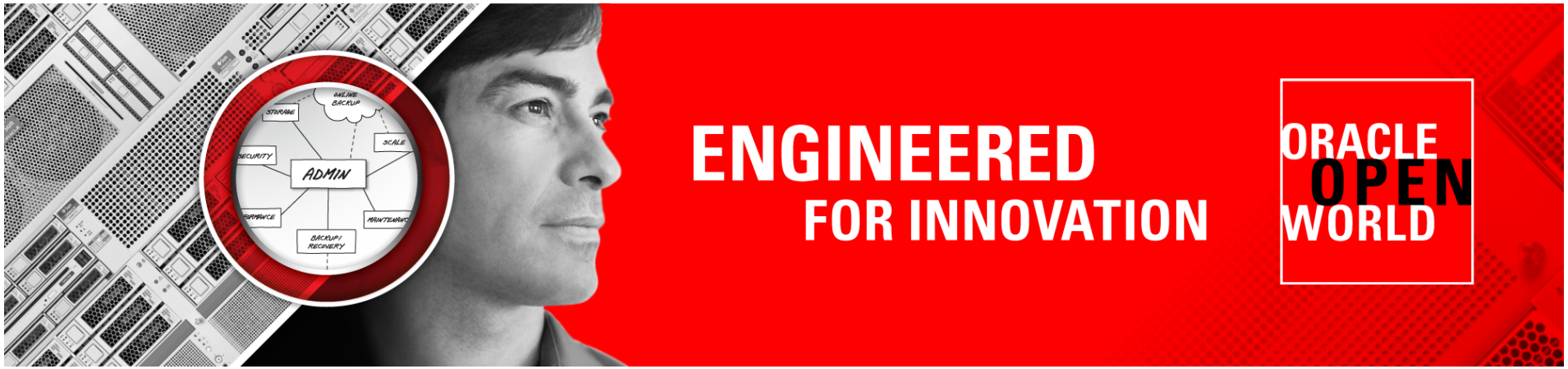
ORACLE®



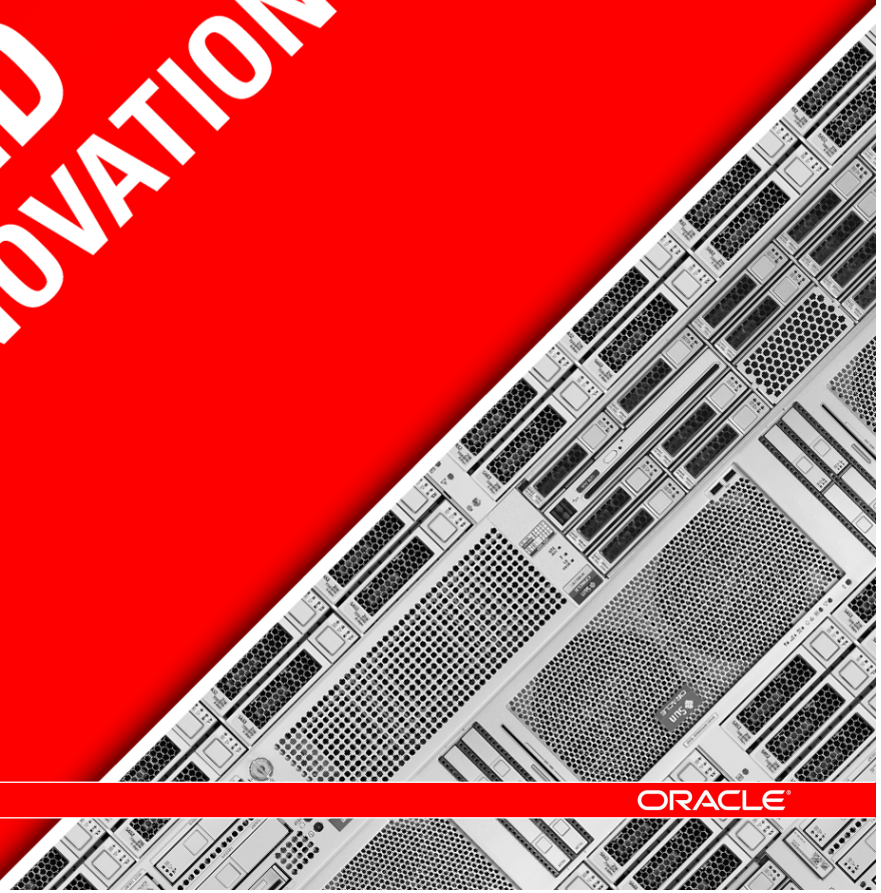
ORACLE®

Oracle Database Semantic Technologies Demos

Oct. 2011

ORACLE OPEN WORLD

ENGINEERED FOR INNOVATION



ORACLE

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 decisions. The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

