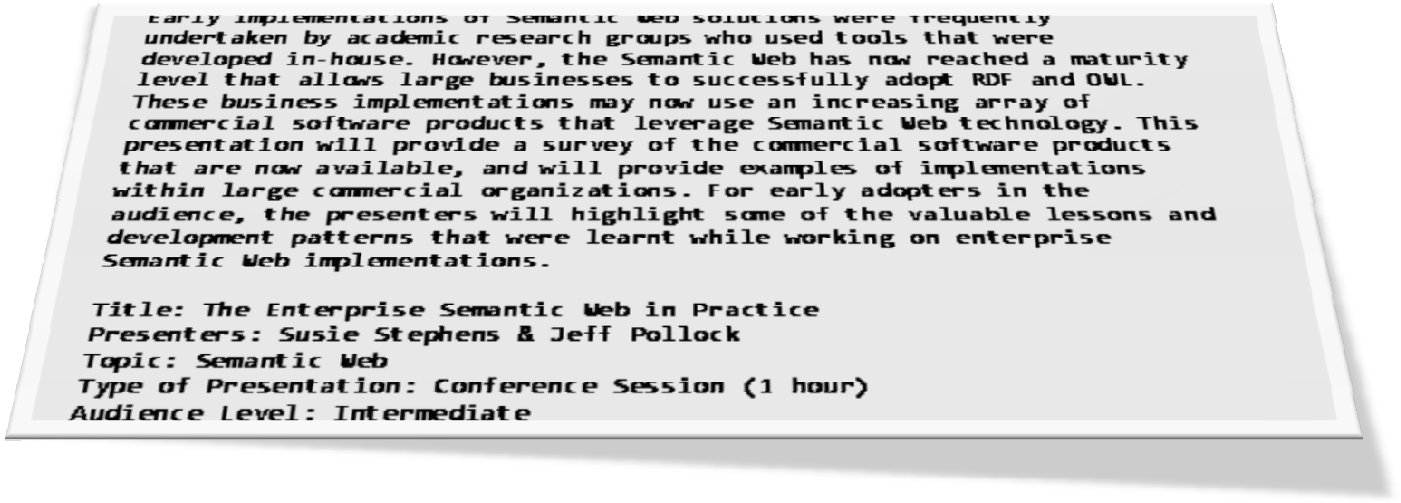

ORACLE®

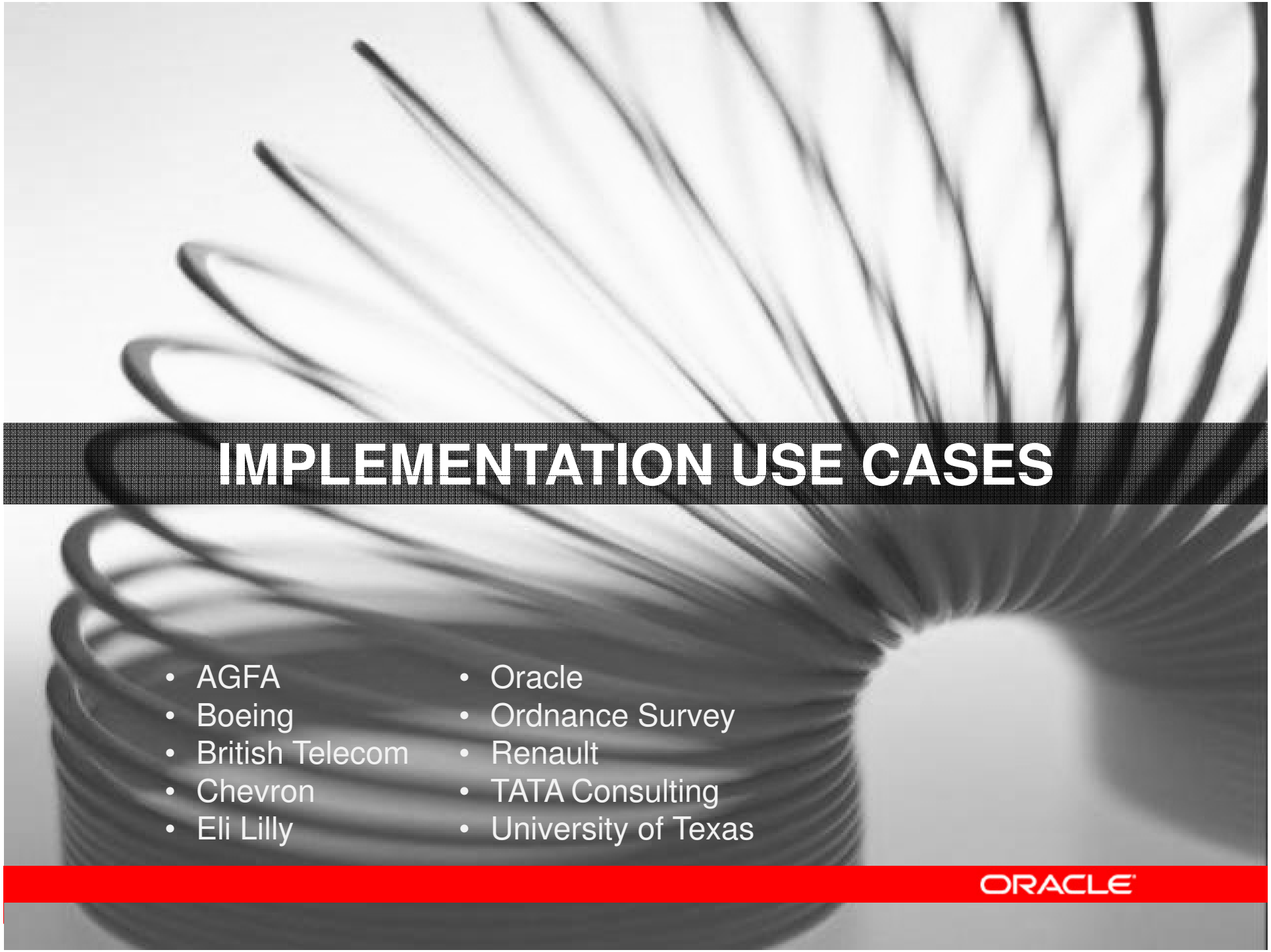
Enterprise Semantic Web in Practice

Jeffrey T. Pollock
Senior Director, Oracle Fusion Middleware

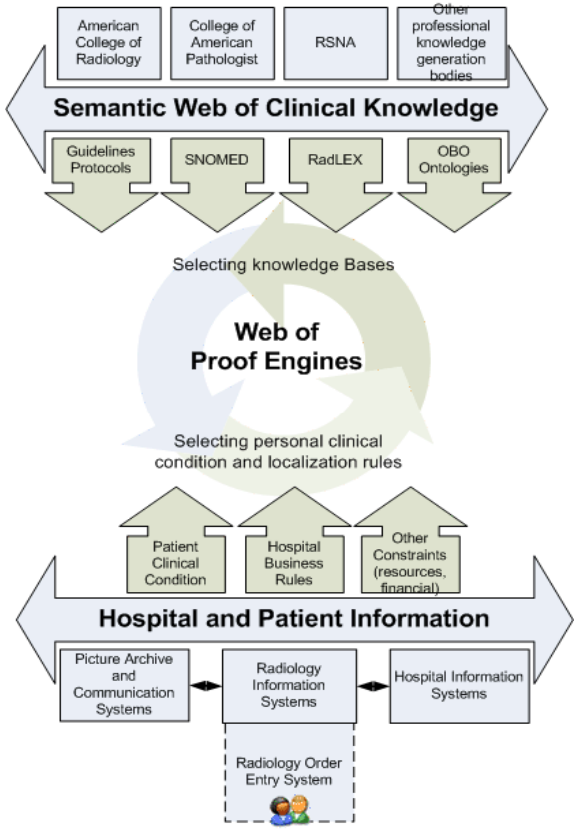
Agenda

- Semantic Web Implementation Use Cases
- Enterprise-ready Semantic Web Tools
- Patterns and Lessons Learned




IMPLEMENTATION USE CASES
• AGFA
• Boeing
• British Telecom
• Chevron
• Eli Lilly
• Oracle
• Ordnance Survey
• Renault
• TATA Consulting
• University of Texas
ORACLE

Implementation Use Case by AGFA: Clinical Health Care Decision Support



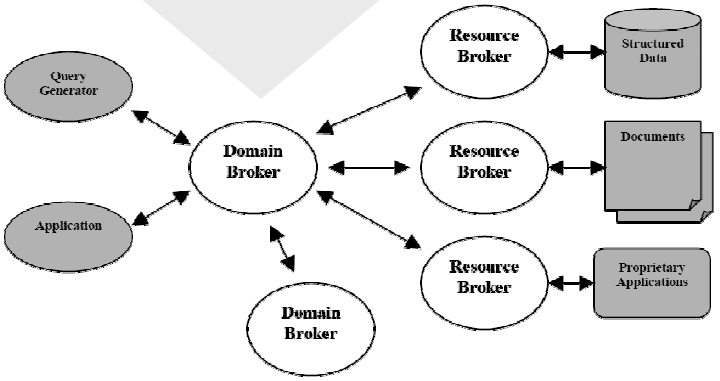
- Supports a natural separation of general medical knowledge captured in the Appropriateness Criteria (guidelines), and the adaptation rules that denote local and execution context. This separation allows knowledge bases to be developed and validated by professional bodies. These knowledge bases end up having better credibility and are easier to keep up to date.
- A standards-based, application neutral platform, for expressing and connecting to the existing corpus of knowledge.
- Ease the burden of developing and maintaining a "complete" knowledge base by one medical organization or vendor.
- Ability to trace the provenance of facts and rules used in medical decision-making, and providing explanation and proof. This is very important in the healthcare domain.

Source: http://www.w3.org/2001/sw/sweo/public/UseCases/

Implementation Use Case by Boeing: Manufacturing and Design Knowledge Exchange

```
<db:DatabaseTable rdf:about="&LDMDEV;AIRCRAFT"
       db:Label="AIRCRAFT">
      <db:PartOfDB rdf:resource="&LDMDEV;LDMDEV"/>
      <db:AssociatedClass rdf:resource="&am;CH-47"/>
</db:DatabaseTable>
<db:DatabaseColumn
       rdf:about="&LDMDEV;AIRCRAFT.AC_SERIAL_NUMBER"
       db:Label="AC_SERIAL_NUMBER">
      <db:PartOfTable rdf:resource="&LDMDEV;AIRCRAFT"/>
      <db:AssociatedSlot rdf:resource="&uo;AircraftSerialNumber"/>
</db:DatabaseColumn>
```

[join db1 "Aircraft" to db2 "CH-47" and db1 "AC_SERIAL_NUMBER to db2 "AircraftSerialNumber"]

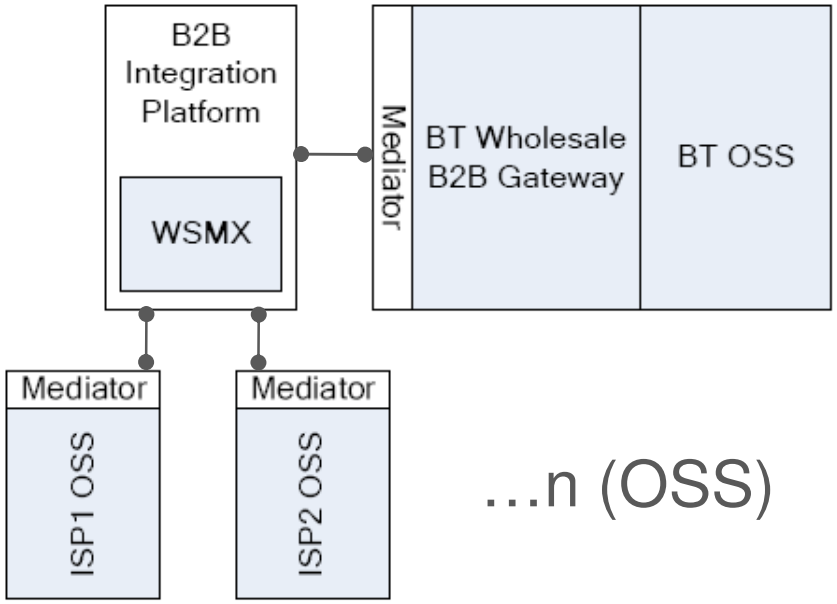


- Reduced cycle time in airplane program development.
- Reduced cycle time in customer service and product support.
- Reduced transaction costs.
- Support for global design and business activities.
- Improved access to content regardless of location or format, allowing for business to be conducted anywhere.
- Flexible infrastructure adaptable to unseen future needs.

