
ORACLE®

ORACLE

You have Terabytes Worth of Triples, Now What?
-- Mining Insights from Your Semantic Data Store

Xavier Lopez
Director, Product Management
Oracle Corporation

Zhe Wu
Consulting Member of Technical Staff
Oracle Corporation

Semantic Tech & Biz
Conference- 5 June 2012

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.

Outline

- Characteristics of a RDF Triple Store
- Challenge: Need for Powerful User-friendly Tools
- Performance and Scalability of Oracle RDF store
- Integration of Business Analytics, Data Mining, and R
- Summary

Basic Characteristics of a Triple Store

- Standards Compliance – W3C
  - RDF, RDFS, OWL 2, SKOS, SPARQL, …
- Fast loading of triple
  - Incremental and bulk loading
- Indexing of triples for fast access
  - Incrementally maintained
- Inferencing
  - Pre-computed inferences (forward chaining)
  - Run-time inferences (backward chaining)
- Querying
  - Allow multiple RDF graphs in SPARQL queries
  - Query execution planning for optimal performance

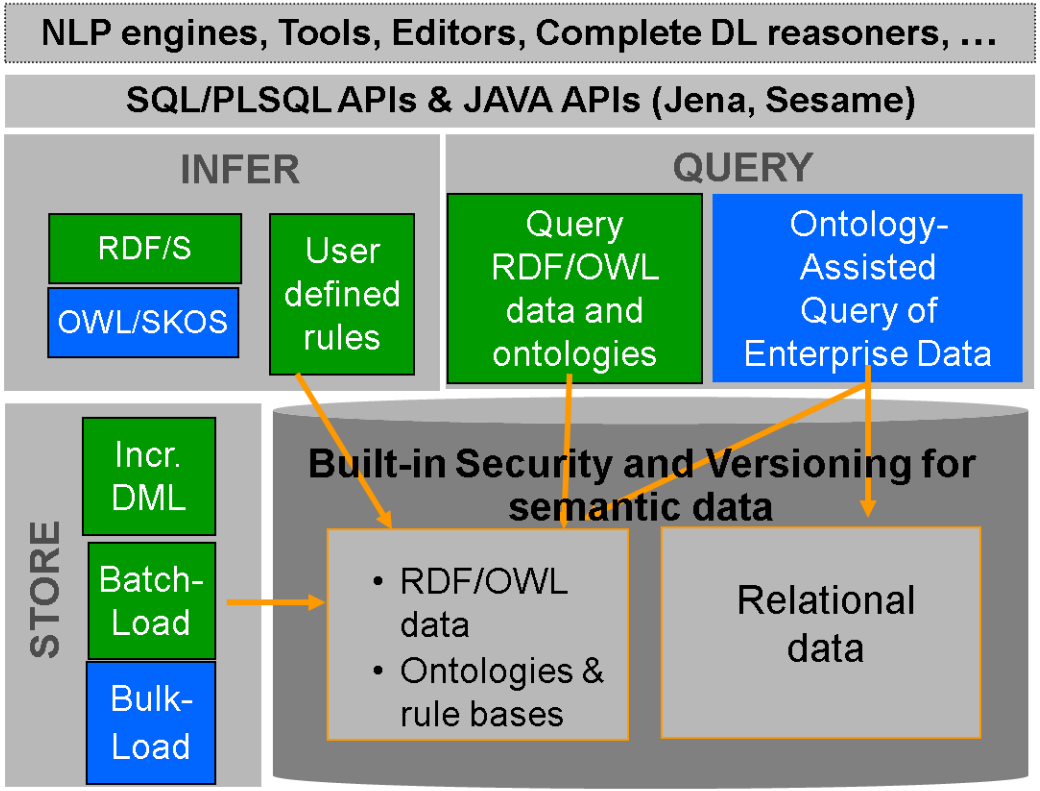
Enterprise Capabilities of a Triple Store

- Scalability to 100s of billions of triples and more
- Integrated access to Relational and RDF data
  - SPARQL query (embedding) in SQL
  - Join SPARQL results with ubiquitous relational data
  - Rich SQL operators (such as aggregates) on triples
- Semantic indexing
  - Index on a source document is an RDF graph with 1 named graph per doc
  - Triples are extracted from a document using NLP and entity/concept extraction
  - RDF graph incrementally updated as new documents entered
- Security: Fine-Grained Access Control (for each triple)
- Querying Text, Spatial and temporal data using SPARQL

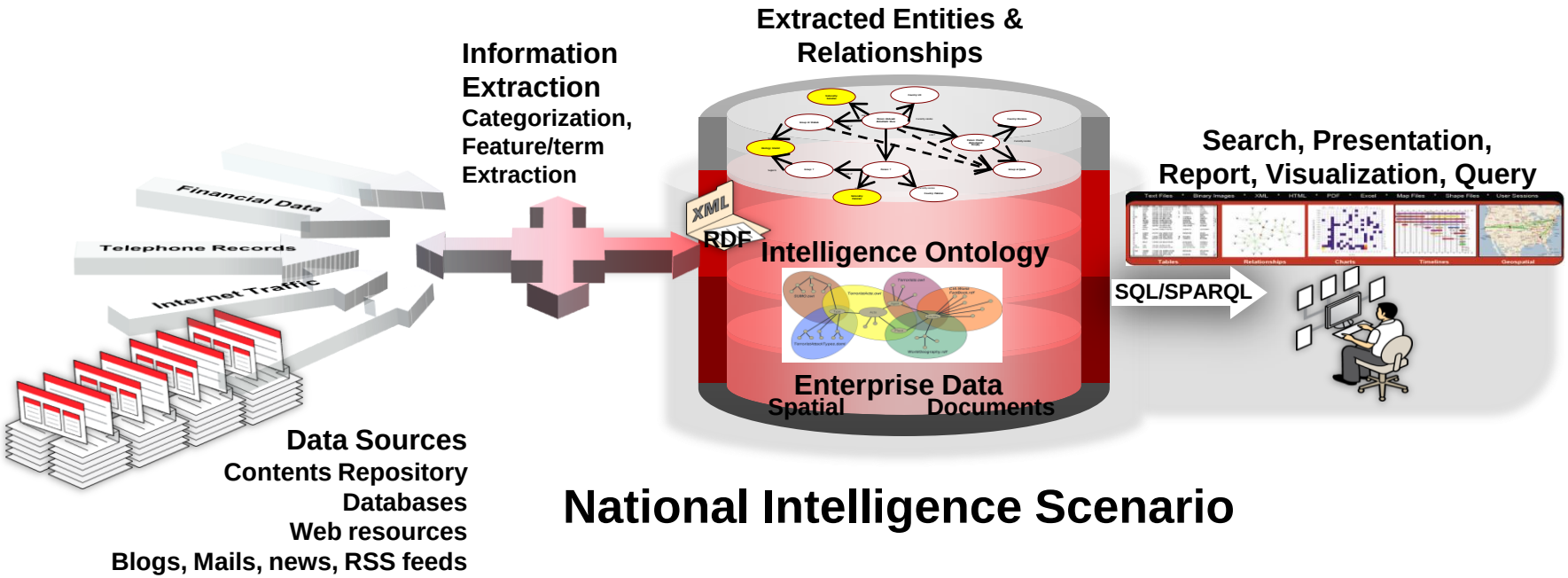
Other Enterprise Features of a Triple Store

- User-defined rules
- Better inferencing
  - Higher order logic beyond OWL 2
  - Incremental inferencing for higher availability
- Tools that need to work with RDF data
  - Navigation and visualization of RDF graphs
  - Graph creation and manipulation
  - Reporting and traditional charting of selected RDF data
  - Exploring & analyzing (testing conjectures)
  - Automated discovery (mining) & predictive analysis

Capabilities Overview of Oracle Database Release 11.2



Role of Semantic-aware Ontology in Intelligence Domain



National Intelligence Scenario

Software Tooling is Key to Adoption

- Need for tools
  - Navigation and visualization of RDF graphs
  - Graph creation and manipulation
  - Reporting and traditional charting of selected RDF data
  - Exploring & analyzing (testing conjectures)
  - Automated discovery (mining) & predictive analysis

Semantic Technologies Partners:

Ontology Engineering

TopQuadrant

protégé

ontoprise
know how to use Know-how

Reasoners

clark&parsia llc
thinking clearly

ontoprise
know how to use Know-how

NLP Entity Extractors

EXPERT SYSTEM
SEMANTIC INTELLIGENCE

LYMBA
the power to answer

GATE
general architecture
for text engineering

Linguamatics

SAILLABS
TECHNOLOGY

CALAIS
Powered by Thomson Reuters

Open Source Frameworks

semantic web framework
jena

openRDF.org
Sesame

Joseki

Standards

OGC
Technical Committee Member

RDF

W3C Semantic Web

Applications

PolarLake

TERANODE

MedTrust

MONDECA

SI / Consulting

infoMENTUM

Raytheon

NORTHROP GRUMMAN

McDONALD BRADLEY

accenture

Cognia

NKA-DECKER

TenForce
The Pragmatic Company

orbis
TECHNOLOGIES, INC.