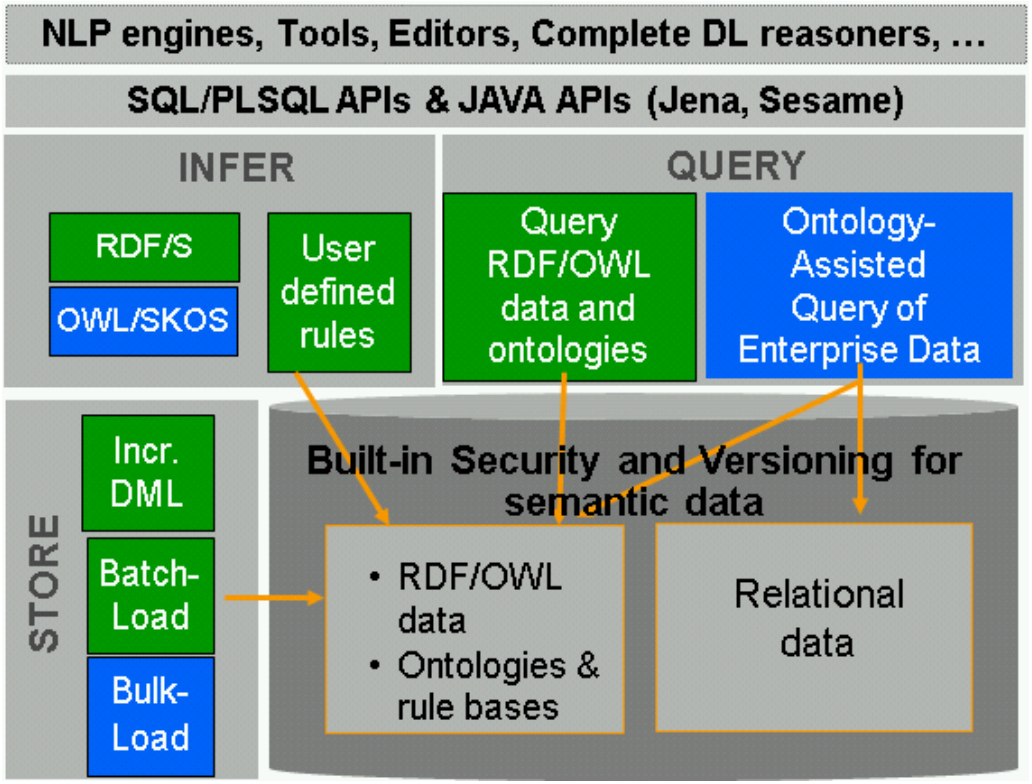
Demos

- Core Database Semantic Technologies Features
  - Load, Inference, Query
- Semantic Indexing of Documents
  - Lymba integration (Video available)
- Semantic Data Visualization/Editing
  - Tom Sawyer integration (Video available)
  - Cytoscape integration
  - TopBraid Composer
- OBIEE Integration with RDF Data
- Federated Query

Demo:
Core Database Semantic Technologies Features

ORACLE

Capabilities Overview of Release 11.2



Semantic Indexing- Key Components

- Extensible Information Extractor
  - Programmable API to plug-in 3rd party extractors into the database.
- SemContext Indextype
  - A custom indexing scheme that interacts with the extractor to manage the metadata extracted from the documents efficiently and facilitates semantic search via SQL queries.
- SEM_CONTAINS Operator
  - To identify documents of interest based on their extracted metadata, using standard SQL queries.
- SEM_CONTAINS_SELECT Ancillary Operator
  - To return additional information (SPARQL Query Results XML) about the documents identified using SEM_CONTAINS operator.

Semantic Indexing- Overview

```
CREATE INDEX ArticleIndex
 ON Newsfeed (Article)
 INDEXTYPE IS SemContext
 PARAMETERS ('my_policy')
```

Triples *table with rowid references*

Subject	Property	Object	graph
p:Marcus	rdf:type	rc::Person	<…/r1>
p:Marcus	:hasName	"Marcus"^^…	<…/r1>
p:Marcus	:hasAge	"38"^^xsd:..	<…/r1>
…	…	…	…

Analytical Queries On Graph Data

Newsfeed *table*

	docId	Article	Source
r1	1	Indiana authorities filed felony charges and a court issued an arrest warrant for a financial manager who apparently tried to fake his death by crashing his airplane in a Florida swamp. Marcus Schrenker, 38 …	CNN
r2	2	Major dealers and investors …	NW
	..		..

Rowid

SemContext index on Article column

```
SELECT Sem_Contains_Select(1)
FROM Newsfeed
 WHERE Sem_Contains (Article,
        '{?x rdf:type rc:Person .
          ?x :hasAge ?age .
          FILTER(?age >= 35)}',1)=1
    AND Source = 'CNN'
```

ORACLE

Performance of Oracle's Semantic Technologies

ORACLE

Performance

Oracle Performs Best on a Balanced Hardware Setup!

Degrees of Parallelism	Data set	Load (B triples/min.)	OWL Inference (M triples/min.)	Query (M answers/min.)
128	LUBM 8K	1.1B / 28' 11" *	869M / 62'	149+ m in 4.3'
64	LUBM 8K	1.1B / 53' 49"	869M / 114'	149+ m in 4.3'

Setup:

Balanced Hardware: Sun M8000 -- 16 SPARC64 VII+ 3.0 GHz CPUs [64 total Cores 128 Parallel Threads] -- 512 GB Ram -- Dual F5100 Flash Arrays (160 total drives)

Storage required: 330GB + 110GB of temporary table space

Software: Oracle Database 11.2.0.2.0 + Patch 9825019: SEMANTIC TECHNOLOGIES 11G R2 FIX BUNDLE 3
SGA_TARGET=256G and PGA_AGGREGATE_TARGET=206G

* 1.1B / 28' 11" means 1.1 billion triples in 28 minutes and 11 seconds

ORACLE

Performance

Oracle Performs Best on a Balanced Hardware Setup!