Source: http://www.w3.org/2001/sw/sweo/public/UseCases/ & "…Semantic Integration of Corporate Information Resources"

Implementation Use Case by British Telecom: B2B Integration using Semantic Mediation

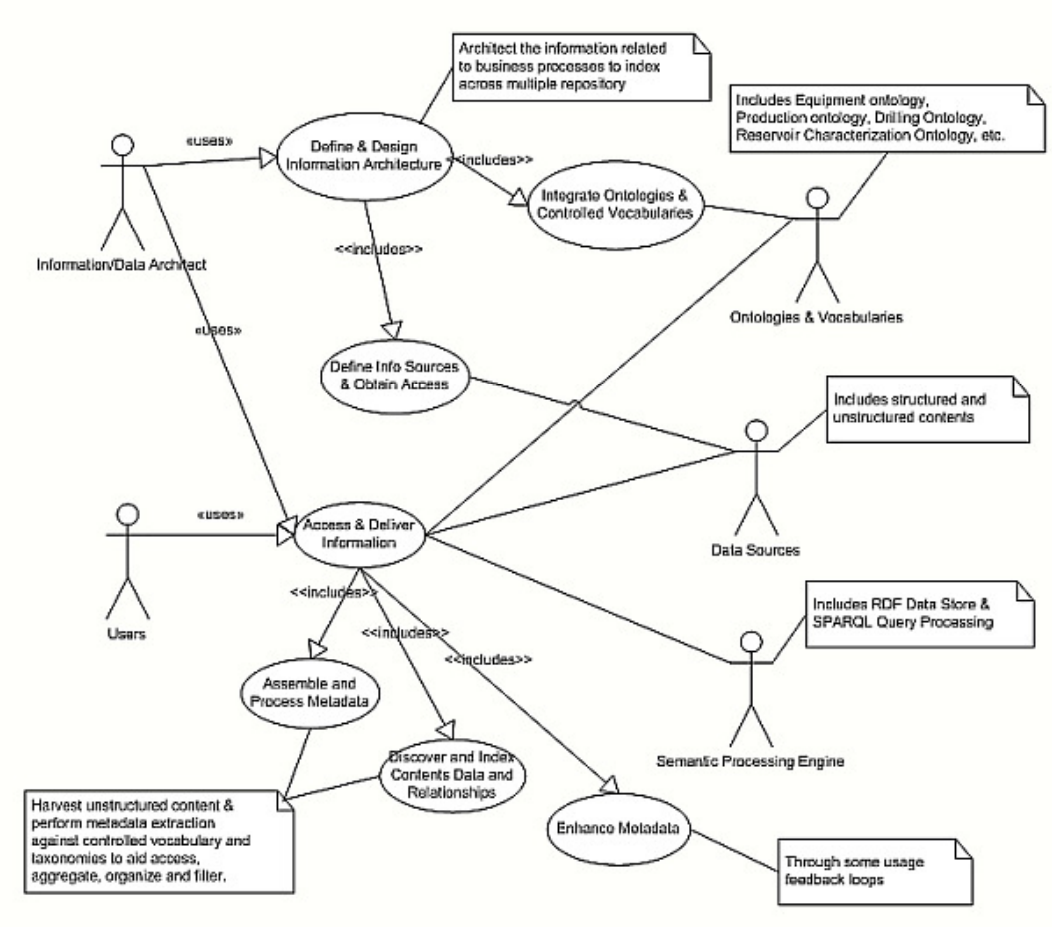
- BT uses semantic descriptions of system interfaces and messages to support integration of Operational Support Systems (OSS)
- Internet Service Providers integrate their OSS with those of BT (via a gateway)



- The approach helps overcome the increasing complexity of supply chains, reduces costs and time-to-market, ontologies allow for a reuse of services

Source: http://www.w3.org/2001/sw/sweo/public/UseCases/

Implementation Use Case by Chevron: Information Integration in Oil and Gas Industry

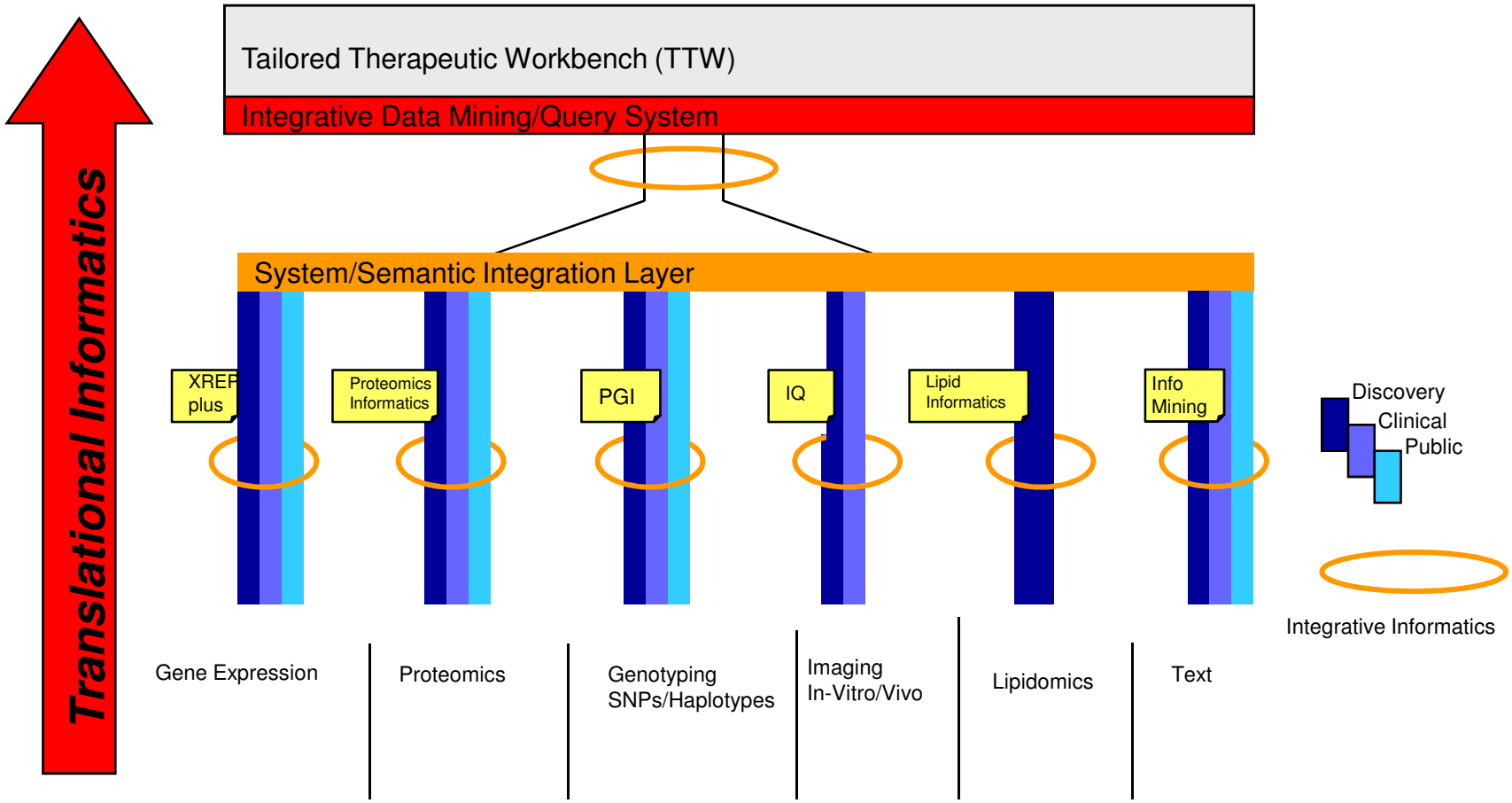


Applications for Oil & Gas:

- Standardization for information exchange between business partners
- Information Integration across applications within a company
- Enable sharing across applications within a company, and with partners and customers
- Collaborative knowledge management for supporting distributed operations and teams

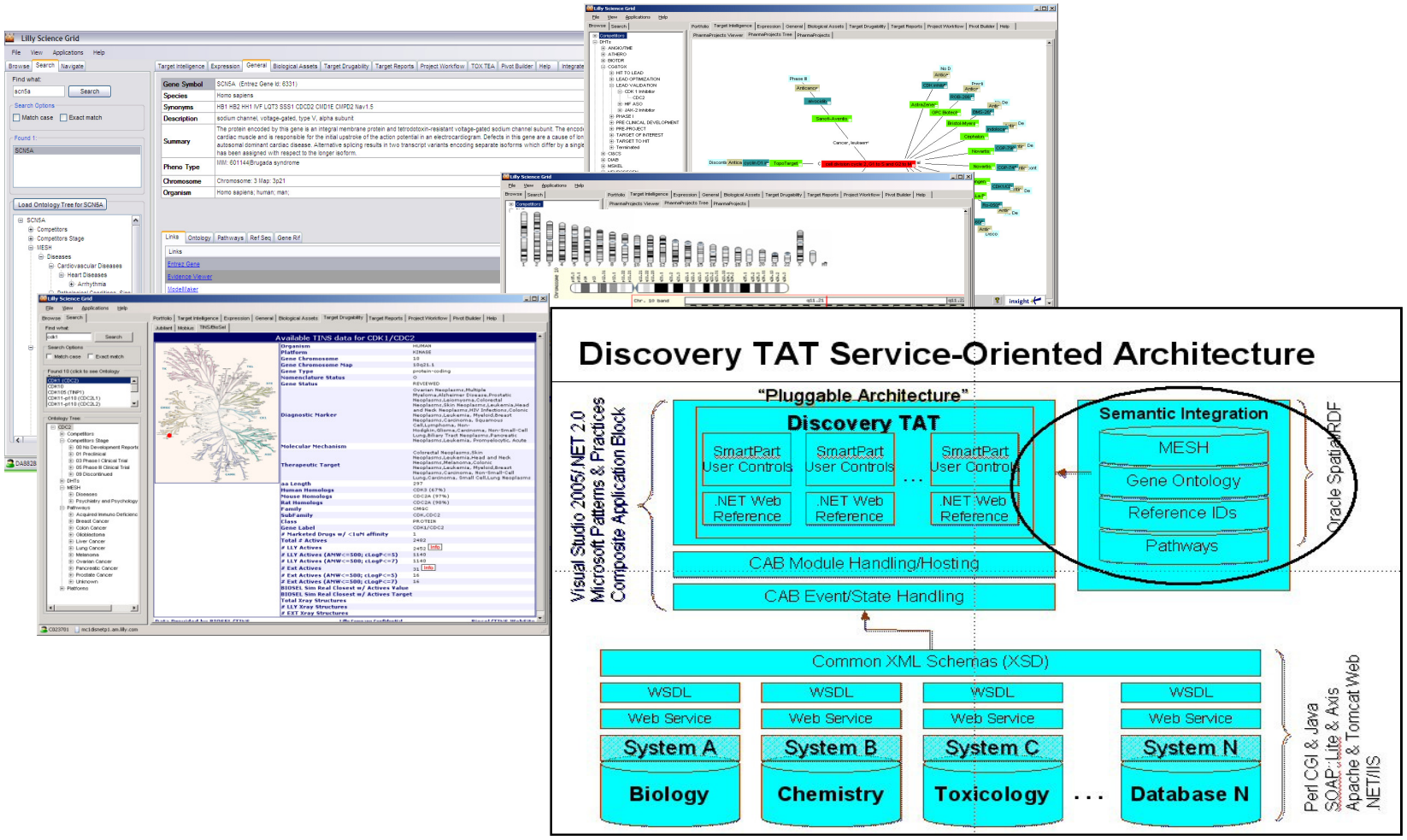
Source: http://www.w3.org/2001/sw/sweo/public/UseCases/

