Semantic Technologies in Oracle Database 11g Release 2:
Capabilities and Architecture Overview

Souripriya Das, Ph.D., Architect, Oracle
## Semantic at OOW 2010 - Sessions

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday, Sept 20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:30 p.m.</td>
<td>How and Why Customers Use Oracle’s Semantic Database Technologies: A Panel</td>
<td>Moscone South Room 200</td>
</tr>
<tr>
<td>2:00 p.m.</td>
<td>Electronic Medical Records with Oracle Semantic Technologies at Cleveland Clinic</td>
<td>Moscone South Room 200</td>
</tr>
<tr>
<td>4:00 p.m.</td>
<td>How Cisco’s Enterprise Collaboration Platform Uses Oracle Semantic Technologies</td>
<td>Hotel Nikko</td>
</tr>
</tbody>
</table>

## Semantic at OOW 2010 – Hands-On Labs

<table>
<thead>
<tr>
<th>Date/Time</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tuesday, Sept 21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:00 p.m.</td>
<td>A Little Semantics Goes a Long Way with Oracle Database 11g</td>
<td>Hilton SF Franciscan A/B/C/D</td>
</tr>
</tbody>
</table>

- **DEMOgrounds**
  - Semantic Database Technologies - *Moscone West, W-045*
Agenda

• Introduction
• Capabilities
• Architecture
THE FOLLOWING IS INTENDED TO OUTLINE OUR GENERAL PRODUCT DIRECTION. IT IS INTENDED FOR INFORMATION PURPOSES ONLY, AND MAY NOT BE INCORPORATED INTO ANY CONTRACT. IT IS NOT A COMMITMENT TO DELIVER ANY MATERIAL, CODE, OR FUNCTIONALITY, AND SHOULD NOT BE RELIED UPON IN MAKING PURCHASING DECISION. THE DEVELOPMENT, RELEASE, AND TIMING OF ANY FEATURES OR FUNCTIONALITY DESCRIBED FOR ORACLE'S PRODUCTS REMAINS AT THE SOLE DISCRETION OF ORACLE.
Introduction

Semantic Technologies in Oracle 11g Release 2
Semantic Technology Stack

User Interface & Applications

Trust

Proof

Unifying Logic

Query: SPARQL

Ontology: OWL

Rule: RIF

Data interchange: RDF

XML

URI/IRI

Crypto

http://www.w3.org/2007/03/layerCake.svg
Extraction, Modeling, Reasoning & Discovery Workflow

Transform & Edit Tools
- Entity Extraction & Transform
- Ontology Engineering
- Categorization
- Custom Scripting

Load/Query/Inference
- RDF/OWL Data Management
- SPARQL Query and Update
- RDFS/OWL Inferencing
- User-def. Rules
- SQL-based Query
- Security (OLS/VPD)
- Semantic Indexing
- Versioning

Applications & Analysis Tools
- BI Analytics
- Graph Visualization
- Social Network Analysis
- Metadata Registry
- Faceted Search
- SPARQL Endpoint

Partner Tools

Partner/Oracle Tools
Overview

Semantic Technologies in Oracle 11g Release 2
## Capabilities: Overview

### Core functionality
- **Loading**
  - Bulk or incremental
- **SPARQL query and update**
  - Jena (w/ Jena Adapter)
  - Sesame (w/ Sesame Adapter)
  - Joseki SPARQL endpoint
  - Sesame/Workbench endpoint
- **Inference**
  - OWL 2 RL, OWLPrime, RDFS
  - User-defined Rules
  - SKOSCORE
  - SNOMED

### Enterprise functionality
- **“SPARQL in SQL”**
  - SEM_MATCH table function
- **Ontology-assisted SQL query**
  - SEM_RELATED operator
- **Semantic indexing of documents**
  - SEM_CONTAINS operator
- **Fine-grained security (@triple)**
  - Oracle Label Security (OLS)
  - Virtual Private Database (VPD)
- **Version management**
  - Workspace Manager
Architecture: Overview

Core functionality:
- LOAD: Bulk-Load, Incr. DML
- QUERY (SQL-based SPARQL): Query RDF/OWL data and ontologies, Ontology-assisted Query of Enterprise Data

Oracle DB:
- RDF/OWL data and ontologies
- Rulebases: OWL, RDF/S, user-defined
- Inferred RDF/OWL data

Security: fine-grained
Versioning: Workspaces

Enterprise (Relational) data

Tools (Cytoscape):
- Joseki / Sesame
- Topbraid Composer

Programming Interface:
- SPARQL: Jena / Sesame
- Java Programs
- JDBC

3rd Party Tools:
- SQL Interface
  - SQLPlus
  - PL/SQL
  - SQLDev

Callouts: Pellet, NLP Extractor: CTX_ENTITY, Calais, GATE
SPARQL Query Architecture