Degrees of Parallelism	Data set	Load (B triples/min.)	OWL Inference (B triples/min.)	Query (B answers/min.)
128	LUBM 25K	3.4B/105’	2.7B/160’	0.47B/8.7’
64	LUBM 25K	3.4B/186’	2.7B/210’	0.47B/12.3’

Setup:

Balanced Hardware: Sun M8000 -- 16 SPARC64 VII+ 3.0 GHz CPUs [64 total Cores 128 Parallel Threads] -- 512 GB Ram -- Dual F5100 Flash Arrays (160 total drives)

Storage required: 900GB + 300GB of temporary table space

Software: Oracle Database 11.2.0.2.0 + Patch 9825019: SEMANTIC TECHNOLOGIES 11G R2 FIX BUNDLE 3
SGA_TARGET=256G and PGA_AGGREGATE_TARGET=206G

ORACLE

Semantics in Oracle Database: Summary

Summary: Oracle Database Semantic Technologies

- Scalable to billions of triples
- RAC & Exadata scalability
- Compression & partitioning
- SQL*Loader direct path load
- Parallel load, inference, query
- Oracle DataGuard availability
- Triple-level DoD-strength security
- Choice of SPARQL or SQL
- Native inference engine
- W3C standards compliance
- Semantic Indexing of text & docs
- Growing ecosystem of 3rd party tools partners

W3C WORLD WIDE WEB consortium

RDF

Key Capabilities:

Load / Storage
- Native RDF graph data store
- Manages billions of triples
- Optimized storage architecture

Query
- SPARQL-Jena/Joseki, Sesame
- SQL/graph query, b-tree indexing
- Ontology assisted SQL query

Reasoning
- RDFS, OWL2 RL, EL+, SKOS
- User-defined SWRL-like rules
- Incremental, parallel reasoning
- Plug-in architecture

ORACLE

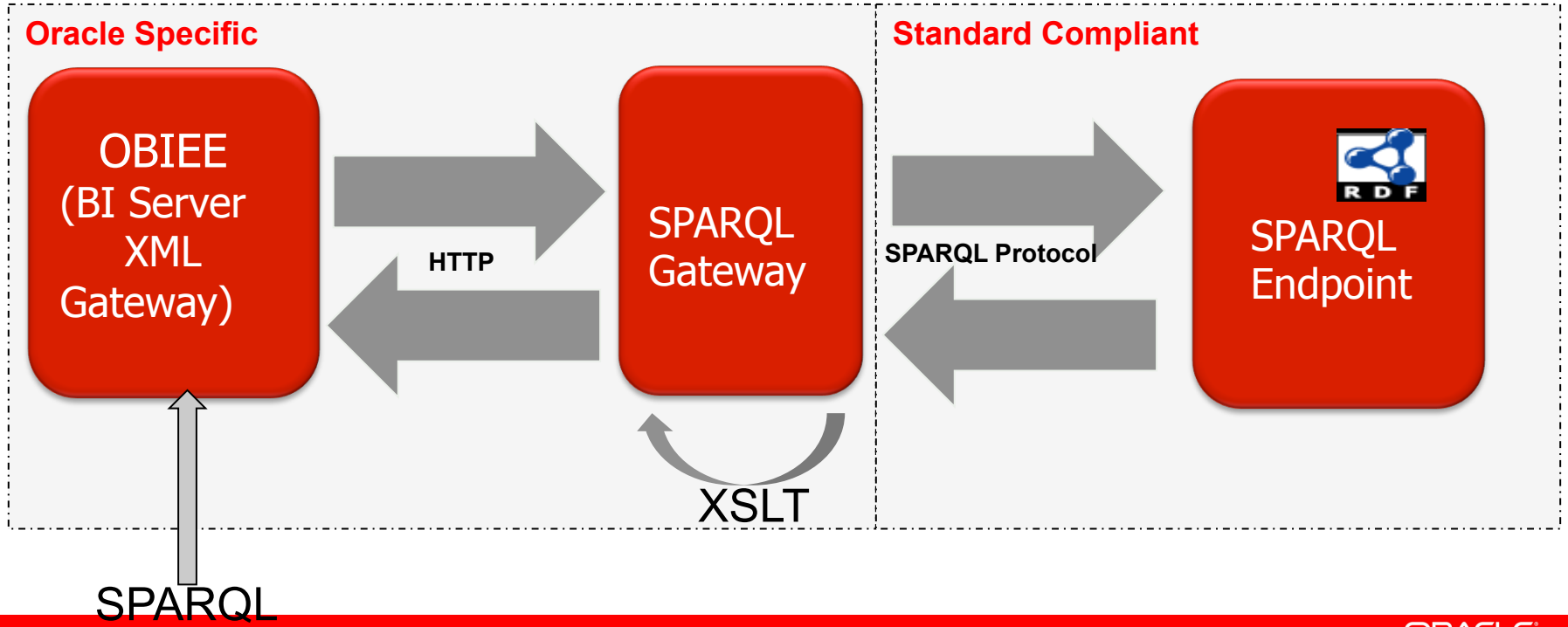
Demo:

OBIEE Integration with RDF Data

(Screenshots)

ORACLE

SPARQL Gateway



Investment RDF Data from dava.gov

http://se...b1.sparql | SPARQL Gateway | http://127...b4.sparql | McAfee Web Gatew... | McAfee Web Gatew...

SPARQL Gateway

home browse sparql xslt

simple navigation interface. Welcome!

SPARQL Endpoint: http://127.0.0.1:8080/joseki/invest local

SPARQL SELECT Query Body: Load example query: Fetch subclass & superclass

```
prefix gov: <http://data-gov.tw.rpi.edu/vocab/p/1565/>
select ?s ?class_ticker ?series_name  ?class_name ?zip_code ?state ?city ?entity_org_type
?address_1 ?address_2
{?s   gov:class_ticker ?class_ticker .
  ?s  gov:series_name ?series_name .
  ?s  gov:class_name ?class_name .
  ?s  gov:zip_code ?zip_code .
  ?s  gov:state ?state .
  ?s  gov:city ?city .
  ?s  gov:entity_org_type ?entity_org_type .
  ?s  gov:address_1 ?address_1 .
```

Submit Query

Hide Advanced Options Timeout (ms) 3000 1000ms (1s) Best effort

View All « Prev 1 2 3 4 Next »