Implementation Use Case by Eli Lilly: Integrative Data Mining and Ad Hoc Query



Source: WWW2007, Eli Lilly, S.Stephens

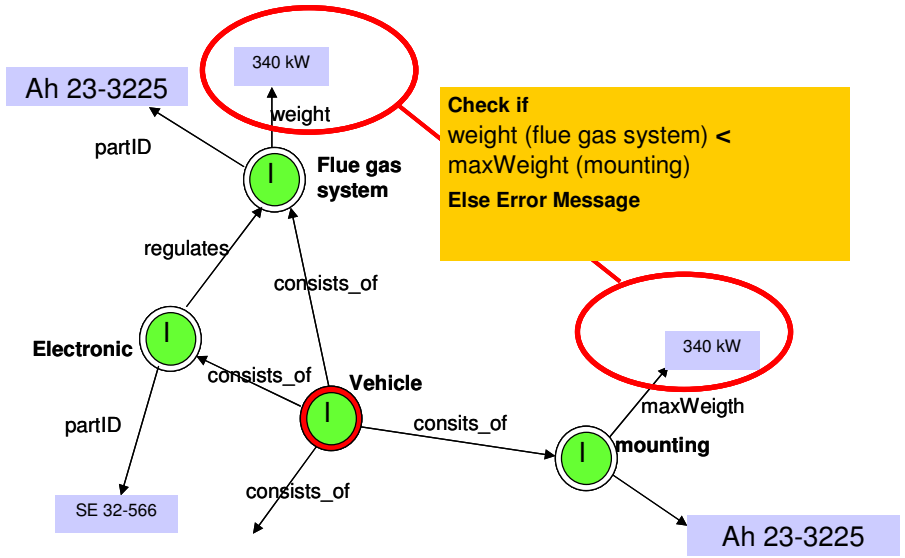
Implementation Use Case by Eli Lilly: Drug Discovery / Target Assessment Tool



Source: WWW2007, Eli Lilly, S.Stephens

Implementation Use Case by Renault: Diagnostic Tools for Complex Machinery

- Use of Semantic Web for generation of a diagnostic engine that can compute necessary procedures on the fly
- A question may be “What are the preliminary steps to test the resistor of the air conditioning engine on a Renault Clio 2006?”
- RDF is used as a flexible bus for exchange of information that originated from engineering departments to repair shops



- OWL is used to model a repository of repairs and diagnostic methods
- Reasoning tools are based on Boolean and probability constraints compilation
- A REST web services architecture is used, with the services returning RDF or HTML depending on content negotiation

Source: S.Stephens STC07, Ontoprise GMBH, Adaptive Information, Pollock & Hodgson

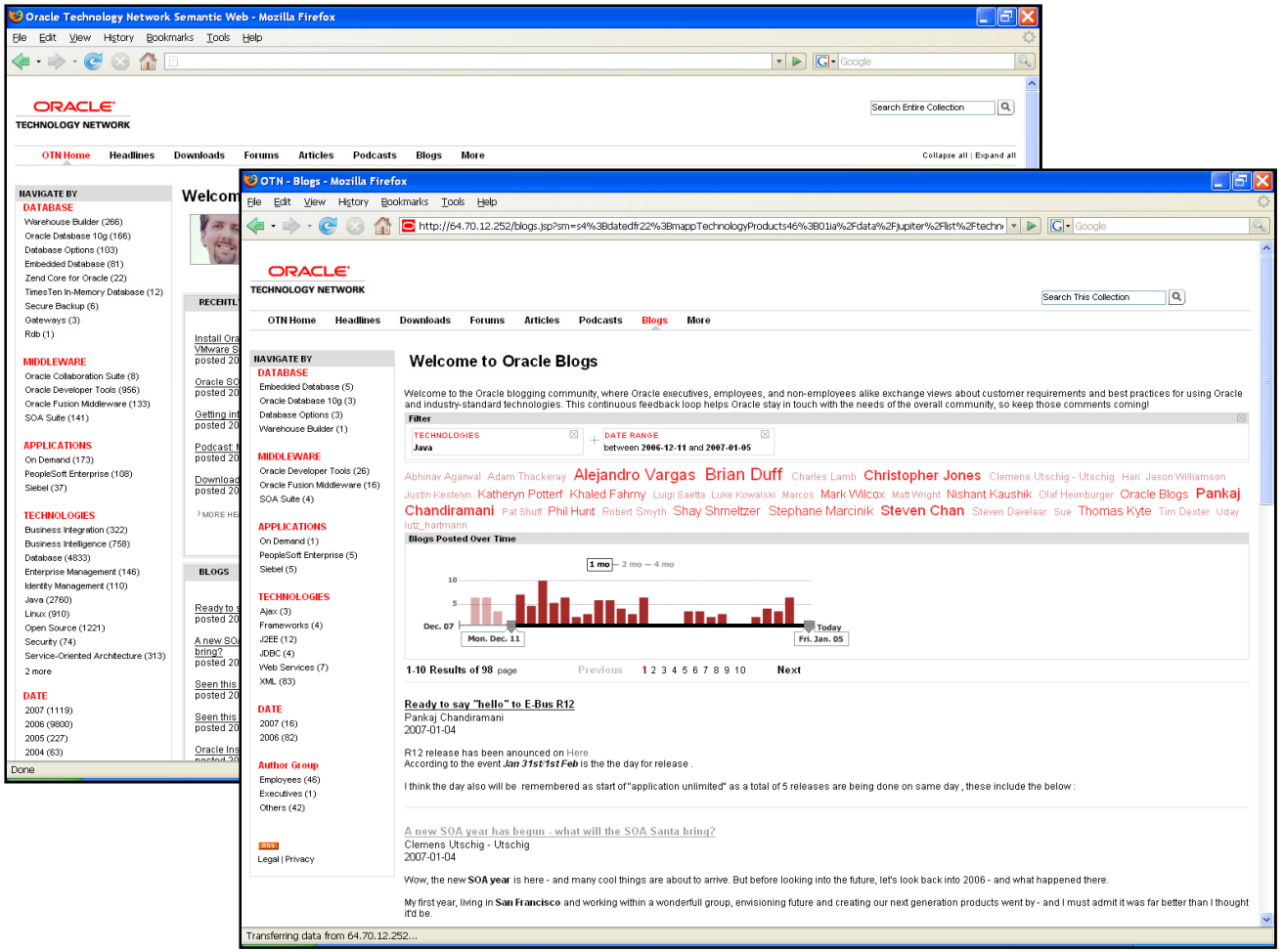
Implementation Use Case by Oracle: Content Search for Collaborative Communities

Oracle Technology Network (OTN) aggregates many sources of content through a single portal

Oracle's taxonomy is used for annotation of news feeds

Semantic Web allows dynamic re-aggregation of results and leads to more comprehensible search results

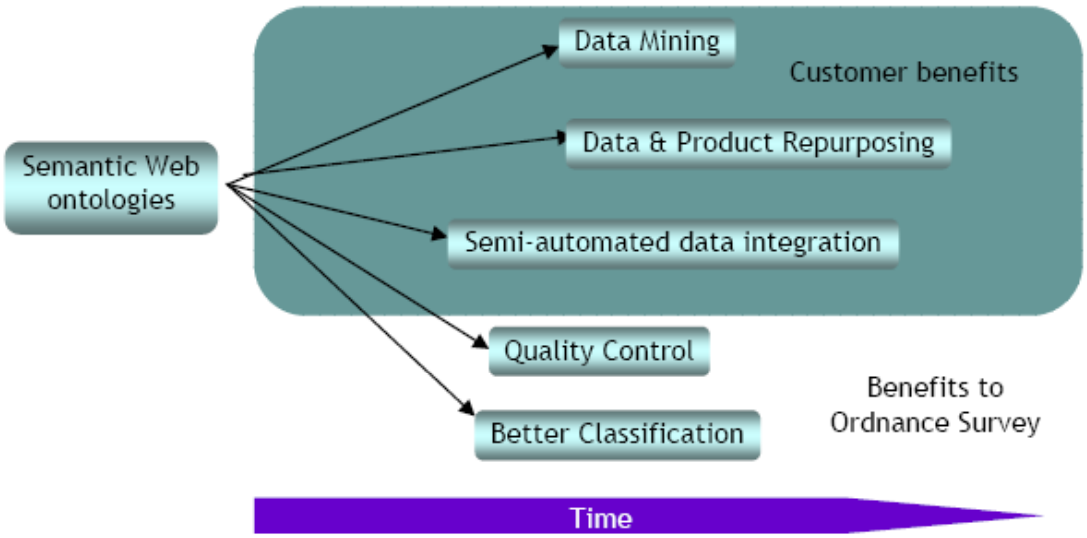
Advantages include enhanced search and navigation, and more powerful user interface



Source: http://www.w3.org/2001/sw/sweo/public/UseCases/

Implementation Use Case by Ordnance Survey: Information Integration using Geographic Data

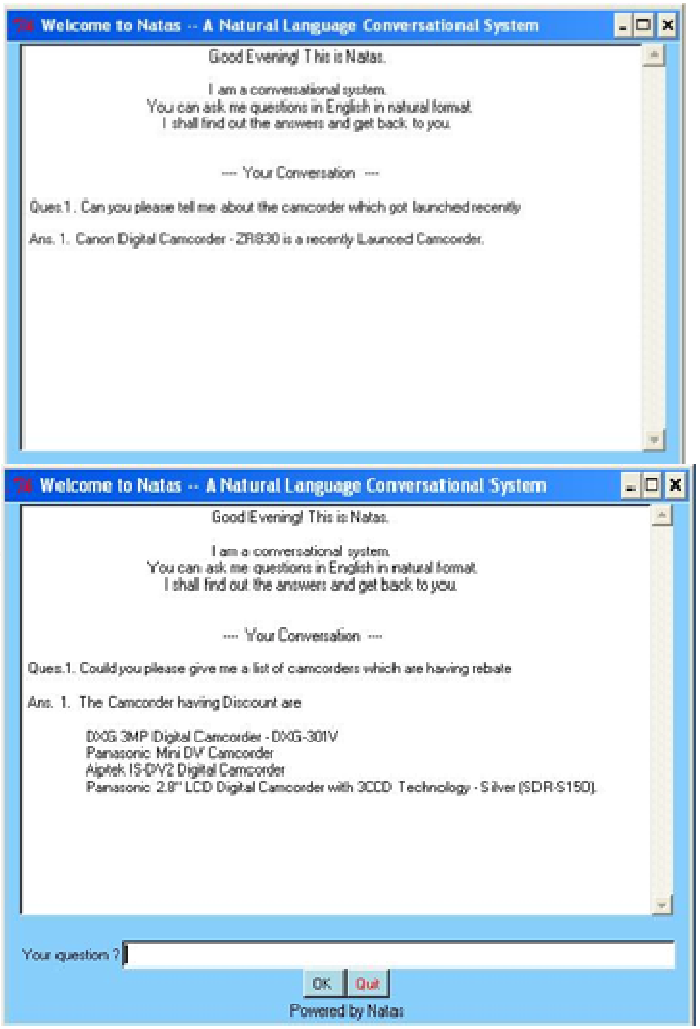
- Ordnance Survey maintains definitive mapping data of Great Britain, the world's largest and most detailed Geo DB
- Semantic Web is used to integrate different, semantically diverse sources of data