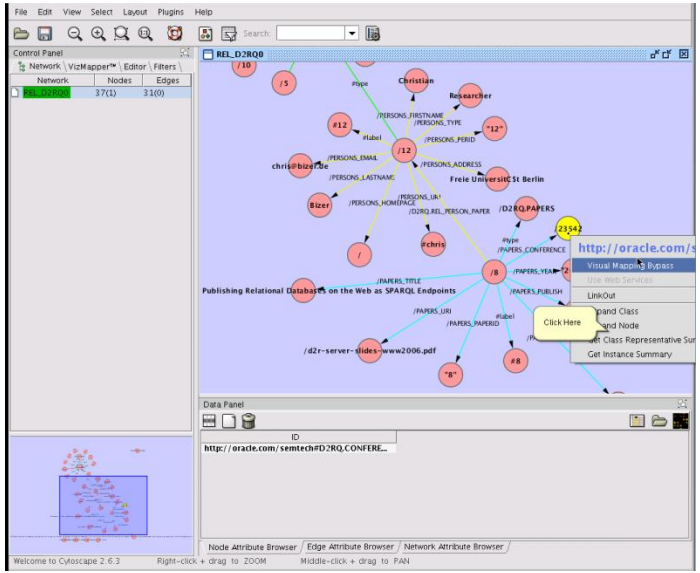
BOEING

ORACLE

Navigation and Visualization of RDF graphs

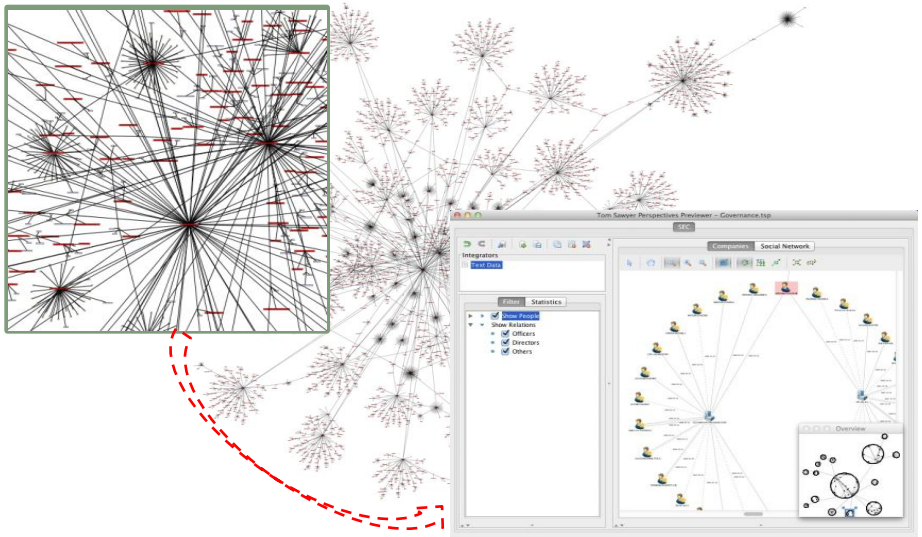
Open Source

- **Adapt** Cytoscape to work with very large RDF graphs



Commercial Products

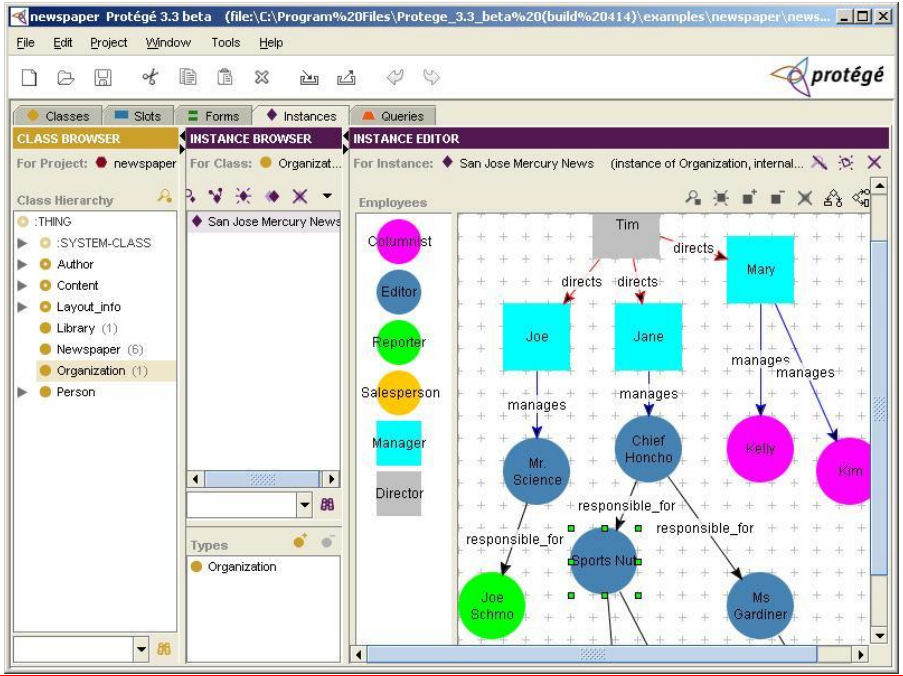
- Tom Sawyer's Perspective now supports RDF