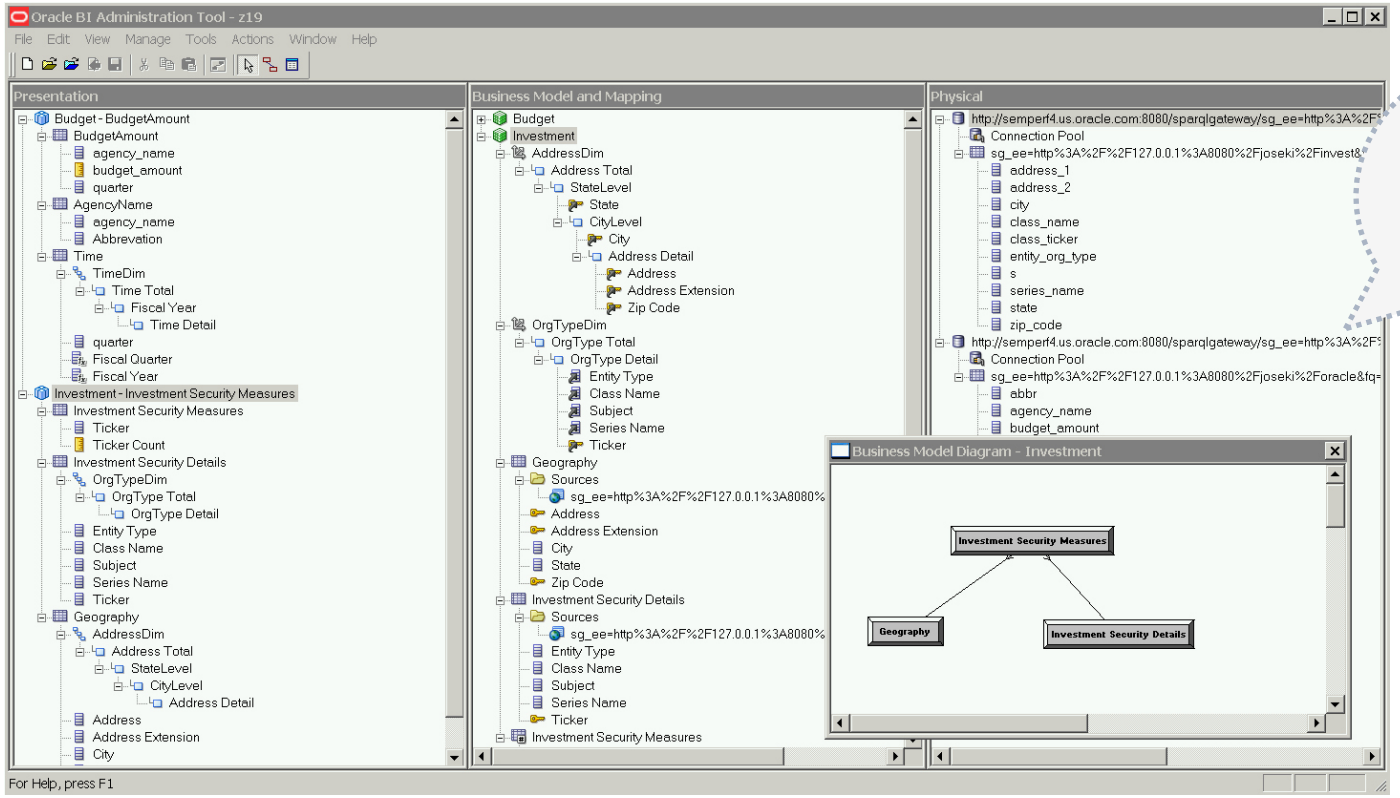
Row Count	S	CLASS_TICKER	SERIES_NAME	CLASS_NAME	ZIP_CODE	STATE	CITY	ENTITY_ORG_TYPE	ADDRESS_1	ADDRESS_
1	entry23631	WPSGX	STEWART W P & CO GROWTH FUND INC	STEWART W P & CO GROWTH FUND INC	10022	NY	NEW YORK	30	527 MADISON AVENUE	
2	entry476	FCNTX	Fidelity Contrafund	Fidelity Contrafund	2109	MA	BOSTON	30	FIDELITY INVESTMENTS COMPANY	82 DEVONSHIR STREET

Transformed SPARQL Query Result

This XML file does not appear to have any style information associated with it. The document tree is shown below.