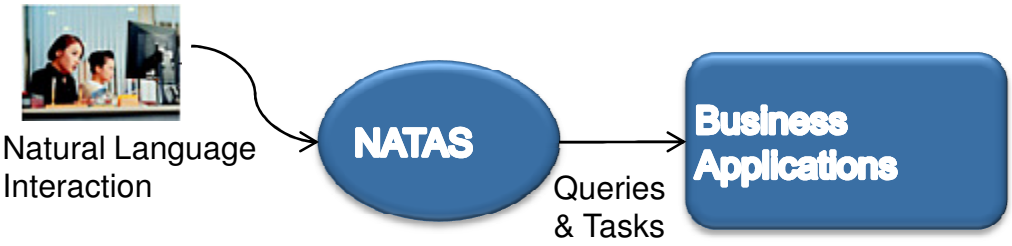


- General ontologies already developed to bridge differences in terminology
- The data is queried efficiently via the ontology or RDF
- Advantages include efficient data integration, data repurposing, and better quality control and classification

Source: http://www.w3.org/2001/sw/sweo/public/UseCases/

Implementation Use Case by Tata Consulting: Natural Language Interface to ERP Applications



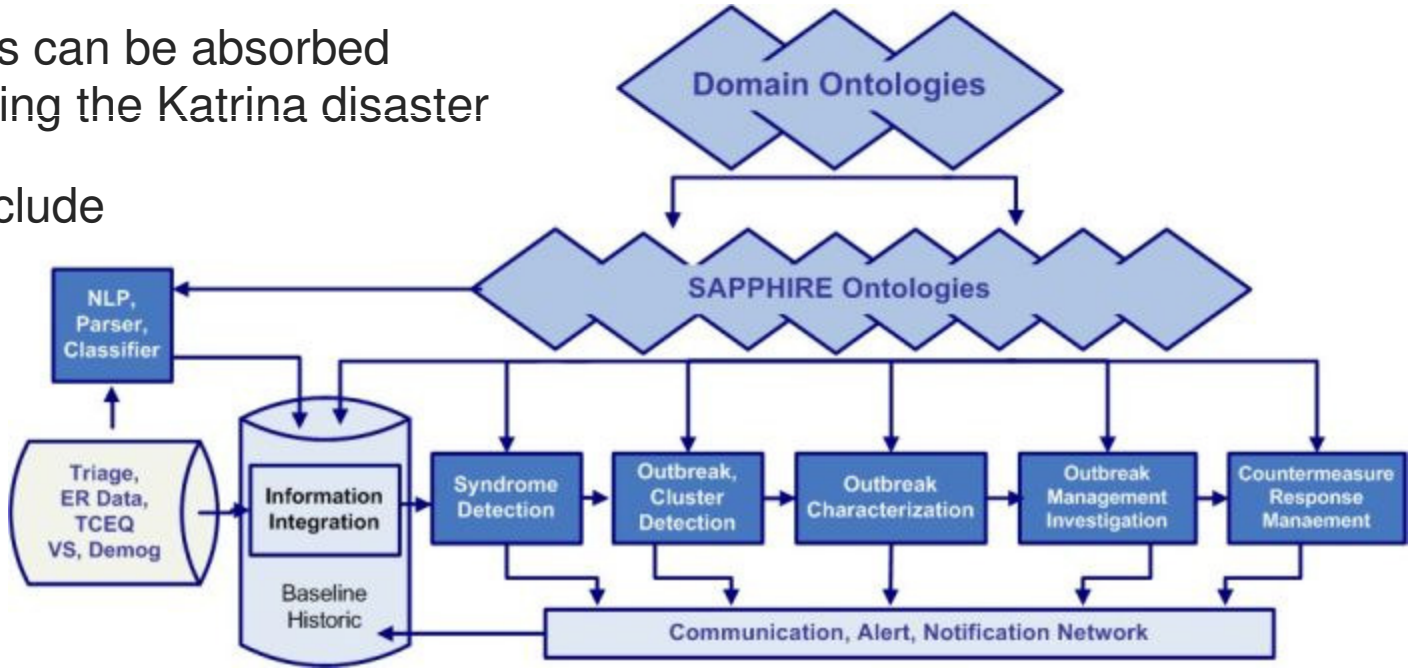


- Enables distinct semantics for the various concepts in the domain, through definition of multiple schemas
- Provides a crisp and simple mechanism to represent an ontology using the <s-p-o> structure of RDF
- Provides mechanisms to formulate generic queries (SPARQL) and instantiate them at runtime in order to answer the queries posed.
- Provides rule evaluation and execution mechanism to create *derived facts*
- Provides mechanisms to link in *external concepts* with existing concepts of the domain through simple <s-p-o> structures

Source: http://www.w3.org/2001/sw/sweo/public/UseCases/

Implementation by University of Texas: Public Health and Disaster Preparedness

- A system that integrates data from multiple disparate sources
- Data can be viewed from many different perspectives, e.g. disease surveillance, environmental protection, biosecurity and bioterrorism, veterinary surveillance
- New data feeds can be absorbed easily, e.g. during the Katrina disaster
- Advantages include dynamic adaptability, and blending of disparate data



Source: http://www.w3.org/2001/sw/sweo/public/UseCases/


VENDOR IMPLEMENTATIONS
• Oracle
• IBM
• Microsoft
• SAP
• Siderean
• Ontoprise
• Software AG / webMethods
• TopQuadrant
ORACLE

Vendor Implementations: Survey of Tools

Triple Stores

- @Semantics RDFStore
- Franz Allegrograph
- IBM Boca
- Intellidimension RDF Gateway
- Northrop Grumman Tucana Suite
- Ontotext OWLIM
- OpenLink Virtuoso
- Oracle Database 10.2

- **Enterprise Search and Collaboration**
- OpenLink Semantic Web Data Spaces
- Radar Networks
- Siderean Seamark Navigator

Reasoners

- RacerPro

Middleware

- IBM WSSR
- Microsoft Connected Services Framework
- Ontology Works
- Ontoprise
- Oracle Fusion Middleware 11g
- Profium Semantic Information Router
- Software AG (webMethods)
- Thetus Publisher