Graph Creation and Manipulation

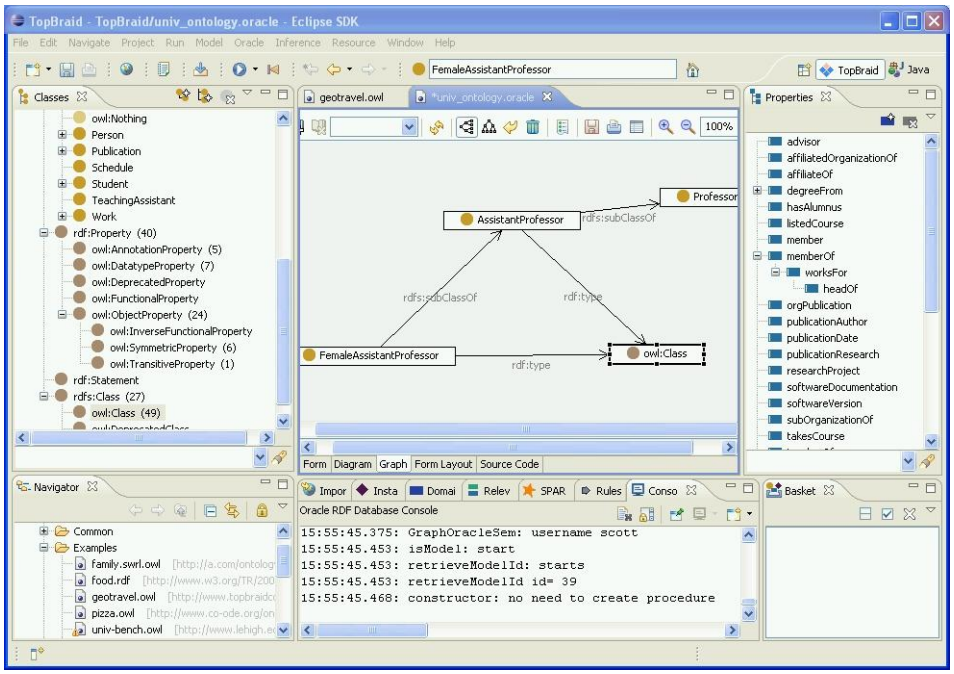
Open Source

- Protege ontology editor



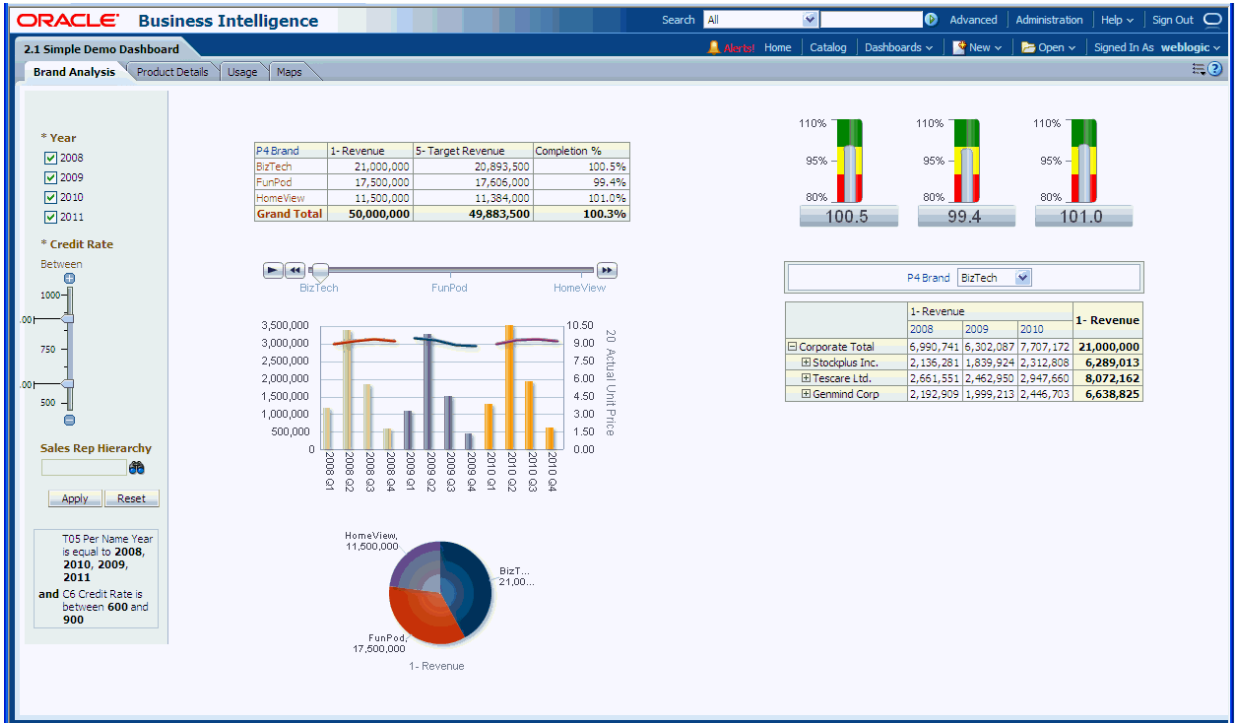
Commercial Products

- Top Quadrant Composer



Reporting and Charting Tools for RDF

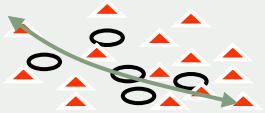
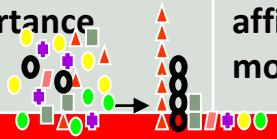
- **Native RDF tools unavailable**
  - for BI style reporting, charting and interactive refinement, eg. Oracle's BI Dashboard, available today primarily for Relational and XML Database



RDF Tools for Exploring & Analyzing

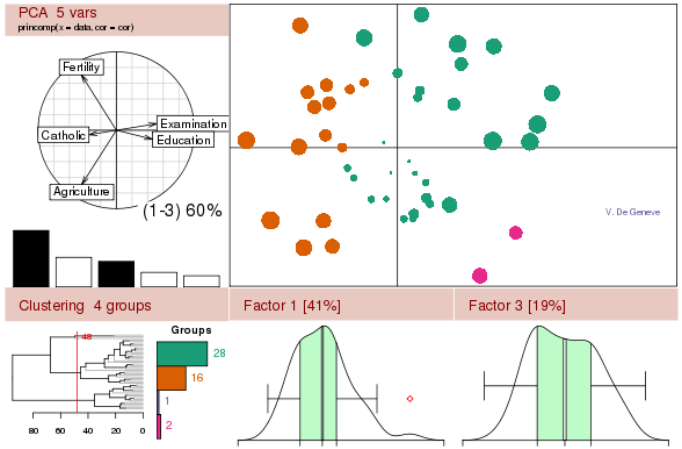
Tools for guided pattern discovery & statistical analysis

Supervised data mining, eg. Oracle Data Miner

Problem Classification	Sample Problem
Classification 	Given demographic data about a set of customers, predict customer response to an affinity card program
Regression 	Given demographic and purchasing data about a set of customers, predict customers' age
Attribute Importance 	Given customer response to an affinity card program, find the most significant predictors

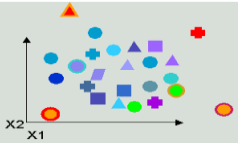
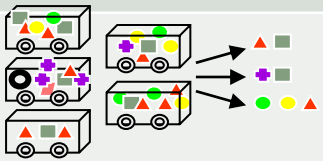
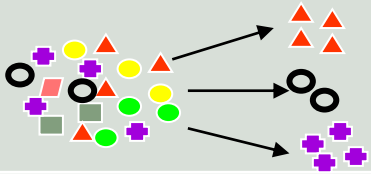
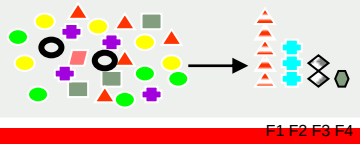
Statistical analysis, eg. Oracle R Enterprise

- Open source language and environment
- Statistical computing and graphics
- Easily produces publication-quality plots
- Highly extensible with open source R packages