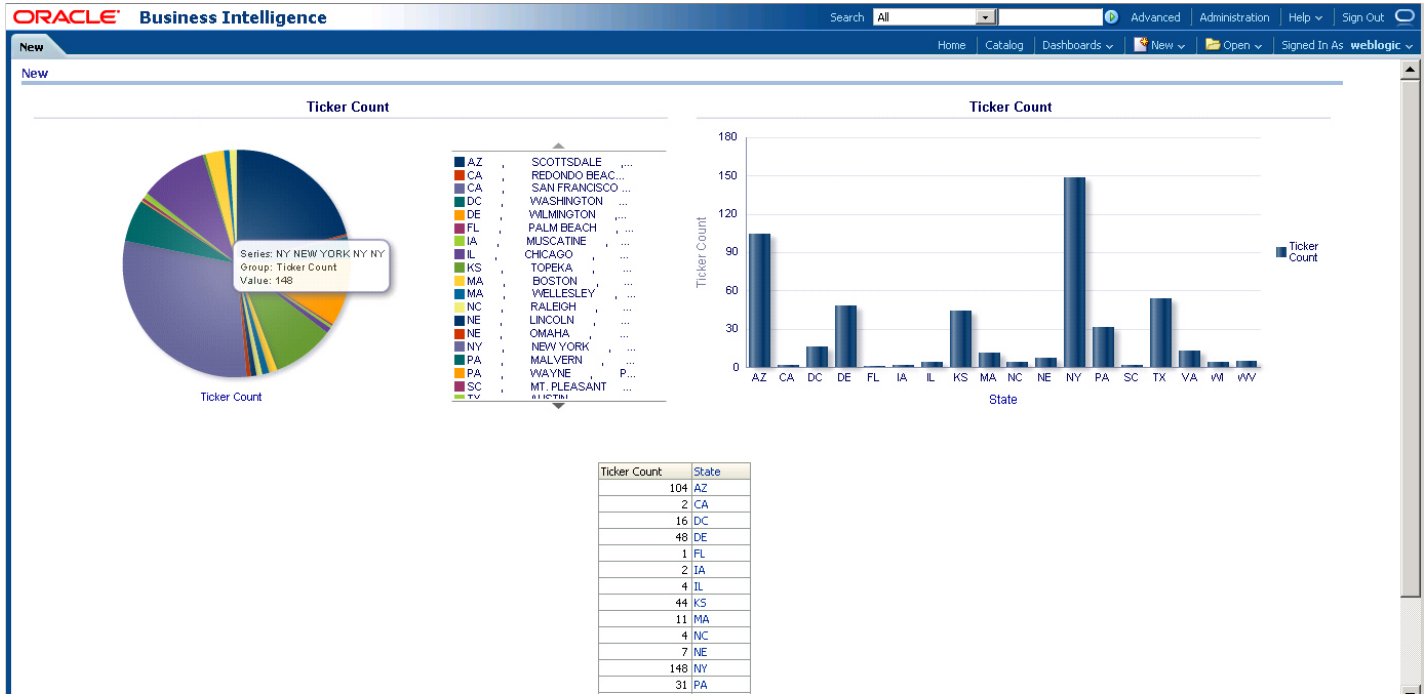

```
- <test>
  - <row>
      <s> entry23631 </s>
      <class_ticker> WPSGX </class_ticker>
      <series_name> STEWART W P & CO GROWTH FUND INC </series_name>
      <class_name> STEWART W P & CO GROWTH FUND INC </class_name>
      <zip_code> 10022 </zip_code>
      <state> NY </state>
      <city> NEW YORK </city>
      <entity_org_type> 30 </entity_org_type>
      <address_1> 527 MADISON AVENUE </address_1>
      <address_2/>
    </row>
  - <row>
      <s> entry476 </s>
      <class_ticker> FCNTX </class_ticker>
      <series_name> Fidelity Contrafund </series_name>
      <class_name> Fidelity Contrafund </class_name>
      <zip_code> 2109 </zip_code>
      <state> MA </state>
      <city> BOSTON </city>
      <entity_org_type> 30 </entity_org_type>
      <address_1> FIDELITY INVESTMENTS COMPANY </address_1>
      <address_2> 82 DEVONSHIRE STREET </address_2>
    </row>
  - <row>
      <s> entry475 </s>
      <class_ticker> FINSX </class_ticker>
      <series_name> Fidelity Advisor New Insights Fund </series_name>
      <class_name> Institutional Class </class_name>
      <zip_code> 2109 </zip_code>
      <state> MA </state>
      <city> BOSTON </city>
      <entity_org_type> 30 </entity_org_type>
      <address_1> FIDELITY INVESTMENTS COMPANY </address_1>
      <address_2> 82 DEVONSHIRE STREET </address_2>
    </row>
  - <row>
      <s> entry473 </s>
```

Business Model

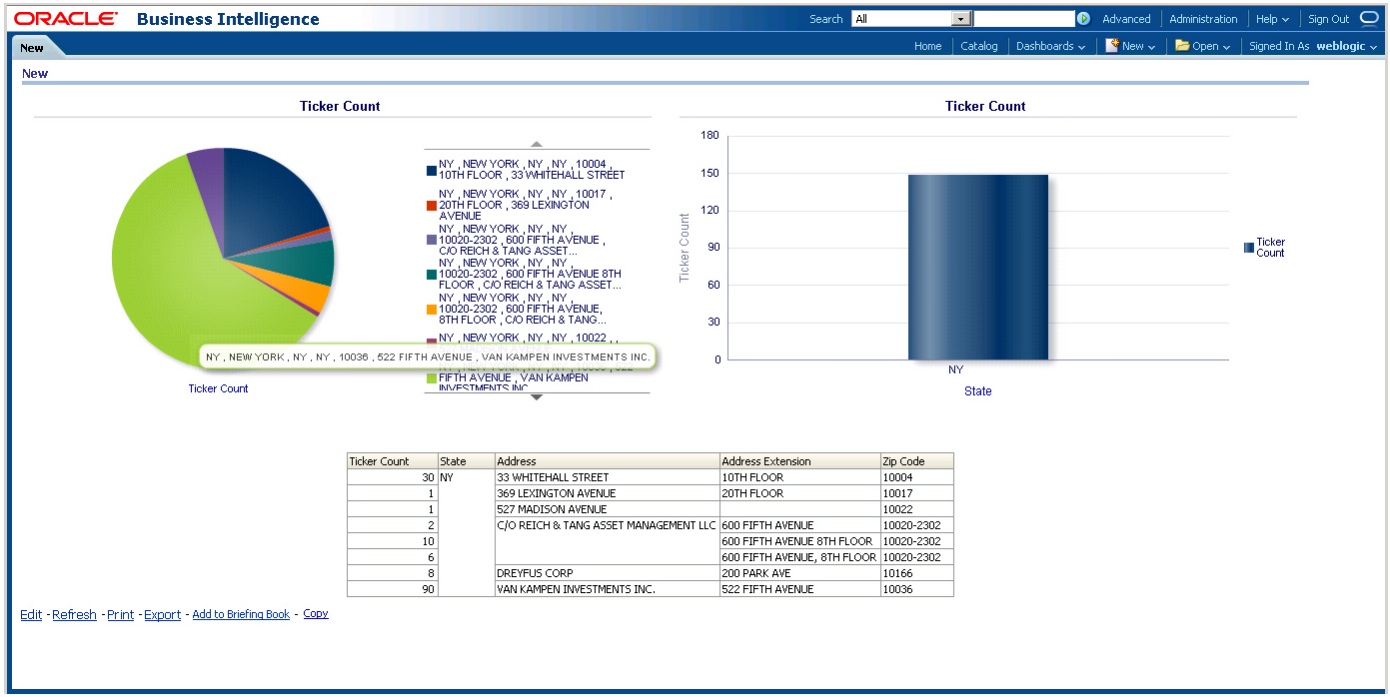


Relational, XML, … or RDF data sources

OBIEE Report



OBIEE Report



Finance RDF Data from DoD

http://se...b1.sparql | SPARQL Gateway | http://127...b4.sparql | McAfee Web Gatew... | SPARQL Gateway

SPARQL Gateway

home browse sparql xslt

A simple navigation interface. Welcome!

