATE
- Ontology Engineering
  - TopQuadrant
  - Ontoprise
- Categorization
  - Cyc
- Custom Scripting

Load, Query & Inference
- RDF/OWL Data Management
- SQL & SPARQL Query
- Inferencing
- Semantic Rules
- Scalability & Security
- Semantic Indexing

Applications & Analysis Tools
- OBIEE, Analytics
- Graph Visualization
  - Cytoscape
- Social Network Analysis
- Metadata Registry
- Faceted Search

Partner Tools

Partner/Oracle Tools
Contacts and Resources:

Product Management: Xavier.Lopez@oracle.com

White papers, documentation, discussion forum, sample code, downloads, press releases:

Google: “Oracle RDF”