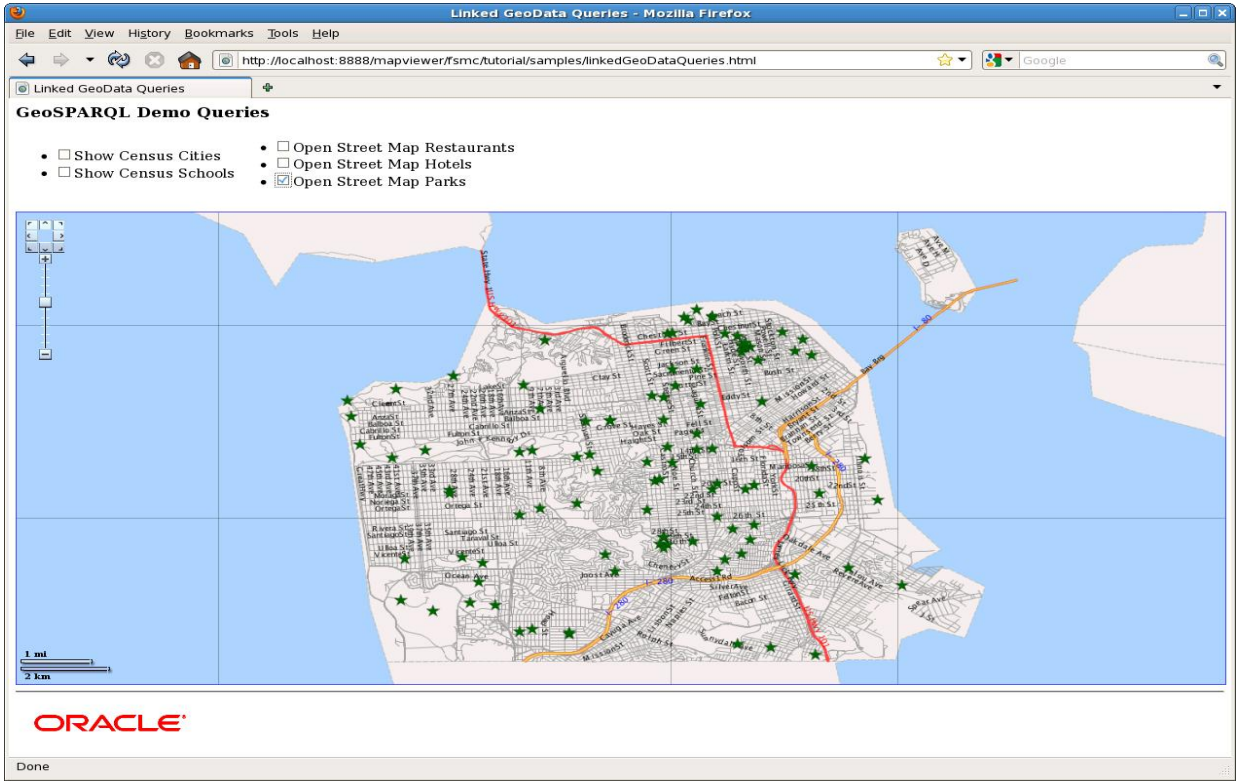


Automated Discovery & Predictive Analysis

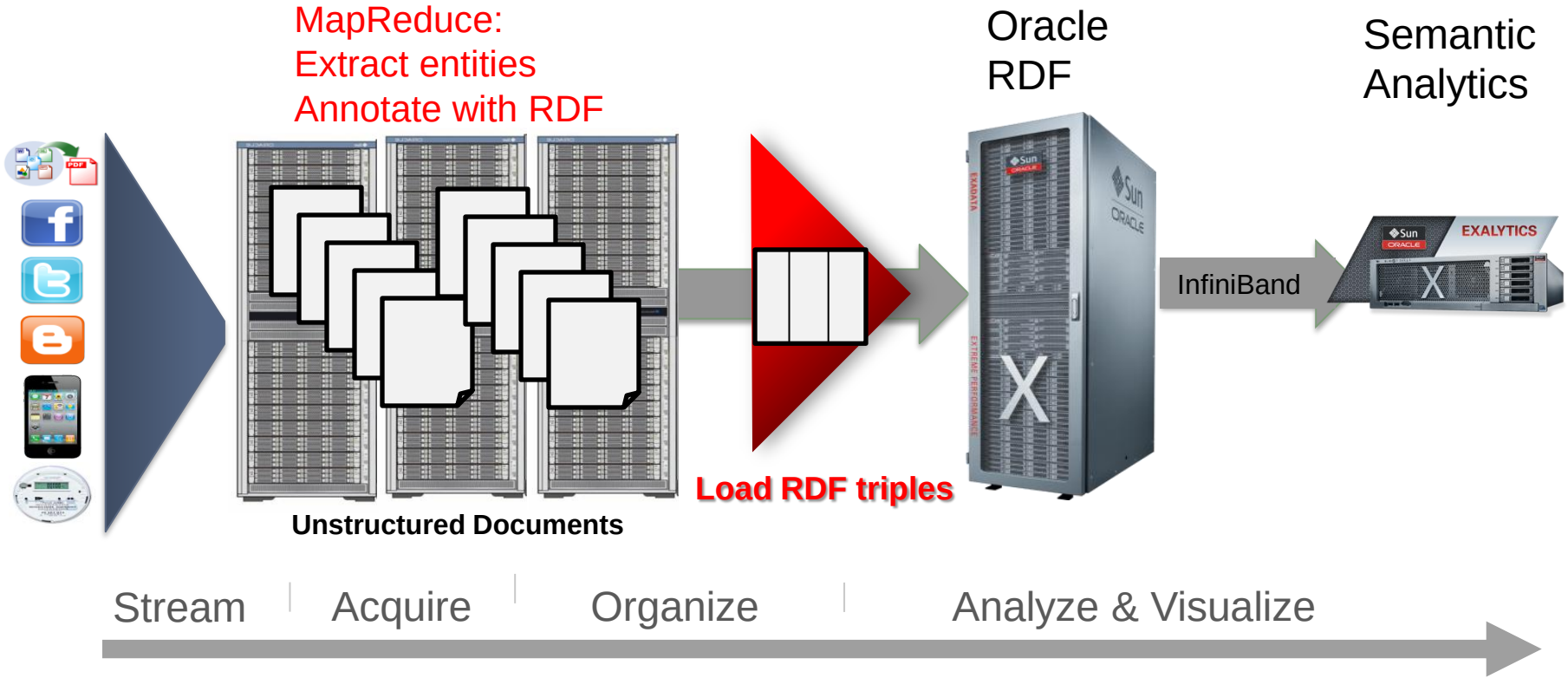
Oracle Data Miner

Problem Classification	Sample Problem
Anomaly Detection 	Given demographic data about a set of customers, identify customer purchasing behavior that is significantly different from the norm
Association Rules 	Find the items that tend to be purchased together and specify their relationship – market basket analysis
Clustering 	Segment demographic data into clusters and rank the probability that an individual will belong to a given cluster
Feature Extraction 	Given demographic data about a set of customers, group the attributes into general characteristics of the customers

Web Mapping with GeoSPARQL



Piping Big Data to RDF Analytics



How to Integrate RDF with Enterprise Reporting, Analysis and Discovery Tools

Zhe Wu

Consulting Member of Technical Staff

Oracle Corporation

Billions of Triples,Terabytes of Storage

Scalability is High, and Getting Better

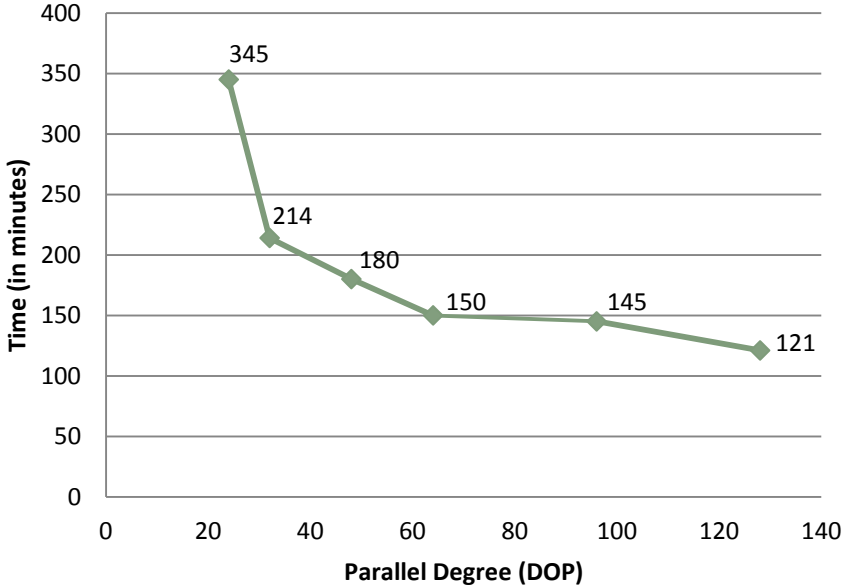
- LUBM 25K tested
  - 3.4 Billion triples in the model
  - 2.7 Billion triples inferred
  - Storage space including indexes: over 1 Terabytes

  - Load speed: 109 minutes on Sun M8000
  - Inference speed: 160 minutes
  - Query throughput: 0.5 Billion in 9 minutes

  - **Balanced hardware**, high parallelism for responsiveness
    - Sun M8000, 512GB RAM, 2TB Flash array, 128 Threads

Parallel Execution Performance on M8000

- LUBM 25K local inference on Sun M8000
  - 6.1B+ quads (3.4B asserted, 2.7B inferred)



Oracle's Parallel Execution is completely transparent!

- **Cross CPUs/Cores on a single node**
- **Cross multiple nodes in a cluster**

Inference Performance on Exadata V2

Data Set (# triples)	Triples Inferred	Time	Degrees of Parallelism
LUBM 100K (13B)	5B	1h, 58' [1]	DOP = 32
LUBM 25K (3.3B)	2.7B	4h, 7' [2]	DOP = 32
LUBM 8K (1.1B)	869M	46' [2]	DOP = 64

[1] **Preliminary result:** 1 round of OWLPrime (OWL Horst semantics)

[2] **Inference:** OWLPrime + components: INTERSECT,INTERSECTSCOH,SVFH,THINGH,THINGSAM,UNION

Setup:

Hardware: Full Rack Sun Oracle Database Machine X2-2 (8 nodes, 72GB RAM per node), and Exadata Storage Server

Storage required: LUBM8K: 330GB or LUBM25K 1TB + 110GB temp table space

Software Oracle Database 11.2.0.1.0 + Patch 9819833: SEMANTIC TECHNOLOGIES 11G R2 FIX BUNDLE 2

Each node: SGA_TARGET=32G and PGA_AGGREGATE_TARGET=31G

Query Performance on Exadata V2

Auto DOP used. **465,849,803 answers generated for LUBM 25K in 274.2 sec.**

Ontology LUBM 25K 3.3 billion triples & 2.7 billion inferred		LUBM Benchmark Queries						
OWLPrime & new inference components	Query	Q1	Q2	Q3	Q4	Q5	Q6	Q7
	# answers	4	2528	6	34	719	260M	67
	Complete?	Y	Y	Y	Y	Y	Y	Y
	Time (sec)	0.01	20.65	0.01	0.01	0.02	23.07	4.99
	Query	Q8	Q9	Q10	Q11	Q12	Q13	Q14
	# answers	7790	6.8M	4	224	15	0.11M	197M
	Complete?	Y	Y	Y	Y	Y	Y	Y
	Time (sec)	0.48	203.06	0.01	0.02	0.02	2.40	19.45