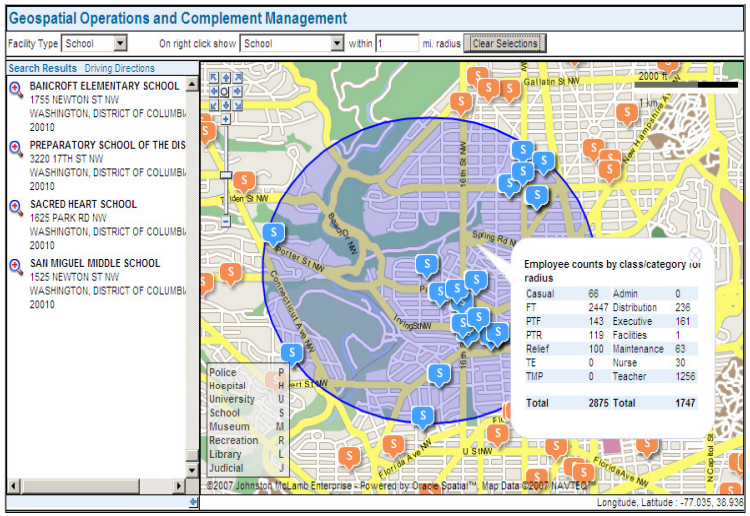
Metadata Tagging

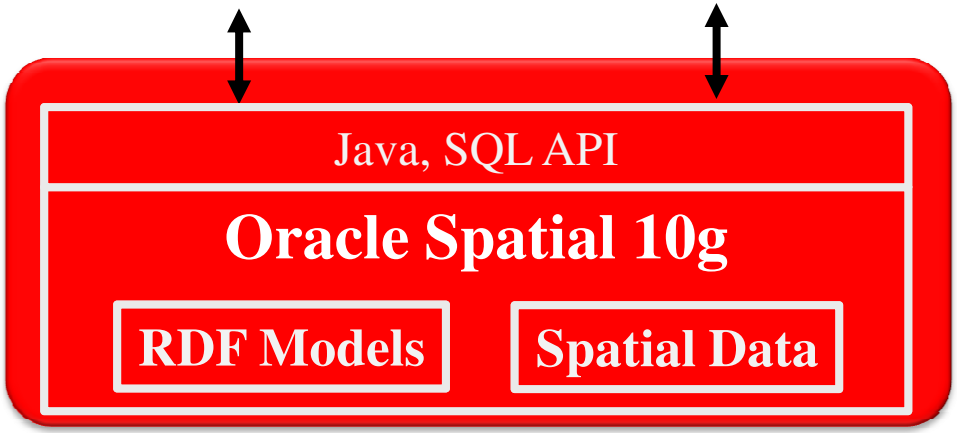
- Adobe XMP & Aduna Metadata Server

Development Environments

- Altova SemanticWorks
- HP Jena
- TopQuadrant TopBraid Composer

Vendor Implementation by Oracle: Oracle Spatial Database with RDF





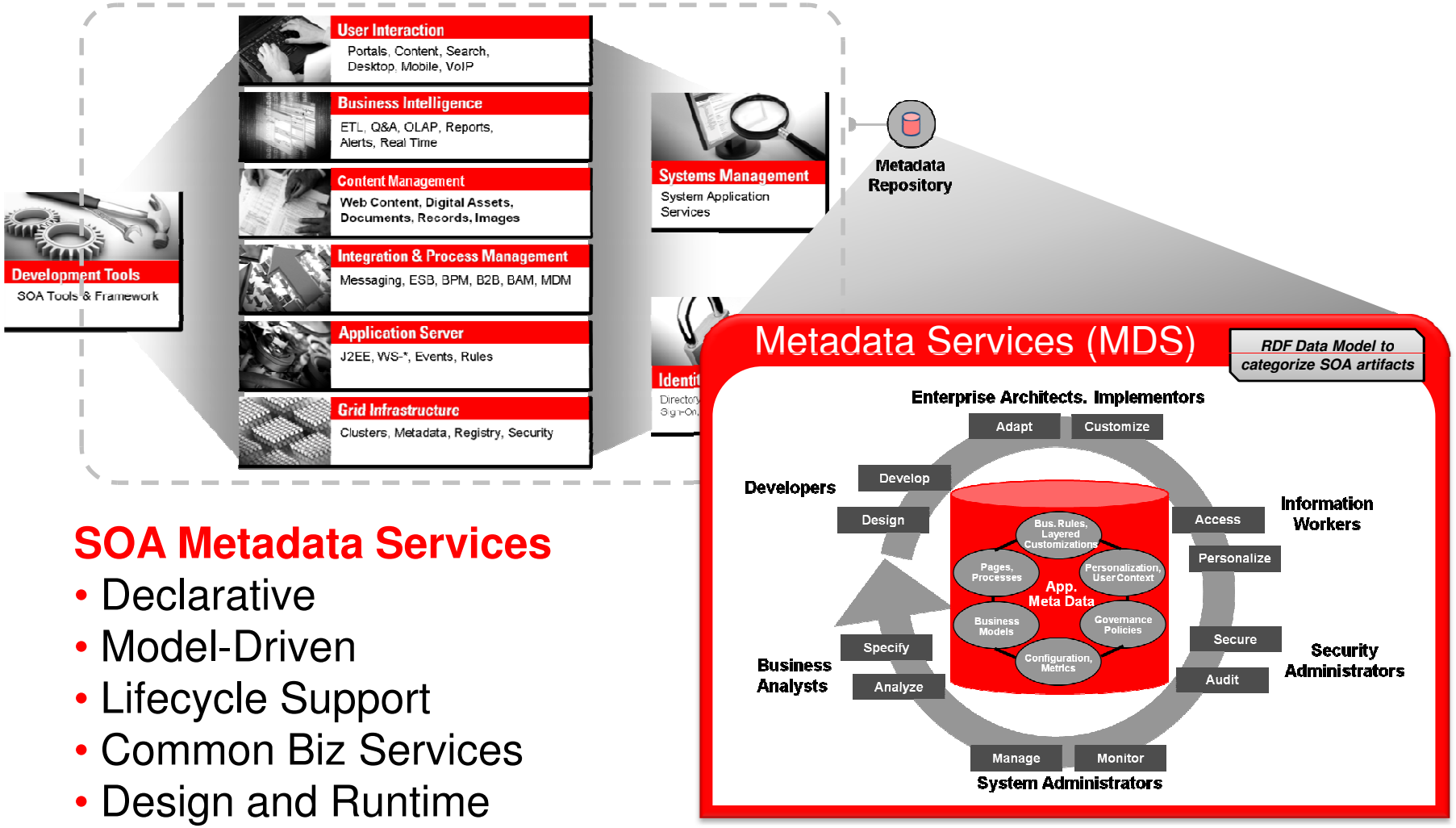
Semantic Tools

- Map Visualization
- Spatial Analysis
- Graph Visualization
- Semantic Search

Data Types

- Persistent RDF/OWL data
- Persistent spatial data
- Persistent raster data

Vendor Implementation by Oracle: Oracle Fusion Middleware Metadata Services

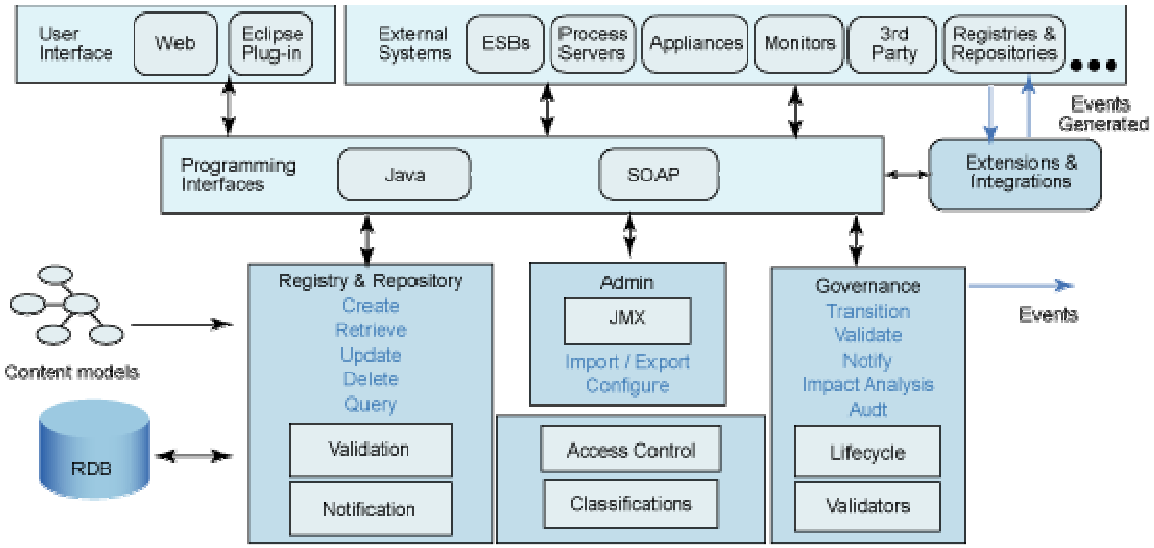


SOA Metadata Services

- Declarative
- Model-Driven
- Lifecycle Support
- Common Biz Services
- Design and Runtime

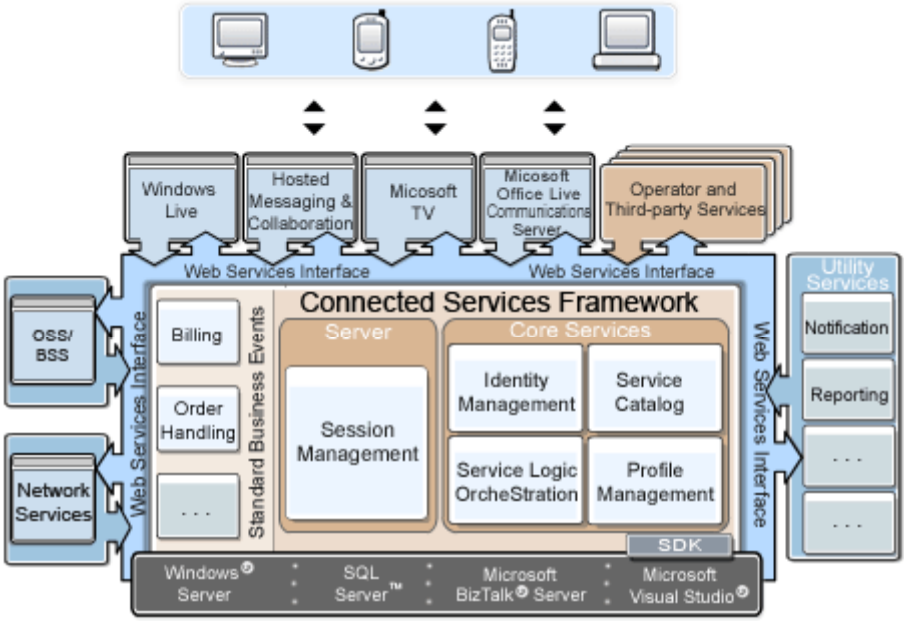
Vendor Implementation by IBM: webSphere Service Registry

"You can load classification systems into Registry and Repository where they can then be used to apply semantic meaning to Service Description Entities. Classification systems are **documents encoded using the Web Ontology Language (OWL)**; see the Resources section for details. While any valid OWL document can be used as a Classification system, at this point in time Registry and Repository exploits only a small subset of the expressiveness of OWL."



http://www-128.ibm.com/developerworks/websphere/library/techarticles/0609_mckee2/0609_mckee2.html

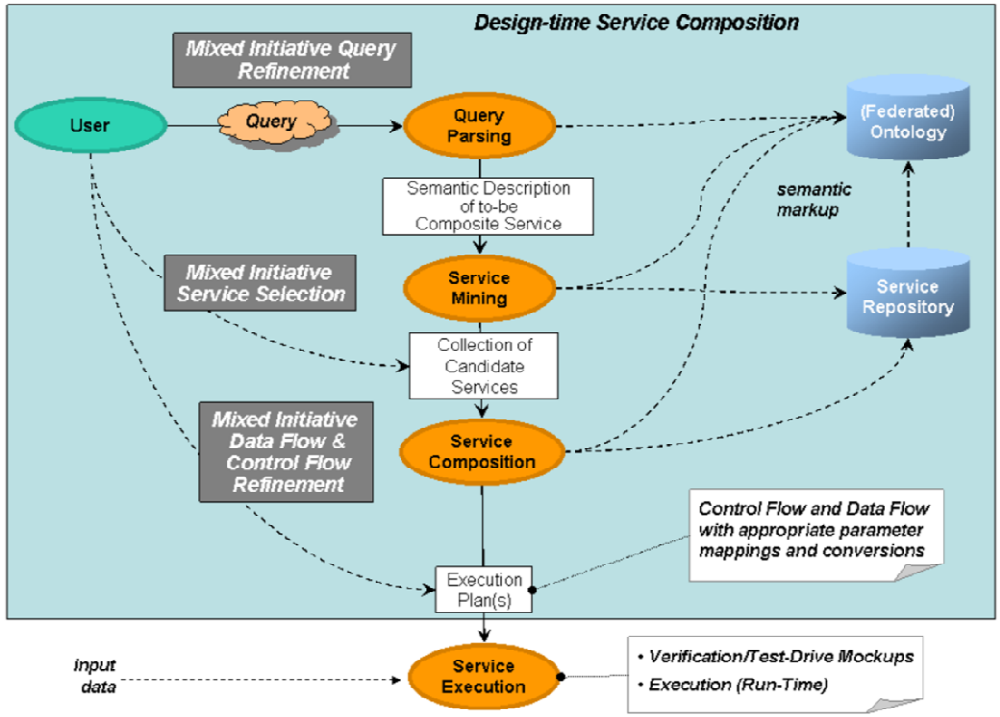
Vendor Implementation by Microsoft: Connected Services Framework



"The Profile Manager component provides profile management services. Connected Services Framework uses Profile Manager to store custom information about users and their preferences. **Profile information is held in a Resource Description Framework (RDF) store**, which is implemented by a Microsoft SQL Server database. Profile Manager provides facilities for creating and managing user profile information and for propagating profile information to Web services that cooperate in a service-oriented application."

http://msdn2.microsoft.com/en-us/library/aa303492.aspx

Vendor Implementation by SAP: Guided Procedures Framework (Netweaver/ESA)