HTTP

Standard SPARQL Endpoint

Java

Jena API
Jena Adapter

Sesame API
Sesame Adapter

SQL

SEM_MATCH
Enterprise Functionality
Semantic Technologies in Oracle 11g Release 2
SEM_MATCH: Adding SPARQL to SQL

**SPARQL**

```
PREFIX foaf: <http://...>  
SELECT ?n1 ?n2  
FROM <http://g1>  
WHERE  
{ ?p foaf:name ?n1  
  OPTIONAL {?p foaf:knows ?f .  
    ?f foaf:name ?n2 }  
  FILTER (REGEX(?n1, "^A")) }
```

**SQL**

```
SELECT n1, n2  
FROM TABLE(SEM_MATCH(  
  '{?p foaf:name ?n1  
    OPTIONAL {?p foaf:knows ?f .  
      ?f foaf:name ?n2 }  
    FILTER (REGEX(?n1, "^A")) }',  
  SEM_MODELS('g1'),...,  
  SEM_ALIASES(  
    SEM_ALIAS('foaf', 'http://...')))
```

**Keywords:**
- **projection**
- **data selection**
- **graph pattern**
- **prefixes**
GeoSpatial Semantic Querying in an RDBMS

SQL-Based Approach

Oracle 11g

Object-Relational Data (Spatial Objects)

Triple Store (Semantic Data)

Geo-Semantic SQL-Based Querying: SEM_MATCH + SDO_RELATE
Semantic Indexing of Documents and Querying using SEM_CONTAINS operator

CREATE INDEX ArticleIndex
ON NewsFeed (Article)
INDEXTYPE IS SemContext
PARAMETERS (‘gate_nlp’)

SELECT docId FROM Newsfeed
WHERE SEM_CONTAINS (Article,
  '{?x rdf:type rc:Person .
   ?x :hasAge ?age .
   FILTER(?age >= 35)}')=1
AND Source = ‘CNN’
Enterprise Security for Semantic Data: OLS (Oracle Label Security)

[OLS] Data classification labels for semantic data

SELECT operation: Labels for triples are used to restrict access to the triples.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
<th>Row Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>projectHLS</td>
<td>Organization</td>
<td>N.America</td>
<td>SE:HLS:US</td>
</tr>
<tr>
<td>projectHLS</td>
<td>ContractValue</td>
<td>1000000</td>
<td>SE:HLS,FIN:US</td>
</tr>
</tbody>
</table>

Label Security Modes

- **Triple-level**: Easier to use. Any one can insert, bulk-load, do inference. (Inserted/inferred triples get user’s label. Query/Infer on visible triples.)
- **Resource-level**: Secures each resource independently.
Enterprise Security for Semantic Data: VPD (Virtual Private Database)

[VPD] Access control policies on semantic data

**Policy ➔ user can access value of projects s/he leads**

**Match pattern ➔**
```
{ ?x :hasValue ?v }
```

**Apply pattern ➔**
```
{ ?x :hasLead “sys_context(...)” }
```

**Query : Get the list of projects and their values**

```
SELECT ?proj ?val
FROM ProjectsGraph
WHERE
{ ?proj :hasValue ?val .
  ?proj :hasLead “sys_context(...)” }
```
Ontology-assisted Query using SQL Operators

SELECT p_id, diagnosis FROM Patients WHERE
WHERE SEMRELATED (diagnosis, ‘rdfs:subClassOf’, ‘Upper_Extremity_Fracture’, sem_models(‘Medical_ontology’), sem_rulebases(‘RDFS’) ...) = 1
AND SEMDISTANCE(123) <= 2;

ID | DIAGNOSIS         |
---|-------------------|
1  | Hand_Fracture     |
2  | Rheumatoid_Arthritis |
Change Mgmt./Versioning for Sem. Data

- Manage public and private versions of semantic data in database workspaces (Workspace Manager)

- An RDF Model is version-enabled by version-enabling the corresponding application table.
  
  ```
  exec DBMS_WM.enableVersioning (table_name => 'contracts_rdf_data');
  ```

- RDF data modified within a workspace is private to the workspace until it is merged.

- SEM_MATCH queries on version-enabled models are version aware and only return relevant data.

- New versions created only for changed data

- Versioning is provisioned for inference
<table>
<thead>
<tr>
<th><strong>Ontology Engineering</strong></th>
<th><strong>Reasoners</strong></th>
<th><strong>Applications</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>TopQuadrant</td>
<td>clarkparsia, llc</td>
<td>TERANODE</td>
</tr>
<tr>
<td>Ontoprise</td>
<td>Ontoprise</td>
<td>metatomix</td>
</tr>
<tr>
<td>protégé</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Query Tool Interfaces</strong></th>
<th><strong>Standards</strong></th>
<th><strong>NLP Entity Extractors</strong></th>
<th><strong>SI / Consulting</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jena</td>
<td>RDF.org</td>
<td>CALAIS</td>
<td>NORTHROP GRUMMAN</td>
</tr>
<tr>
<td>Sesame</td>
<td></td>
<td>Linguamatics</td>
<td>Raytheon</td>
</tr>
<tr>
<td>Joseki</td>
<td></td>
<td>LYMBA</td>
<td>ORBIS TECHNOLOGIES</td>
</tr>
<tr>
<td>cYcorp</td>
<td>W3C</td>
<td>GATE</td>
<td>BOEING</td>
</tr>
</tbody>
</table>

|               |               |                           | Boeing              |
|               |               |                           | McDonald Bradley    |
|               |               |                           | Cognia              |
|               |               |                           | Accenture           |