A Strategy to Introduce Analytic Tools

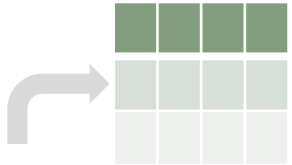
- Great need for RDF analytic tools
  - But the field of analytic tools for RDF databases is a barren one
- In contrast, the mature field of relational & XML analytics is abundant
  - It is still a major undertaking to add native RDF/Sparql support to relational/XML analytic and mining tools
- Our strategy
  - Provide a simple way to make target RDF data available in XML or relational

Use BI Tool against Semantic Data

- Make the semantic data available to a BI tool in an appropriate format
  - Turn a semantic data store into yet another data source for BI tool
    - Logical inference can be used to “enrich” asserted facts
  - Relational and XML are popular formats
- Static versus dynamic data transformation
  - Static data transformation is acceptable if data seldom updates
  - Dynamic data transformation is crucial if data frequently updates
    - Eliminates synchronization
    - Maintains a single source of truth, better security

Use BI Tool: Semantic Data to Relational

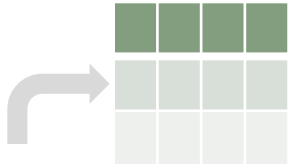
- Data transformation: semantic data ➔ relational
  - Bindings from a SPARQL query can naturally be viewed as "columns"



SPARQL ➔

```
select ?agency_name ?label ?budget ?quarter
 where {
         ?agency    :hasQuarter    ?quarter .
         ?agency    rdfs:label     ?label .
         ?agency  :hasBudgetAmount ?budget .
         ?agency    :hasAbbrev     ?agency_name
       }
```

Use BI Tool: Semantic Data to Relational



- Data transformation: semantic data ➔ relational
  - **create view** AGENCY_BUDGET as select * from

SQL ➔

```
(table(sem_match('
    select ?agency_name ?label ?budget ?quarter
     where {
            ?agency    :hasQuarter    ?quarter .
            ?agency    rdfs:label     ?label .
            ?agency  :hasBudgetAmount ?budget .
            ?agency    :hasAbbrev     ?agency_name
          }'  … )))
```

- Using materialized view is possible
- Add post-processing (e.g. un-escaping) of column values if necessary

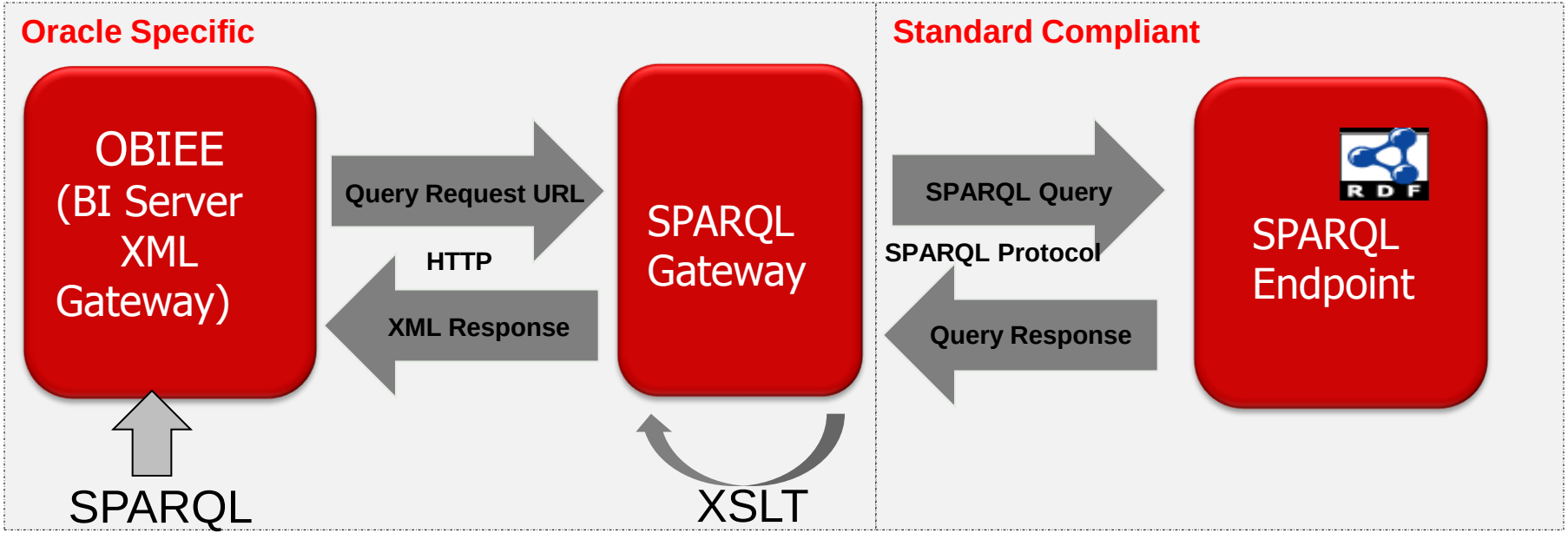
Using BI Tool: Semantic Data to XML

- **XML is the default format of SPARQL Query response from a web service endpoint. Transformation is necessary**
  - BI Tool may require a different XML format
  - May need to remove namespaces, data type URIs, etc.

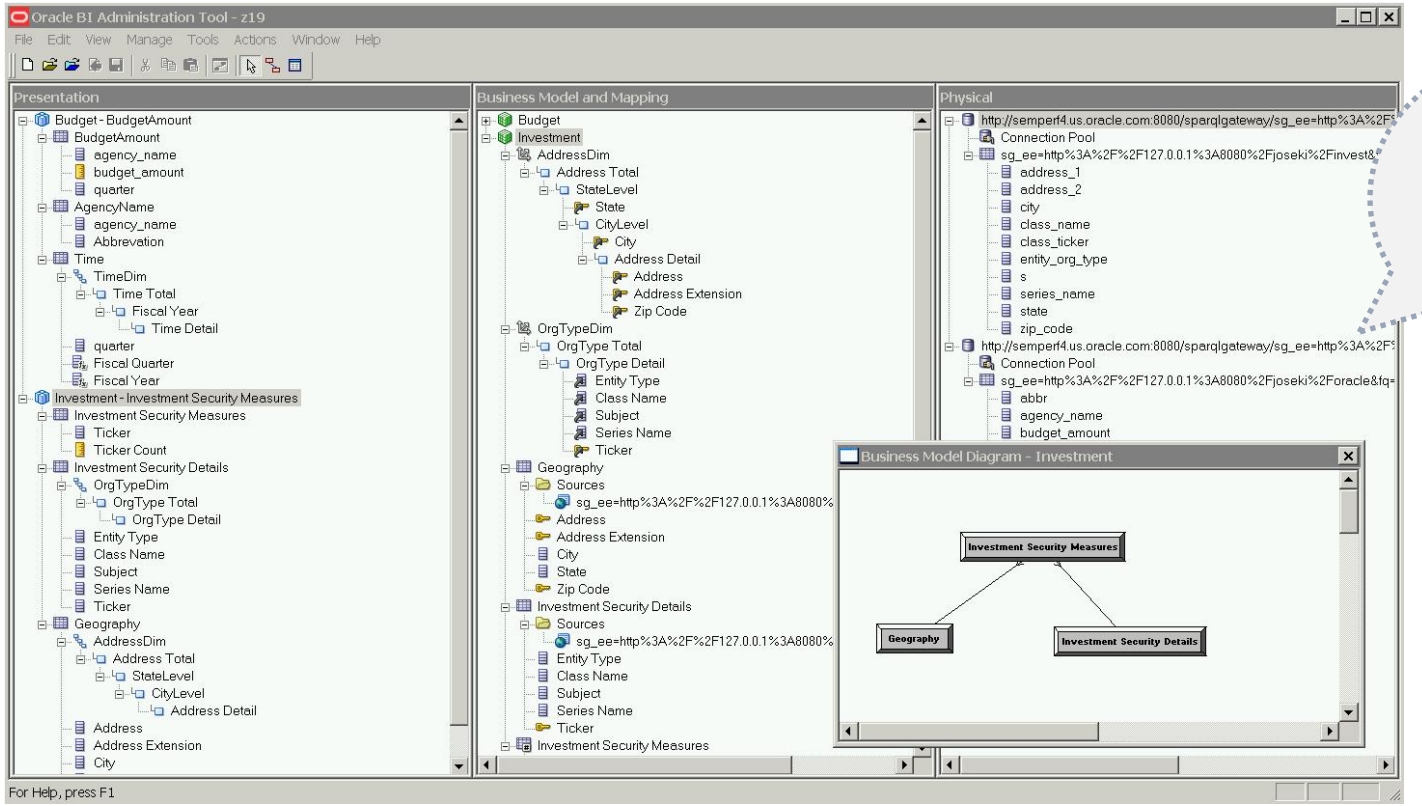
SPARQL Query Response XML	OBIEE expected XML
<sparql xmlns="http://www.w3.org/2005/sparql-results#"> <head> <variable name="agency_name"/> … <results> <result> <binding name="agency_name"> <literal>DEF. ADV. RESEARCH PROJ.</literal> </binding> …	<test> <row> <agency_name> DEF. ADV. RESEARCH PROJ. </agency_name> <budget> … </budget> </row> …