SPARQL Endpoint: http://127.0.0.1:8080/joseki/oracle local

SPARQL SELECT Query Body: Load example query: Fetch subclass & superclass

```
PREFIX owl:    <http://www.w3.org/2000/07/owl#>
PREFIX fn:     <http://www.w3.org/2005/xpath-functions#>
PREFIX ouext:  <http://oracle.com/semtech/jena-adaptor/ext/user-def-function#>
PREFIX oext:   <http://oracle.com/semtech/jena-adaptor/ext/function#>
PREFIX ORACLE_SEM_FS_NS: <http://oracle.com/semtech#timeout=100,qid=123>
SELECT ?agency_name ?abbr ?budget_amount ?quarter
{?agency_name <http://localhost/ontologies/budget.owl#hasAbbreviation> ?abbr .
 ?agency_name <http://localhost/ontologies/budget.owl#hasBudgetAmount> ?budget_amount .
 ?agency_name <http://localhost/ontologies/budget.owl#hasQuarter> ?quarter .
}
LIMIT 100
```

Submit Query

Show Advanced Options

Row Count	AGENCY_NAME	ABBR	BUDGET_AMOUNT	QUARTER
1	..se_Advanced_Research_Projects_Agency4	DEF. ADV. RESEARCH PROJ.	4630383 datatype: xsd:decimal	FY2010Q4
2	http://localhost/United_States_Navy4	UNITED STATES NAVY	1083024 datatype: xsd:decimal	FY2010Q4
3	http://localhost/United_States_Navy4	UNITED STATES NAVY	392679 datatype: xsd:decimal	FY2010Q4
4	http://localhost/United_States_Navy4	UNITED STATES NAVY	235980 datatype: xsd:decimal	FY2010Q4