**Oracle’s Partners for Semantic Technologies**

**Integrated Tools and Solution Providers:**
Some Oracle Database Semantics Customers

Life Sciences
- Lilly
- Pfizer

Defense/Intelligence
- National Geospatial-Intelligence Agency

Education
- The University of Michigan

Clinical Medicine & Research
- Swiss Institute of Bioinformatics
- Health Science Center at Houston
- Cleveland Clinic

Telecomm & Networking
- Hutchinson 3G
- Austria

Publishing
- Thomson Reuters

Telecomm & Networking
- CISCO

Publishing
- Westlaw
For More Information

search.oracle.com

Semantic Technologies

or

oracle.com
ORACLE IS THE INFORMATION COMPANY
Concepts
RDF: Integration and Inference
### Relational table vs. RDF

<table>
<thead>
<tr>
<th>Name</th>
<th>hasFather</th>
<th>hasMother</th>
<th>hasSister</th>
</tr>
</thead>
<tbody>
<tr>
<td>:Kate</td>
<td>:Sam</td>
<td>:Sue</td>
<td></td>
</tr>
<tr>
<td>:Jack</td>
<td>:Sam</td>
<td>:Sue</td>
<td>:Kate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>:Kate</td>
<td>:hasFather</td>
<td>:Sam</td>
</tr>
<tr>
<td>:Kate</td>
<td>:hasMother</td>
<td>:Sue</td>
</tr>
<tr>
<td>:Jack</td>
<td>:hasFather</td>
<td>:Sam</td>
</tr>
<tr>
<td>:Jack</td>
<td>:hasMother</td>
<td>:Sue</td>
</tr>
<tr>
<td>:Jack</td>
<td>:hasSister</td>
<td>:Kate</td>
</tr>
</tbody>
</table>
### Integration

<table>
<thead>
<tr>
<th>Name</th>
<th>hasFather</th>
<th>hasMother</th>
<th>hasSister</th>
</tr>
</thead>
<tbody>
<tr>
<td>:Kate</td>
<td>:Sam</td>
<td>:Sue</td>
<td></td>
</tr>
<tr>
<td>:Jack</td>
<td>:Sam</td>
<td>:Sue</td>
<td>:Kate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Plays</th>
</tr>
</thead>
<tbody>
<tr>
<td>:Kathryn</td>
<td>:Chess</td>
</tr>
<tr>
<td>:Kathryn</td>
<td>:Frisbee</td>
</tr>
</tbody>
</table>

In RDF:

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>:Kate</td>
<td>:hasFather</td>
<td>:Sam</td>
</tr>
<tr>
<td>:Kate</td>
<td>:hasMother</td>
<td>:Sue</td>
</tr>
<tr>
<td>:Jack</td>
<td>:hasFather</td>
<td>:Sam</td>
</tr>
<tr>
<td>:Jack</td>
<td>:hasMother</td>
<td>:Sue</td>
</tr>
<tr>
<td>:Jack</td>
<td>:hasSister</td>
<td>:Kate</td>
</tr>
<tr>
<td>:Kathryn</td>
<td>:plays</td>
<td>:Chess</td>
</tr>
<tr>
<td>:Kathryn</td>
<td>:plays</td>
<td>:Frisbee</td>
</tr>
<tr>
<td>:Kathryn</td>
<td>owl:sameAs</td>
<td>:Kate</td>
</tr>
</tbody>
</table>

Kathryn = Kate
Inference

<table>
<thead>
<tr>
<th>Name</th>
<th>hasFather</th>
<th>hasMother</th>
<th>hasSister</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>:Kate</td>
<td>:Sam</td>
<td>:Sue</td>
<td></td>
<td>:Female</td>
</tr>
<tr>
<td>:Jack</td>
<td>:Sam</td>
<td>:Sue</td>
<td>:Kate</td>
<td></td>
</tr>
<tr>
<td>:Sam</td>
<td></td>
<td></td>
<td>:Kate</td>
<td>:Male</td>
</tr>
<tr>
<td>:Sue</td>
<td></td>
<td></td>
<td></td>
<td>:Female</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Predicate</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>:hasFather</td>
<td>rdfs:range</td>
<td>:Male</td>
</tr>
<tr>
<td>:hasMother</td>
<td>rdfs:range</td>
<td>:Female</td>
</tr>
<tr>
<td>:hasSister</td>
<td>rdfs:range</td>
<td>:Female</td>
</tr>
</tbody>
</table>

supplied domain semantics