Semantic Data to XML via SPARQL Gateway

- **SPARQL Gateway is a feature of Jena Adapter**



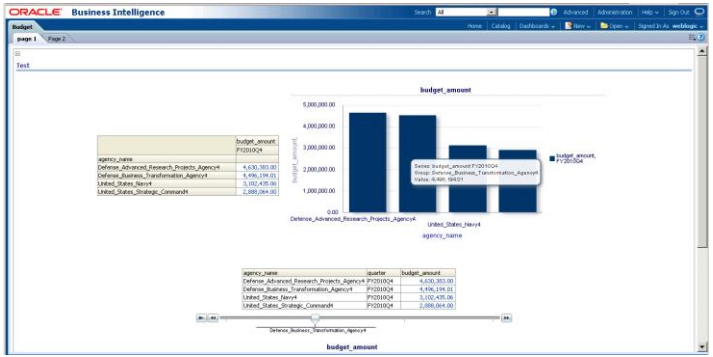
Using BI tool: Create Business Model

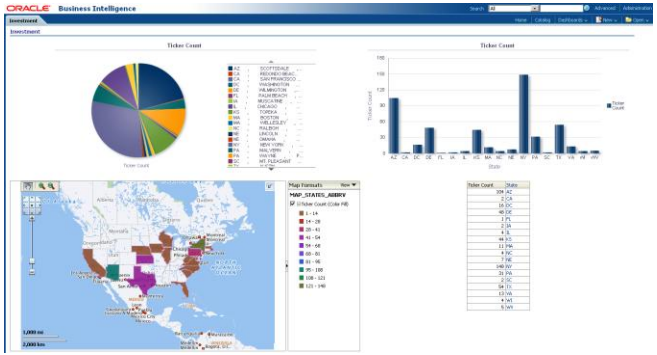


Relational, XML, … or RDF data sources

Use BI Tool against Semantic Data

- Tie it all together
  - Turn a semantic data store into yet another data source to BI
  - Perform conventional BI modeling
  - Define presentation layer, build report/dashboard
  - This is one example of what you may get:





Using Data Mining Tool against Semantic Data

- Make the semantic data available to a data mining tool in an appropriate format
  - Turn a semantic data store into yet another data source for DM tool

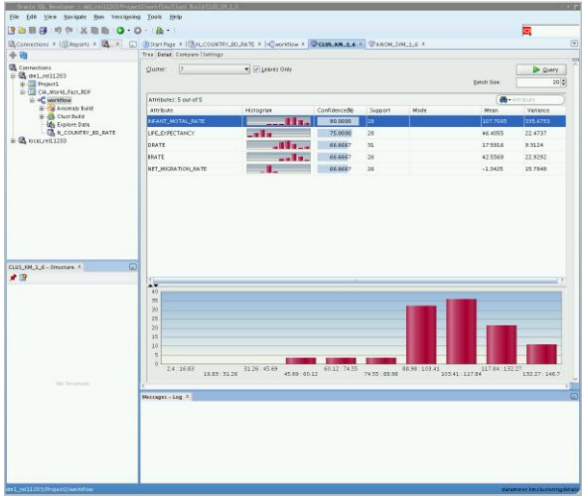
```
create view N_COUNTRY_BD_RATE as select name
  , to_number(brate) as  brate
  , to_number(drate) as  drate
  , to_number(popu)  as  population
  , to_number(mig)    as  net_migration_rate
  , to_number(imr)    as   infant_motal_rate
  , to_number(leab)  as   life_expectancy
 from table(sem_match('{
  ?subject <http://www.cia.gov/cia/publications/factbook#Birth_rate>                    ?brate .
  ?subject <http://www.cia.gov/cia/publications/factbook#Death_rate>                    ?drate .
  ?subject <http://www.cia.gov/cia/publications/factbook#Name>                          ?name .
  ?subject <http://www.cia.gov/cia/publications/factbook#Population>                    ?popu .
  ?subject <http://www.cia.gov/cia/publications/factbook#Net_migration_rate>          ?mig .
  ?subject <http://www.cia.gov/cia/publications/factbook#Infant_mortality_rate>       ?imr .
  ?subject <http://www.cia.gov/cia/publications/factbook#Life_expectancy_at_birth> ?leab .
 }‘ … ))
```

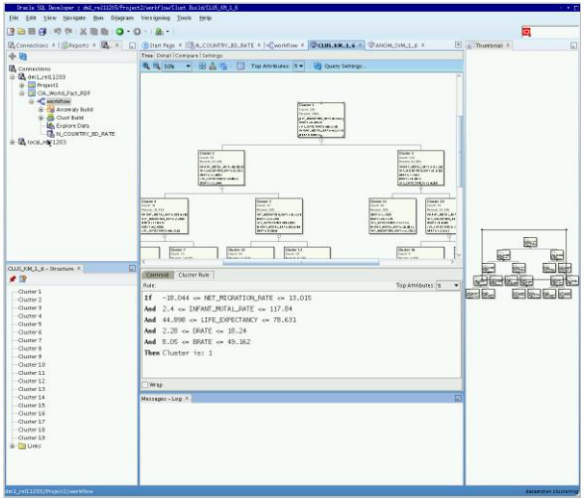
SPARQL

SQL

Using Data Mining Tool against Semantic Data

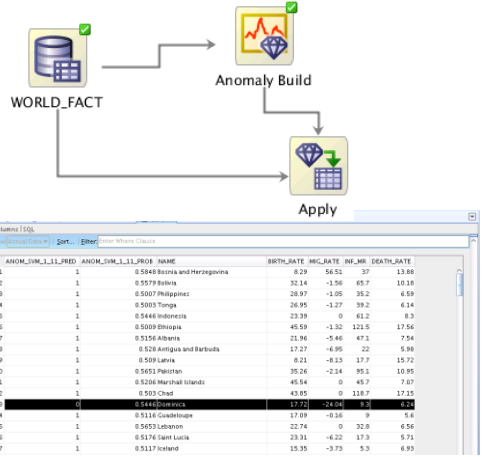
- Tie it all together
  - Turn a semantic data store into yet another data source to DM
  - Follow the conventional DM process:
    - Data preparation, build/evaluate model, deployment
    - This is one example of what you may get:

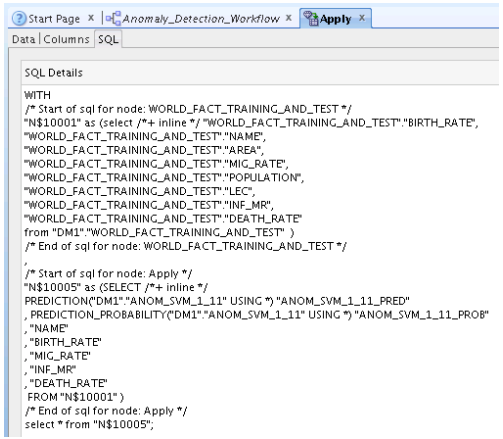




Using Data Mining Tool against Semantic Data

- Tie it all together
  - Turn a semantic data store into yet another data source to DM
  - Follow the conventional DM process:
    - Data preparation, build/evaluate model, deployment
    - Some Mining results can be **saved back as RDF** into Oracle database





Anomaly Detection output in SQL

Convert into RDF

```
:AbnormalCase1   :hasSubject
                          :Dominica   .
:AbnormalCase1   :probability
                          "0.54"
```