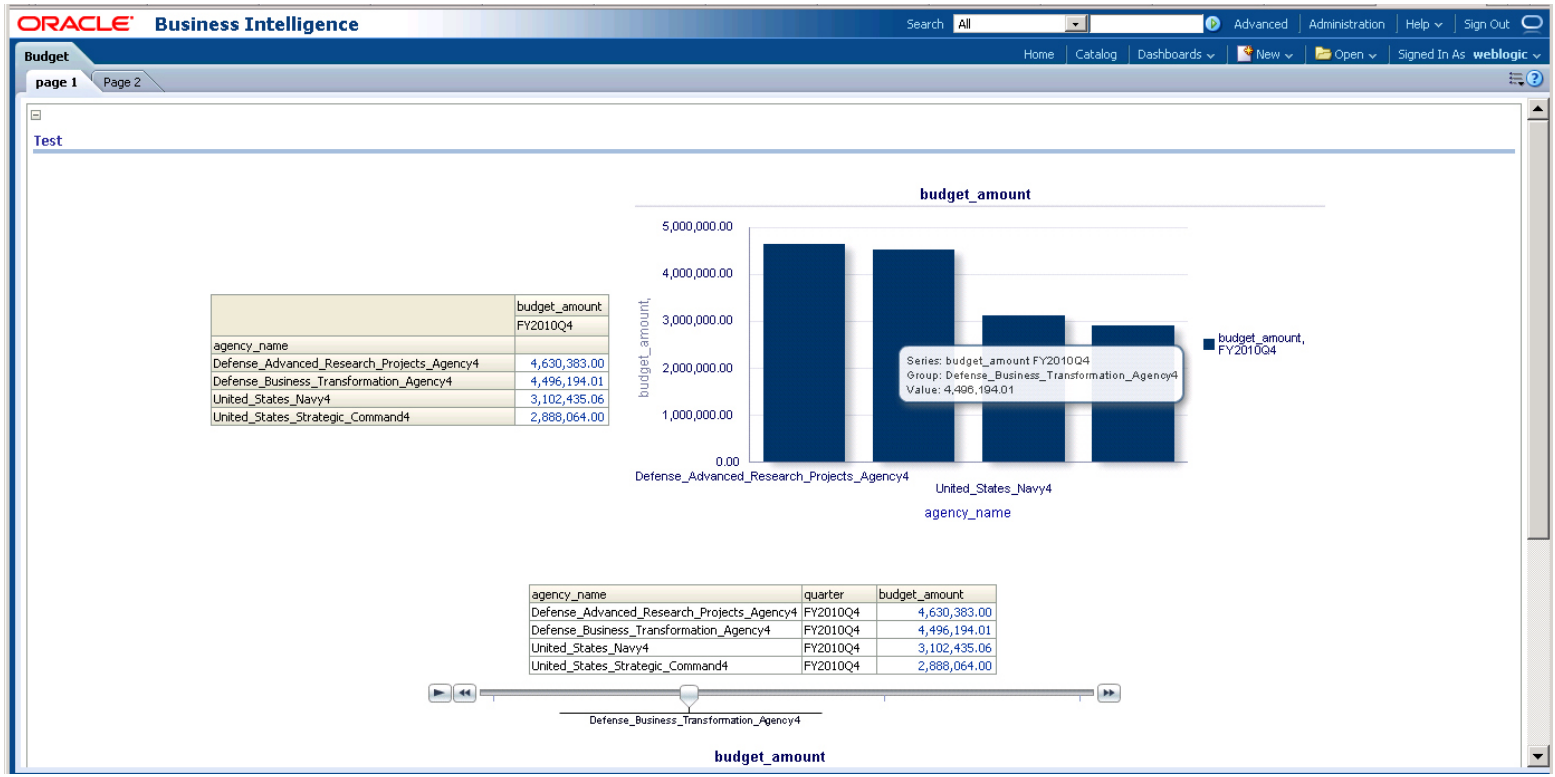
Transformed SPARQL Query Results

This XML file does not appear to have any style information associated with it. The document tree is shown below.


```
- <test>
  - <row>
    - <agency_name>
        Defense_Advanced_Research_Projects_Agency4
      </agency_name>
      <abbr> DEF. ADV. RESEARCH PROJ. </abbr>
      <budget_amount> 4630383 </budget_amount>
      <quarter> FY2010Q4 </quarter>
    </row>
  - <row>
      <agency_name> United_States_Navy4 </agency_name>
      <abbr> UNITED STATES NAVY </abbr>
      <budget_amount> 1083024 </budget_amount>
      <quarter> FY2010Q4 </quarter>
    </row>
  - <row>
      <agency_name> United_States_Navy4 </agency_name>
      <abbr> UNITED STATES NAVY </abbr>
      <budget_amount> 392679 </budget_amount>
      <quarter> FY2010Q4 </quarter>
    </row>
  - <row>
      <agency_name> United_States_Navy4 </agency_name>
      <abbr> UNITED STATES NAVY </abbr>
      <budget_amount> 235980 </budget_amount>
      <quarter> FY2010Q4 </quarter>
    </row>
  - <row>
      <agency_name> United_States_Navy4 </agency_name>
      <abbr> UNITED STATES NAVY </abbr>
      <budget_amount> 3645.06 </budget_amount>
      <quarter> FY2010Q4 </quarter>
    </row>
  - <row>
      <agency_name> United_States_Navy4 </agency_name>
      <abbr> UNITED STATES NAVY </abbr>
      <budget_amount> 0 </budget_amount>
      <quarter> FY2010Q4 </quarter>
    </row>
```


 |

OBIEE Report

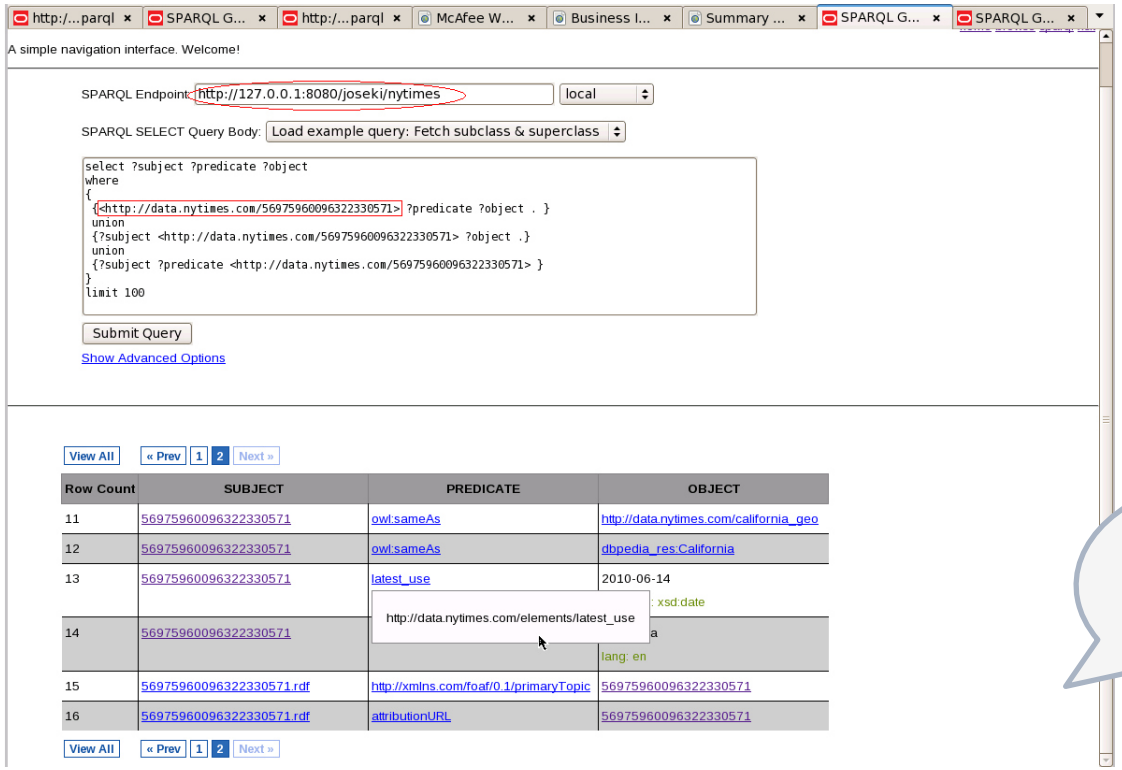


Demo:

Federated Queries across Oracle Database and Remote SPARQL Web Service endpoint

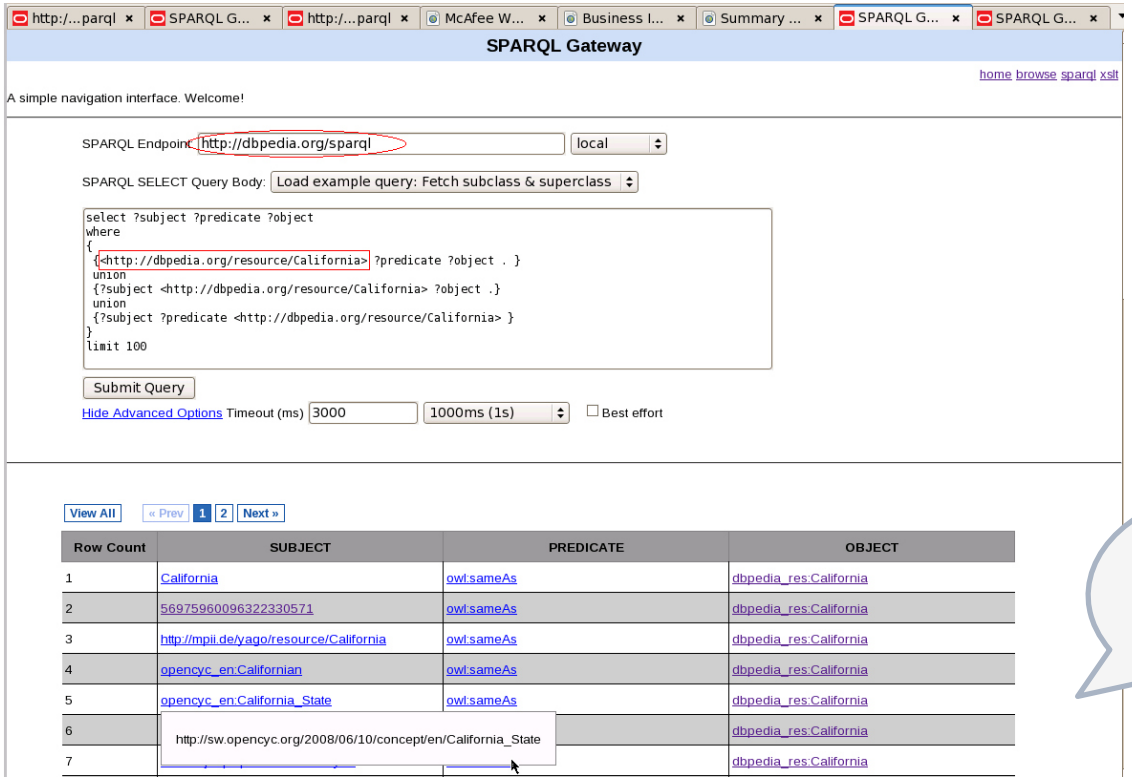
ORACLE

NYTimes RDF Data about CA



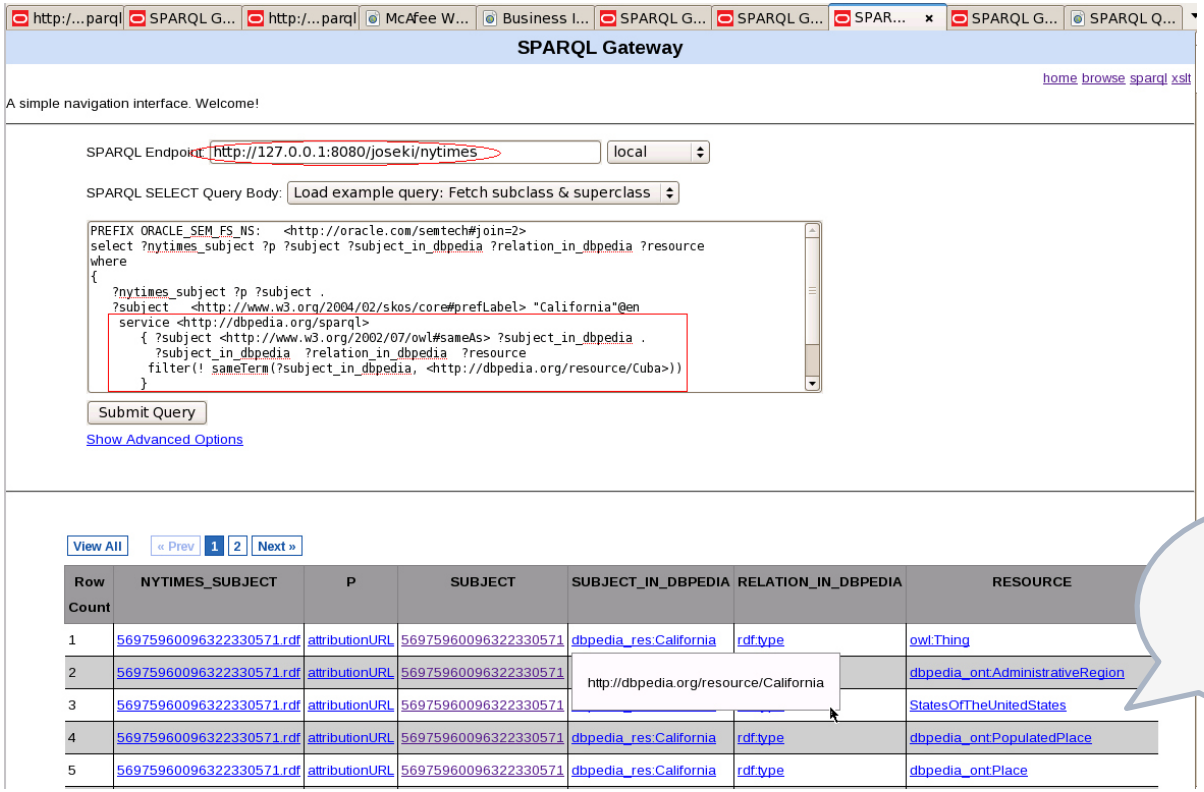
Data stored In Oracle Database

DBPedia RDF Data about CA



Data stored
On the Web

Federated Query



Data Merged together!