



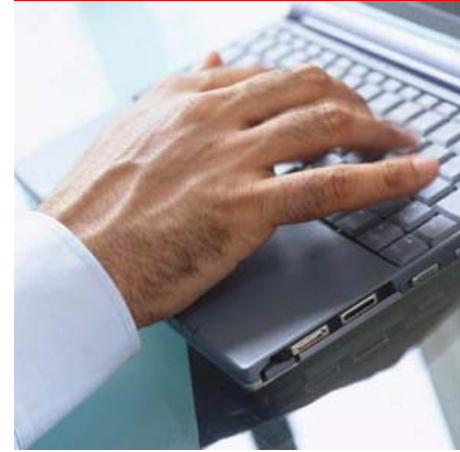
# **Semantic Technology: A Challenge or Opportunity for Building Enterprise Systems**

David Pearson

VP Architecture, Oracle Corporation

# Agenda

- Semantic Web Vision
- Semantic Technology Challenge
- Use Cases
- Semantic Technology Opportunities
- Summary & Conclusions



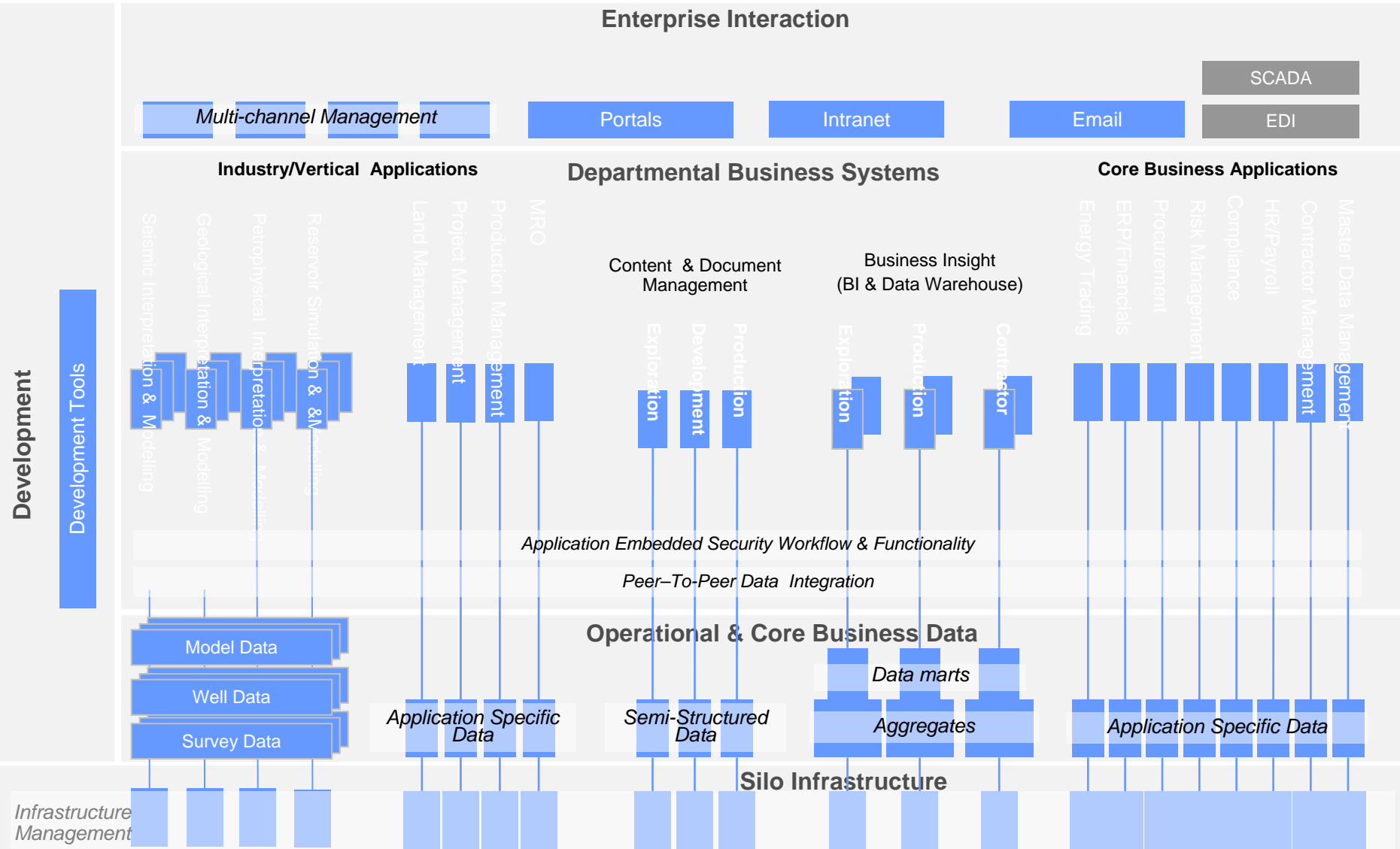
# Semantic Web Vision

*“I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web – the content, links, and transactions between people and computers. A ‘Semantic Web’, which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The ‘intelligent agents’ people have touted for ages will finally materialize.”*