Using Oracle R Enterprise with Semantic Data

- Make the semantic data available to ORE in relational format
  - Turn a semantic data store into yet another data source for R tool

```
create view  COUNTRY_INFO  as  select  name
  , to_number(brate) as  brate
  , to_number(drate) as  drate
  , to_number(popu)  as  population
  , to_number(mig)    as  net_migration_rate
  , to_number(imr)    as   infant_motal_rate
  , to_number(leab)  as   life_expectancy
 from table(sem_match('{
  ?subject <http://www.cia.gov/cia/publications/factbook#Birth_rate>                  ?brate .
  ?subject <http://www.cia.gov/cia/publications/factbook#Death_rate>                  ?drate .
  ?subject <http://www.cia.gov/cia/publications/factbook#Name>                        ?name .
  ?subject <http://www.cia.gov/cia/publications/factbook#Population>                  ?popu .
  ?subject <http://www.cia.gov/cia/publications/factbook#Net_migration_rate>          ?mig .
  ?subject <http://www.cia.gov/cia/publications/factbook#Infant_mortality_rate>       ?imr .
  ?subject <http://www.cia.gov/cia/publications/factbook#Life_expectancy_at_birth> ?leab .
 }‘ … ))
```

SPARQL

SQL

Using Oracle R Enterprise with Semantic Data

- Tie it all together
  - Turn a semantic data store into yet another data source to ORE
  - Use R to analyze semantic data
    - This is one example of what you may get:

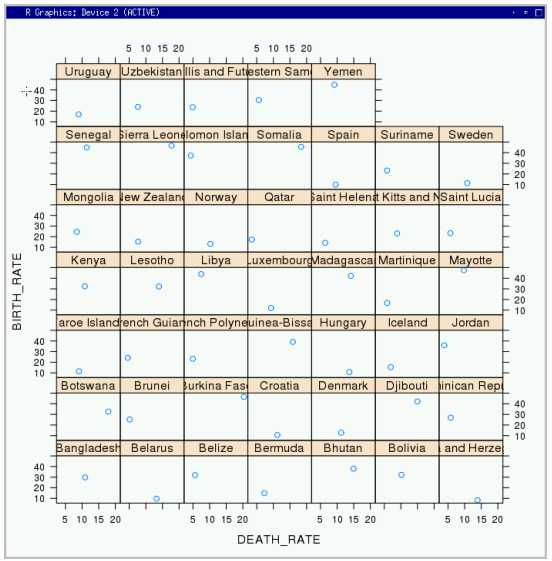
```
Oracle Enterprise R

Attaching package: 'OREbase'

The following object(s) are masked from 'package:base':

    cbind, data.frame, eval, interaction, order, paste, pmax, pmin,
    rbind, table

Loading required package: OREstats
Loading required package: MASS
Loading required package: OREgraphics
Loading required package: OREeda
Loading required package: ORExml
> ore.connect("rquser", "rel11203", "localhost", "rquser")
> ore.sync()
Warning message:
"RQUSER"."R_TPL" contains unsupported data types
> ore.attach()
> library(lattice)
>
> COUNTRY_LOCAL <- ore.pull(COUNTRY_BD_RATE)
> xyplot(BIRTH_RATE ~ DEATH_RATE | NAME, data = COUNTRY_LOCAL)
> head(COUNTRY_INFO)
                NAME BIRTH_RATE DEATH_RATE LIFE_EXP INF_MR MIG_RATE
1 Bosnia and Herzegovina       8.29      13.88    59.42   37.0    56.51
2                Bolivia      32.14      10.18    60.34   65.7    -1.56
3    Antigua and Barbuda      17.27       5.98    70.93   22.0    -6.95
4            Saint Lucia      23.31       5.71    71.36   17.3    -6.22
5                Iceland      15.35       6.93    78.73    5.3    -3.73
6               Cameroon      42.22      13.64    52.27   77.6     0.00
  POPULATION    AREA
1    3222584   51233
2    7669868 1098580
3      63739     440
4     150630     620
5     269697  103000
6   14677510  475440
>
> []
```



Lattice graph

Using Oracle R Enterprise with Semantic Data

- Tie it all together
  - Turn a semantic data store into yet another data source to ORE
  - Use R to analyze semantic data
    - This is one example of what you may get:

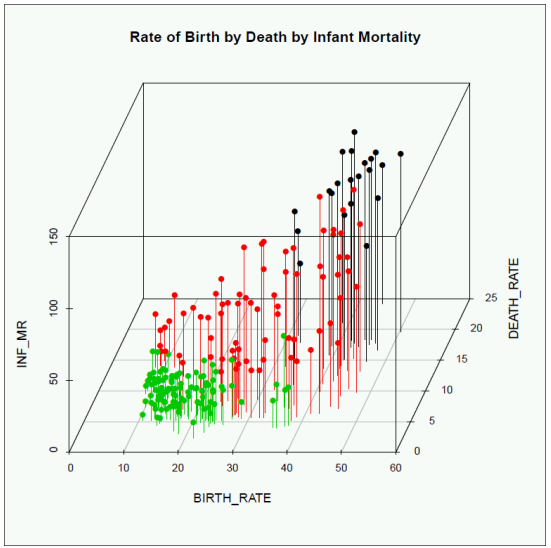
```
Oracle Enterprise R

Attaching package: 'OREbase'

The following object(s) are masked from 'package:base':

    cbind, data.frame, eval, interaction, order, paste, pmax, pmin,
    rbind, table

Loading required package: OREstats
Loading required package: MASS
Loading required package: OREgraphics
Loading required package: OREeda
Loading required package: ORExml
> ore.connect("rquser", "rel11203", "localhost", "rquser")
> ore.sync()
Warning message:
"RQUSER"."R_TPL" contains unsupported data types
> ore.attach()
> library(lattice)
>
> COUNTRY_LOCAL <- ore.pull(COUNTRY_BD_RATE)
> xyplot(BIRTH_RATE ~ DEATH_RATE | NAME, data = COUNTRY_LOCAL)
> head(COUNTRY_INFO)
                  NAME BIRTH_RATE DEATH_RATE LIFE_EXP INF_MR MIG_RATE
1 Bosnia and Herzegovina       8.29      13.88    59.42   37.0    56.51
2              Bolivia      32.14      10.18    60.34   65.7    -1.56
3    Antigua and Barbuda      17.27       5.98    70.93   22.0    -6.95
4            Saint Lucia      23.31       5.71    71.36   17.3    -6.22
5                Iceland      15.35       6.93    78.73    5.3    -3.73
6               Cameroon      42.22      13.64    52.27   77.6     0.00
  POPULATION    AREA
1    3222584   51233
2    7669868 1098580
3      63739     440
4     150630     620
5     269697  103000
6   14677510  475440
>
> []
```



Scatterplot3d

Using Oracle R Enterprise with Semantic Data

- Tie it all together
  - Turn a semantic data store into yet another data source to ORE
  - Use R to analyze semantic data
    - This is one example of what you may get:

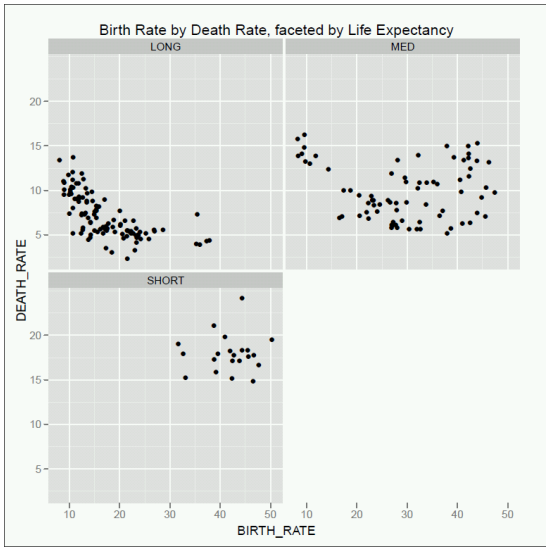
```
Oracle Enterprise R

Attaching package: 'OREbase'

The following object(s) are masked from 'package:base':

    cbind, data.frame, eval, interaction, order, paste, pmax, pmin,
    rbind, table

Loading required package: OREstats
Loading required package: MASS
Loading required package: OREgraphics
Loading required package: OREeda
Loading required package: ORExml
> ore.connect("rquser", "rel11203", "localhost", "rquser")
> ore.sync()
Warning message:
"RQUSER"."R_TPL" contains unsupported data types
> ore.attach()
> library(lattice)
>
> COUNTRY_LOCAL <- ore.pull(COUNTRY_BD_RATE)
> xyplot(BIRTH_RATE ~ DEATH_RATE | NAME, data = COUNTRY_LOCAL)
> head(COUNTRY_INFO)
                NAME BIRTH_RATE DEATH_RATE LIFE_EXP INF_MR MIG_RATE
1 Bosnia and Herzegovina       8.29      13.88    59.42   37.0    56.51
2                Bolivia      32.14      10.18    60.34   65.7    -1.56
3    Antigua and Barbuda      17.27       5.98    70.93   22.0    -6.95
4            Saint Lucia      23.31       5.71    71.36   17.3    -6.22
5                Iceland      15.35       6.93    78.73    5.3    -3.73
6               Cameroon      42.22      13.64    52.27   77.6     0.00
  POPULATION    AREA
1    3222584   51233
2    7669868 1098580
3      63739     440
4     150630     620
5     269697  103000
6   14677510  475440
>
> []
```



qplot

Using Oracle R Enterprise with Semantic Data

- Tie it all together
  - Turn a semantic data store into yet another data source to ORE
  - Use R to analyze semantic data
    - This is one example of what you may get:

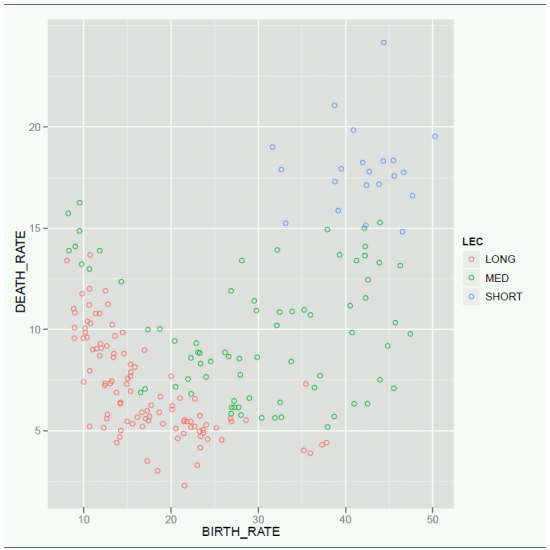
```
Oracle Enterprise R

Attaching package: 'OREbase'

The following object(s) are masked from 'package:base':

    cbind, data.frame, eval, interaction, order, paste, pmax, pmin,
    rbind, table

Loading required package: OREstats
Loading required package: MASS
Loading required package: OREgraphics
Loading required package: OREeda
Loading required package: ORExml
> ore.connect("rquser", "rel11203", "localhost", "rquser")
> ore.sync()
Warning message:
"RQUSER"."R_TPL" contains unsupported data types
> ore.attach()
> library(lattice)
>
> COUNTRY_LOCAL <- ore.pull(COUNTRY_BD_RATE)
> xyplot(BIRTH_RATE ~ DEATH_RATE | NAME, data = COUNTRY_LOCAL)
> head(COUNTRY_INFO)
                  NAME BIRTH_RATE DEATH_RATE LIFE_EXP INF_MR MIG_RATE
1 Bosnia and Herzegovina       8.29      13.88    59.42   37.0    56.51
2              Bolivia      32.14      10.18    60.34   65.7    -1.56
3    Antigua and Barbuda      17.27       5.98    70.93   22.0    -6.95
4            Saint Lucia      23.31       5.71    71.36   17.3    -6.22
5                Iceland      15.35       6.93    78.73    5.3    -3.73
6               Cameroon      42.22      13.64    52.27   77.6     0.00
  POPULATION    AREA
1    3222584   51233
2    7669868 1098580
3      63739     440
4     150630     620
5     269697  103000
6   14677510  475440
>
> []
```



ggplot

Using Oracle R Enterprise with Semantic Data

- Tie it all together
  - Turn a semantic data store into yet another data source to ORE
  - Use R to analyze semantic data
    - This is one example of what you may get:

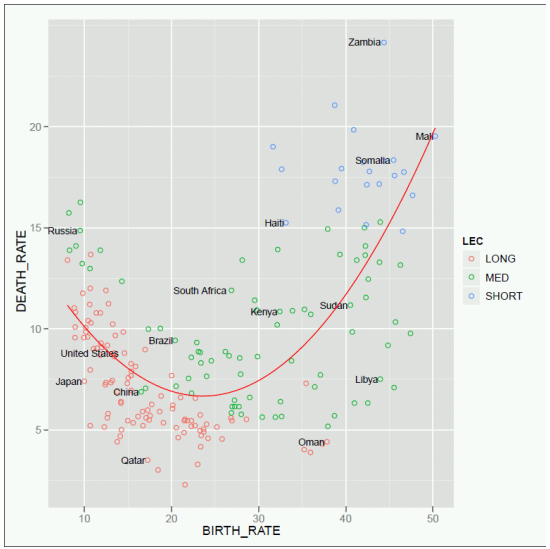
```
Oracle Enterprise R

Attaching package: 'OREbase'

The following object(s) are masked from 'package:base':

    cbind, data.frame, eval, interaction, order, paste, pmax, pmin,
    rbind, table

Loading required package: OREstats
Loading required package: MASS
Loading required package: OREgraphics
Loading required package: OREeda
Loading required package: ORExml
> ore.connect("rquser", "rel11203", "localhost", "rquser")
> ore.sync()
Warning message:
"RQUSER"."R_TPL" contains unsupported data types
> ore.attach()
> library(lattice)
>
> COUNTRY_LOCAL <- ore.pull(COUNTRY_BD_RATE)
> xyplot(BIRTH_RATE ~ DEATH_RATE | NAME, data = COUNTRY_LOCAL)
> head(COUNTRY_INFO)
                  NAME BIRTH_RATE DEATH_RATE LIFE_EXP INF_MR MIG_RATE
1 Bosnia and Herzegovina       8.29      13.88    59.42   37.0     56.51
2              Bolivia      32.14      10.18    60.34   65.7     -1.56
3    Antigua and Barbuda      17.27       5.98    70.93   22.0     -6.95
4            Saint Lucia      23.31       5.71    71.36   17.3     -6.22
5                Iceland      15.35       6.93    78.73    5.3     -3.73
6               Cameroon      42.22      13.64    52.27   77.6      0.00
  POPULATION    AREA
1    3222584   51233
2    7669868 1098580
3      63739     440
4     150630     620
5     269697  103000
6   14677510  475440
>
> []
```



geom_smooth

Summary

- Oracle delivers enterprise-class semantic data management
  - Manage RDF data w/ Oracle's scalability, performance, availability and triple-level label security
  - Incorporate popular open source technologies:
    - Jena, Sesame, Pellet, GATE, Cytoscape, Protégé
  - Exploit W3C and OGC standards
    - Query using SPARQL 1.1, SQL, GeoSPARQL
    - Inference in the database w/ OWL2 RL, EL+, SKOS rules
    - Develop in Java, PL/SQL
  - Semantically index documents and unstructured text
  - Use leading commercial tools: TopQuadrant, Tom Sawyer, Lymba…
- Oracle provides enterprise-class tools to help you mine insight from semantic data
  - OBIEE
  - Oracle Data Mining
  - Oracle R Enterprise

For More Information

Oracle RDF 

Xavier.Lopez@oracle.com

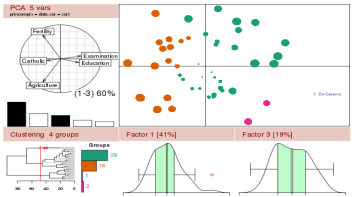
Alan.Wu@oracle.com

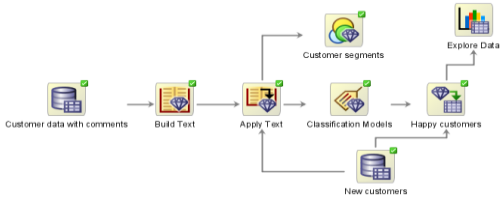
oracle.com


ORACLE®

Oracle In-Database Advanced Analytics

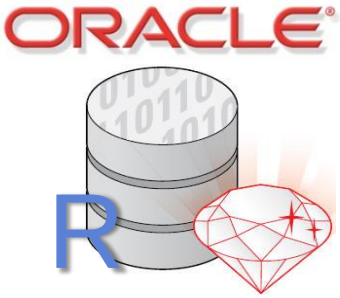
Comprehensive Advanced Analytics Platform





Oracle R Enterprise

- Popular open source statistical programming language & environment
- Integrated with database for scalability
- Wide range of statistical and advanced analytical functions
- R embedded in enterprise appls & OBIEE
- Exploratory data analysis
- Extensive graphics
- Open source R (CRAN) packages
- Integrated with Hadoop for HPC



Oracle Data Mining

- Automated knowledge discovery inside the Database
- 12 in-database data mining algorithms
- Text mining
- Predictive analytics applications development environment
- Star schema and transactional data mining
- Exadata "scoring" of ODM models
- SQL Developer/Oracle Data Miner GUI

Statistics Advanced Analytics Data & Text Mining Predictive Analytics