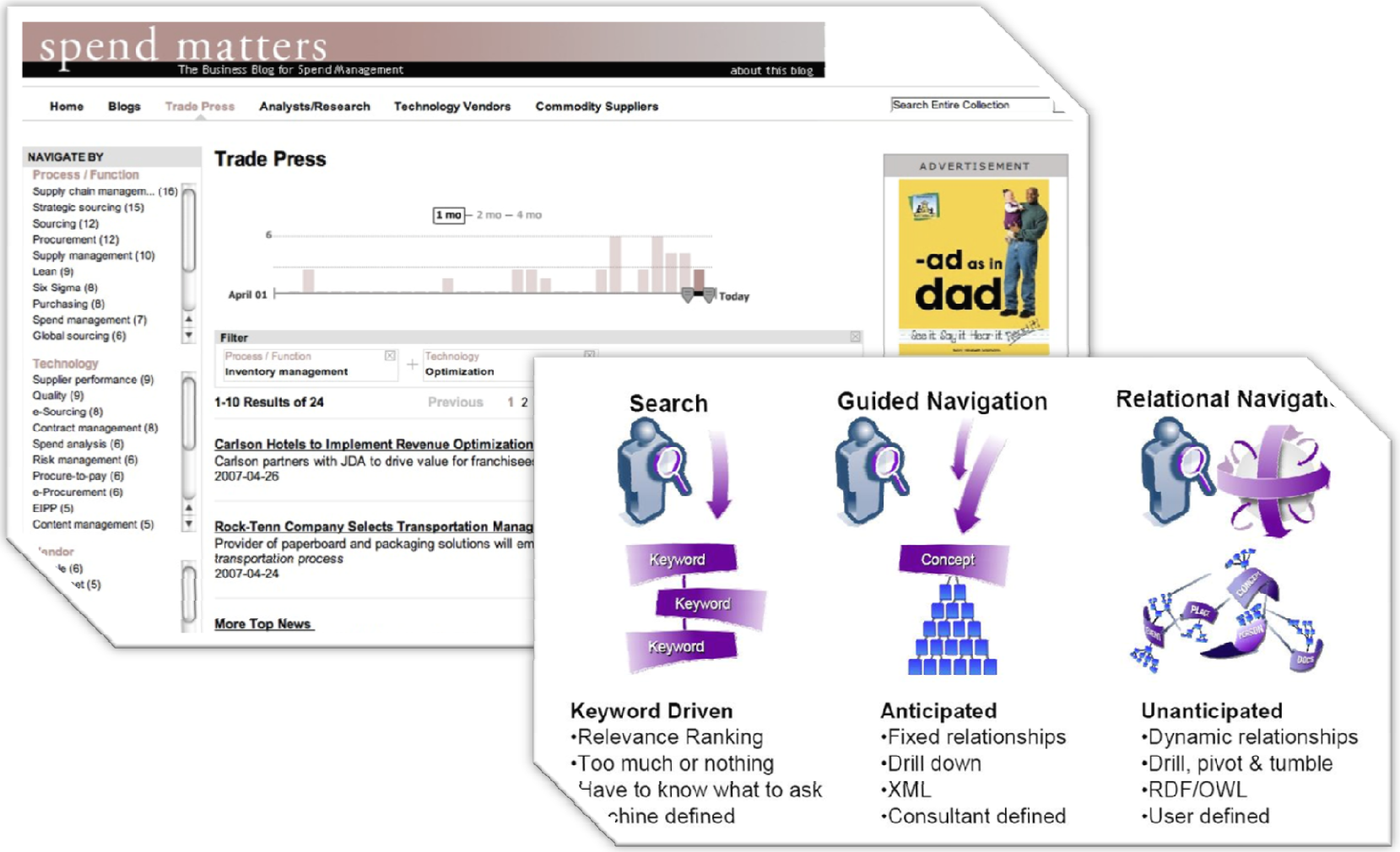


"…we describe a mixed initiative framework for semantic web service discovery and composition that aims at flexibly interleaving human decision making and automated functionality in environments where annotations may be incomplete and even inconsistent.

An initial version of this framework has been implemented in SAP's Guided Procedures, a key element of SAP's Enterperise Service Architecture (ESA)."

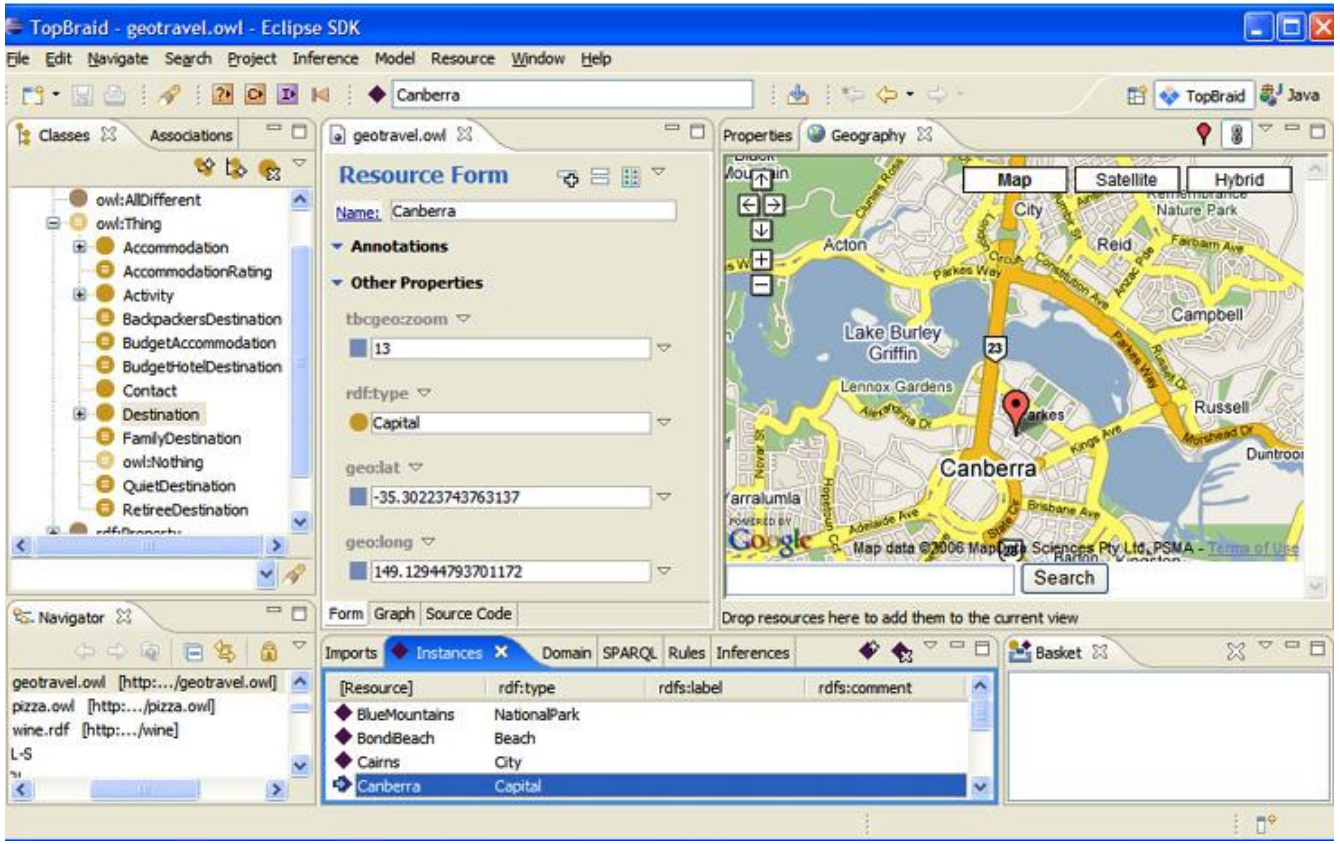
http://iswc2006.semanticweb.org/items/Rao2006lr.pdf

Vendor Implementation by Siderean: Siderean Navigator, MAPP, and Analytics



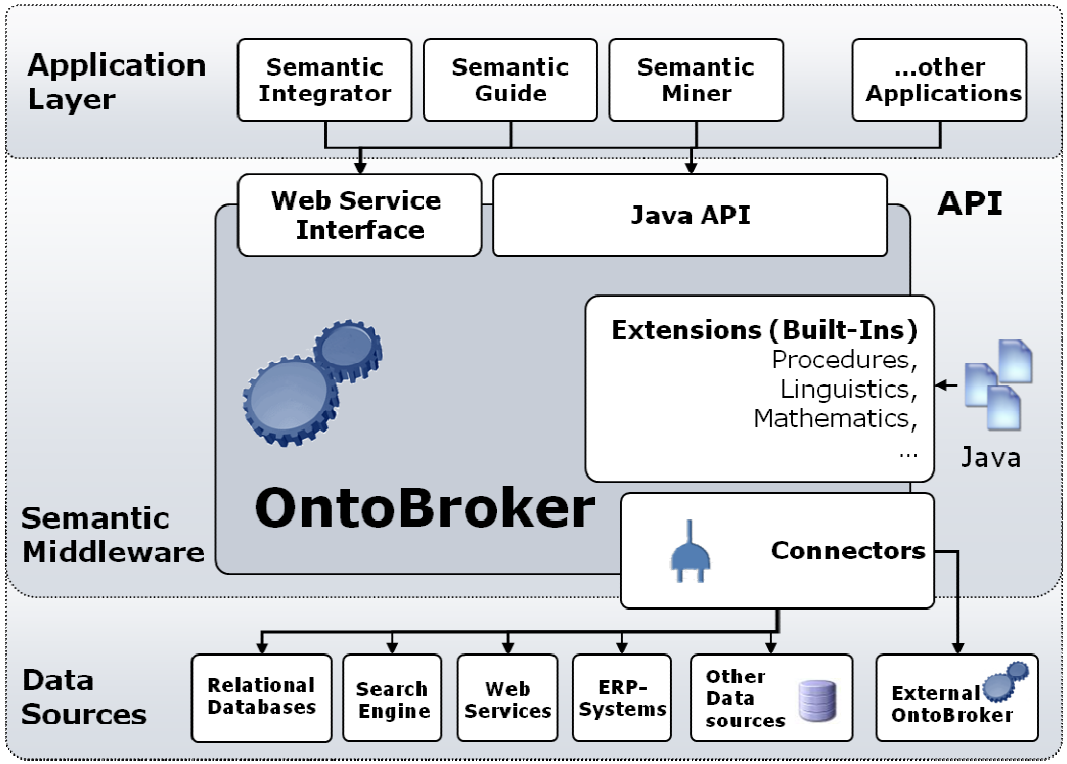
Vendor Implementation by TopQuadrant: TopBraid Composer

- Multi-user ontology editor based on Eclipse
- Jena and SWRL rules are supported
- Interface for editing SPARQL queries
- Create diagrams from class definitions in RDFS/OWL
- Classification and consistency checking
- Semantic mapping of models
- Geography and location mapping
- Support for mashups
- Visual RDF graphs



Vendor Implementation by Ontoprise: OntoBroker, OntoStudio and Ontoprise Apps

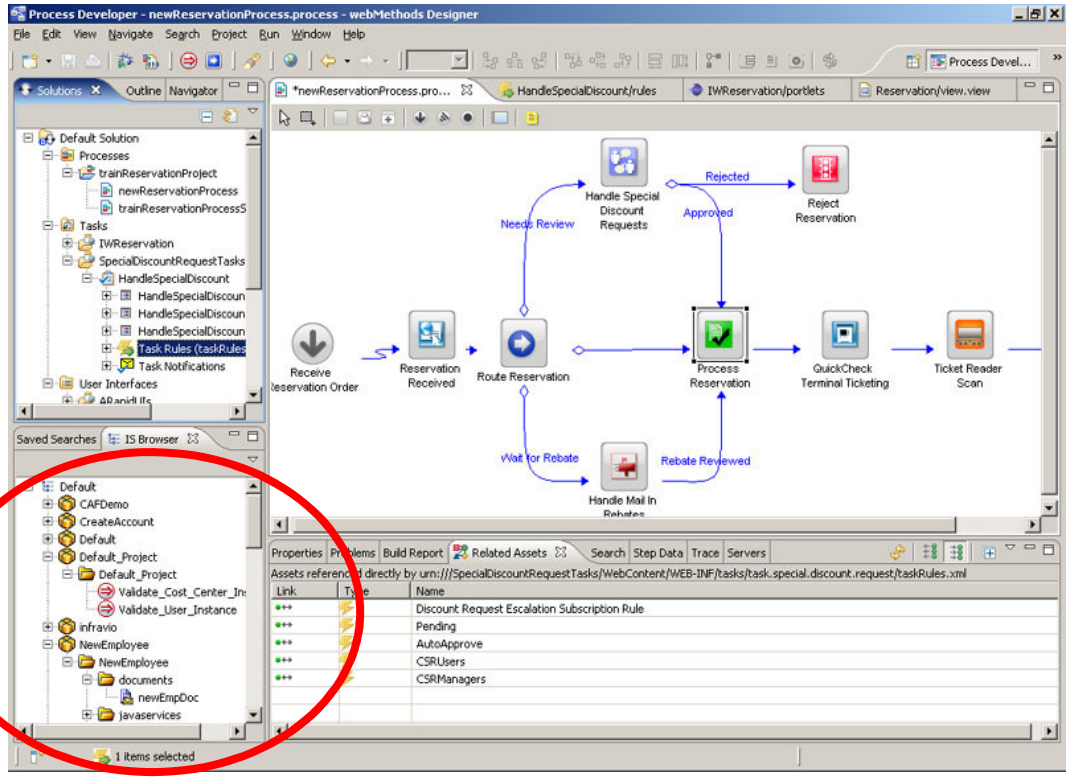
OntoBroker 5.0
The leading inference engine OntoBroker processes ontologies and the logic represented inside them. With version 5.0 distributed OntoBroker-architectures can be employed to answer very complex queries very fast or to deal with high query rates. The information and the query processing is hereby distributed over multiple OntoBroker instances.



This makes OntoBroker a highly-scalable semantic middleware solution that e.g. allows for integrating data from different locations. The support of the W3C Standards OWL and RDF(s) as well as F-Logic guarantees a very high interoperability.

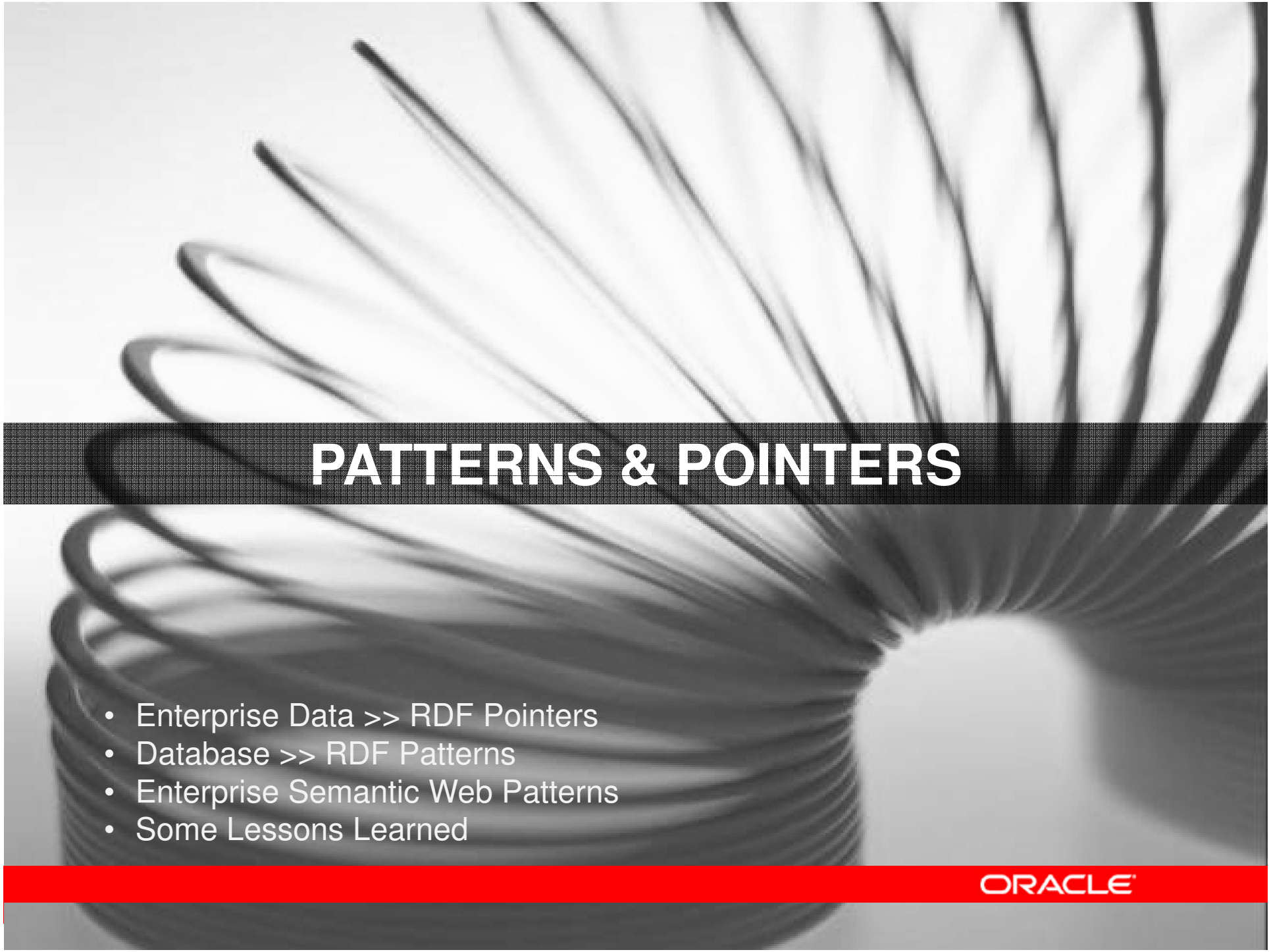
Vendor Implementation by SoftwareAG: Centrasite & webMethods Metadata Repository

"What Customer Needs are Being Addressed?
SOA and BPM inherently increase the complexity of the IT environment by introducing more moving parts. To realize long-term ROI from SOA and BPM, organizations have to manage this complexity in the form of understanding what SOA/BPM assets exist, what they do, how they are used, and how they inter-relate—this is the role that semantic metadata technology plays.



The Cerebra technology will **address the need to track, manage, and aggregate information about components such as Web services, document definitions, and the processes and applications that utilize them**. Without this information, organizations will be challenged to increase reuse, manage change (e.g., what happens if I change this Web service?), and promote consistency across implementations."

http://www1.webmethods.com/PDF/CerebraFAQs.pdf


PATTERNS & POINTERS
• Enterprise Data >> RDF Pointers
• Database >> RDF Patterns
• Enterprise Semantic Web Patterns
• Some Lessons Learned
ORACLE

Pointers:
Converting Enterprise Data to RDF

- Relational -> RDF
- XML -> RDF
- Excel -> RDF
- JPEG -> RDF
- BibTEX -> RDF
- Java -> RDF
- Weather -> RDF
- Palm -> RDF
- Outlook -> RDF
- Flickr -> RDF

A directory of RDFizers is provided at:
http://simile.mit.edu/wiki/RDFizers

Patterns:
Using Relational Data as RDF

Commercial Tools:

- Siderean
- Ontoprise
- Franz Technologies
- OpenLink

Key Patterns:

- RDF as Views
- RDF as Proxy

Other Approaches:

- D2RQ
- SPASQL
- DartGrid
- SquirrelRDF

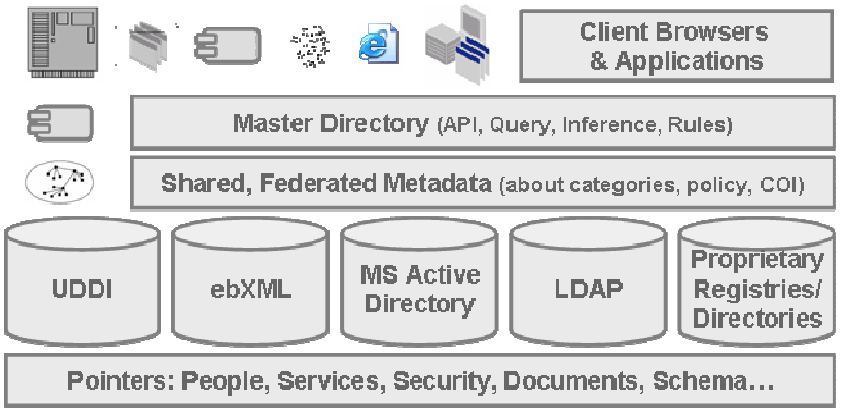
New Work:

- W3C Workshop in Oct 2007 concurrent with ANSI SQL Stds

Patterns: Sem-Tech for Registry / Repository

Registry and Repository services may leverage the advanced classification and taxonomic features of semantic technology for **automatic classification** and more **agile change management** in highly dynamic data environments (such as multi-community registries)

conceptual architecture:



operational capabilities:

- Discover Services
  - UDDI style registries with much better discovery
- Discover other Assets (IT, HR, Doc…)
  - Use the registry for other assets, anything with a pointer/key
- Align Categories and Keywords
  - From Different Registries etc.
- Apply Consistent Policy & Governance
  - Across Communities of Interest

technical approach:

- Open Source Technical Approach
  - OWL/RDF Models | Domain Info
  - Pellet/KAON Reasoner | Query KB
  - SOA using SCA patterns | Utility API
  - Apache JUDDI core
- Custom Tooling
  - Category Management GUI (+policy)
  - Import/Export to OWL/RDF

order of magnitude:

Approximate Level of Effort

Phase 1: ~18 man-months*

- Assuming ~3 sources & ~3 COI's

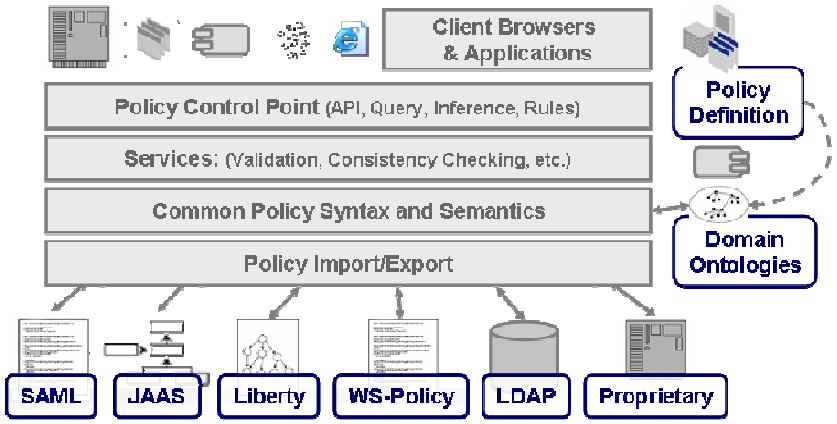
Known Implementations

OWL aware ebXML, SD Supercomputer Center, webMethods UDDI/RDF paper, Software AG Centrasite Registry, IBM, SAP…

Patterns: Sem-Tech for Data Governance

Data Governance is first and foremost a "people process," but IT can leverage the advanced classification and business rule features of semantic technology for **inferring policies** and **checking data consistency** across heterogeneous application environments

conceptual architecture:



operational capabilities:

- Common Hub for Policy Definitions
  - Rule taxonomy and logics
- Unified Policy & Compliance Visibility
  - One place to view and change policy
- Central Governance Checkpoint
  - Validate semantic/structural consistency
- Automatic, Reliable Impact Analysis
  - Reasoners and rules automate tasks

technical approach:

- Open Source Technical Approach
  - OWL+Rules | General Policy Formats
  - Pellet/KAON Reasoner | Query KB
  - SOA using SCA patterns | Utility API
- Custom Tooling
  - Design: Policy Definition and Mapping
  - Runtime: Import/Export policies to OWL/RDF and create a PEP

order of magnitude:

Approximate Level of Effort

Phase 1: ~36 man-months*

- Assuming low level of spec maturity

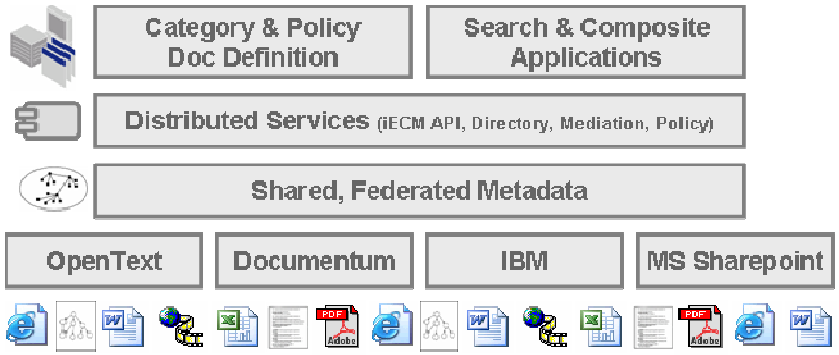
Known Implemenations

University of Maryland Mindlab, European Union Research, OntoGrid & University Manchester, KAON and Ontoprise

Patterns: Sem-Tech for Content Management

Content Management Systems are mature technology, but most companies have too many. The Sem-Tech approach is to layer ontology metadata between the CMS' to provide a **unified classification scheme**, then NLP or search engines **automatically organize documents by ontology keywords**

conceptual architecture:



operational capabilities:

- Bolt-on to Existing CMS Systems
  - Sharepoint, Documentum, IBM, Stellent/Oracle, etc.
- Automatic Classification of Content
  - Based on document text/attributes and declared taxonomy
- Discovered Content in Local Views
  - Local context for new document content
- Unified Interface to Disparate CMS/ECM
  - iECM API in front of semantic technology

technical approach:

- Open Source Technical Approach
  - OWL/RDF | Domain Models
  - Pellet/KAON Reasoner | Query KB
  - Instance Store & Mapping (Kowari etc.)
  - SOA using SCA patterns | Utility API
- Custom Tooling
  - Ontology bridging (axioms++) heuristic
  - Runtime: ECM adapters for generating document metadata and pointer system

order of magnitude:

Approximate Level of Effort

Phase 1: ~36 man-months*

- Main issue: adaptors & instances

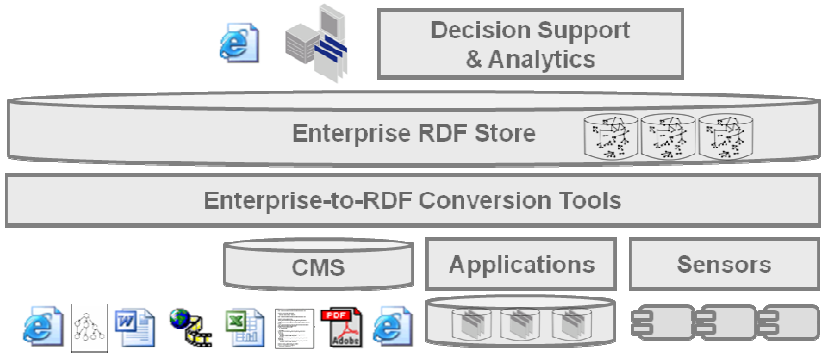
Known Implementations

EMC R&D, Networked Planet, Cerebra, Semagix, Eastman Kodak, SchemaLogic, IBM, Siderean/Oracle

Patterns: Sem-Tech for Decision Support

Decision Support systems come in many flavors. Data mining, Cube-based OLAP, and Star-based Business Intelligence are but a few. Graph data structures may also be used for analytics, which turns out to **work very well when the data structure, relations, or sources are prone to change**

conceptual architecture:



operational capabilities:

- Augmentation for Business Intelligence
  - Different "style" of navigating data (vs. Cubes, Snowflake, etc.)
- Leveraged for "Hard" Analytic Problems
  - Schema and relations change frequently and dramatically
- Automatic (re) Classification
  - Continually asserting new facts and adding data and re-categorizing
- Bulk Data Import is Remarkably Easy
  - Tuning the automatic classification features is usually hard

technical approach:

- Open Source Technical Approach
  - Scalable RDF DB| Domain Models
  - Pellet/KAON Reasoner | Query KB
  - RDF'izer | RDB to RDF Mapper
- Custom Tooling
  - Design: Analytics front-end | Visualization
  - Runtime: Business rules and tuning of the bulk import classification heuristics

order of magnitude:

Approximate Level of Effort

Phase 1: ~36 man-months*

- Main issue: tuning the classifiers

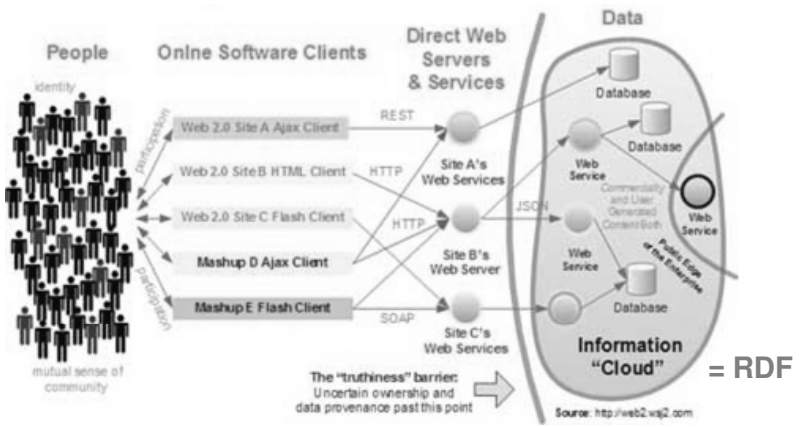
Known Implementations

Eli Lilly, Renault, University of Texas, Ordnance Survey, GlaxoSmithKline, DoD (various)

Patterns: Sem-Tech for Collaboration

Collaboration Software might leverage semantic technology at several layers. Most frequently, it will use RDF/OWL markup to identify classes of things in the data network: People, Places, Resumes, Articles, etc. From there, **everything may be linked, navigated, and re-combined in new ways**

conceptual architecture:



operational capabilities:

- Web 2.0 or Web 3.0 ???
  - If Web 2.0 was for the UI, then Web 3.0 is for the data
- Recombinant Data (AKA: Mashups)
  - Use other people's data for your own purposes - easily
- Automate the Simple Stuff
  - Eliminate the need to ever cut-and-past in your collaboration software
- Harness and Preserve Group Knowledge
  - Web 2.0 empowered the group, but loses fidelity on the data. Semantic technology takes it one step further and can preserve the group knowledge in a 'knowledge base' (aka: RDF)

technical approach:

- Open Source Technical Approach
  - OWL/RDF | Domain/FOAF Models
  - Haystack | DISCO RDF Browser
  - Inference Engine, Pellet (optional)
  - Tomcat, or other servlet implementation
- Custom Tooling
  - User interface and navigation
  - Runtime: most likely a Java or .NET back-end to wire the server and RDF-DB together

order of magnitude:

Approximate Level of Effort

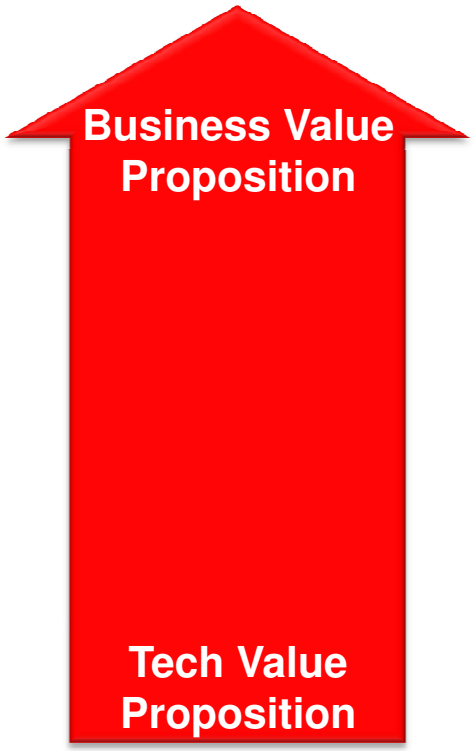
Phase 1: ~18 man-months*

- Main issue: proof & trust (for open nets)

Known Implementations

Radar Networks, Boeing, Chevron, Oracle Technology Network, Blue Chip Experts, Semantic Wiki, Haystack, FOAF etc.

Pointers:
Describe Value to Several Organizations



Business Transformation Enabler

- Gain confidence in consistency and faster process support for organizational change initiatives – reuse data to new tasks
- See business data more clearly: more transparency, more traceability
- Measure change impacts: Impact Analysis, Data Lineage etc.
- Break the ties that bind (brittle, inflexible, expensive data)

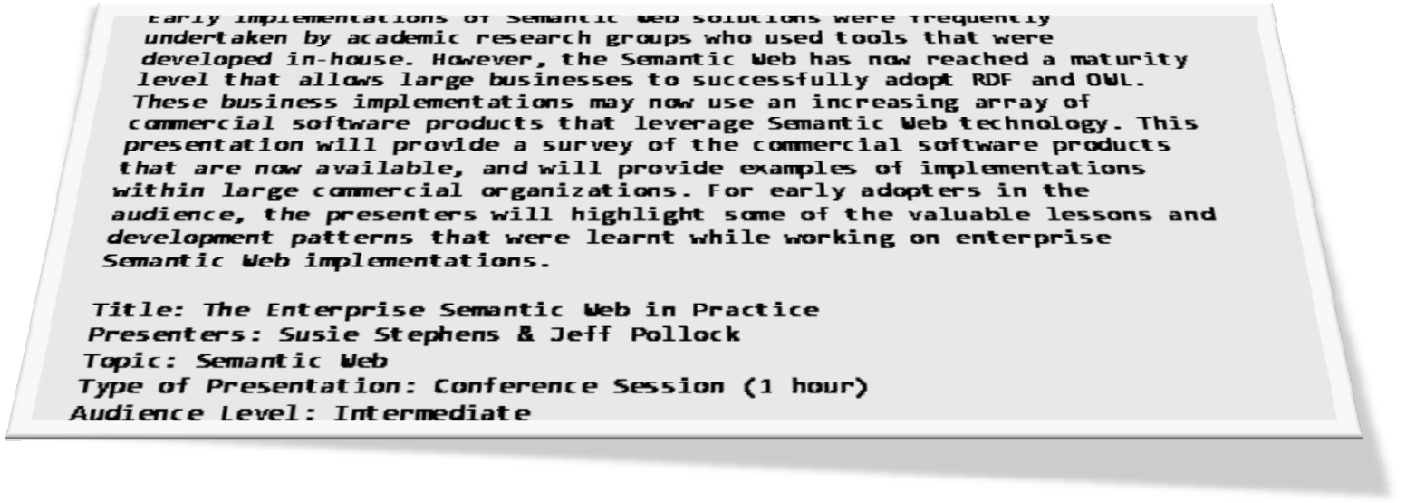
- **IT Service Capabilities**
  - Distribute query, xform, lookup, and search utilities to the edge
- **Centralized Policies**
  - Get a handle on governance, compliance, SLA's and security
- **Automation**
  - Impacts; know how a local change ripples to other communities
- **Inference-Driven Systems**
  - Works with explicit and implicit knowledge, more declarative
- **Better Metadata**
  - Modern, based on logic and math, web-ready, AI capable

Patterns:
Lessons Learned

- Use robust URIs
- Reuse existing data and ontologies
- A little semantics goes a long way
- Model the real world rather than data artifacts
- Build upon your infrastructure incrementally

Summary

- The Semantic Web tool environment is maturing with many commercial offerings throughout the infrastructure stack
- Growing number of large enterprises are implementing solutions that incorporate Semantic Technologies
- Implementations follow similar value scenarios and patterns




Q&A



ORACLE IS THE INFORMATION COMPANY


ORACLE®