**Tim Berners-Lee, 1999**

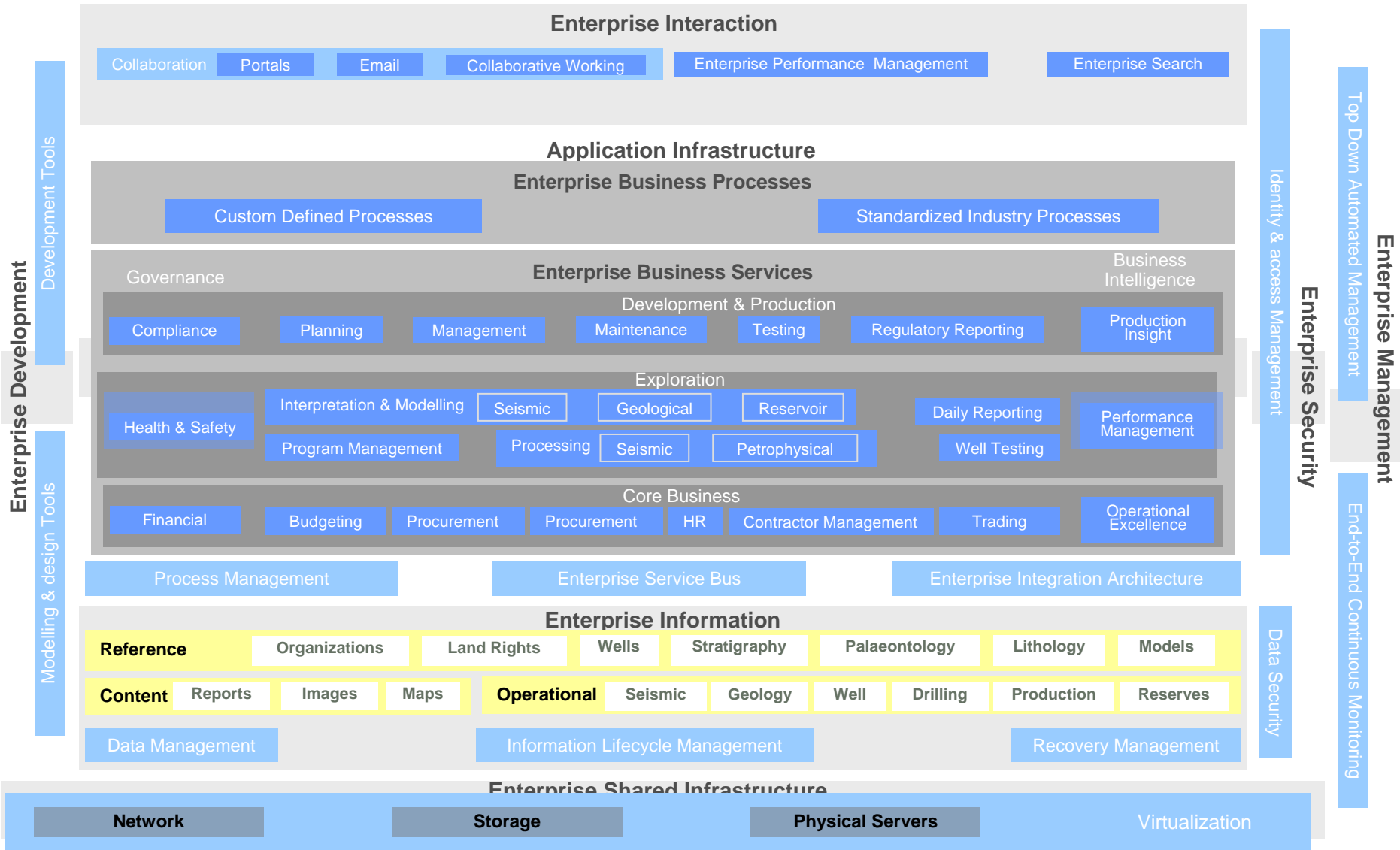
# Upstream Architecture Challenge

## Current State Example



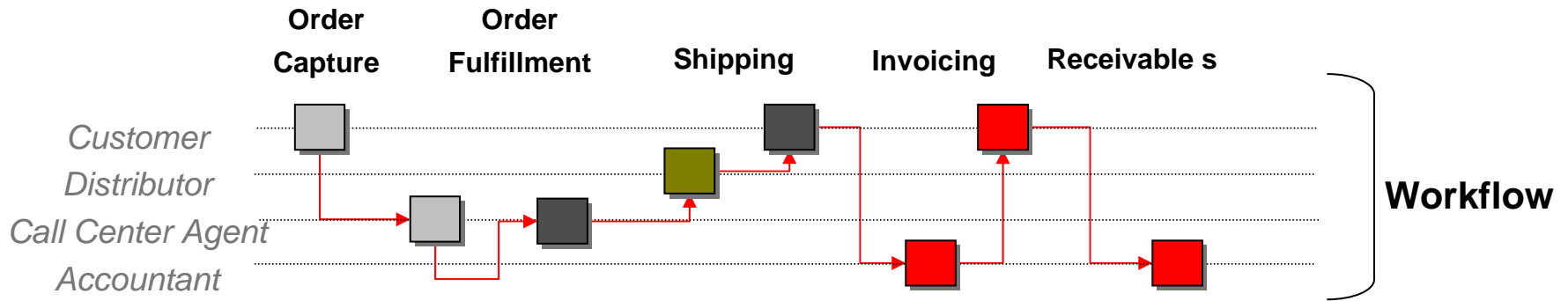
# Upstream Reference Architecture Challenge

## Vision Example

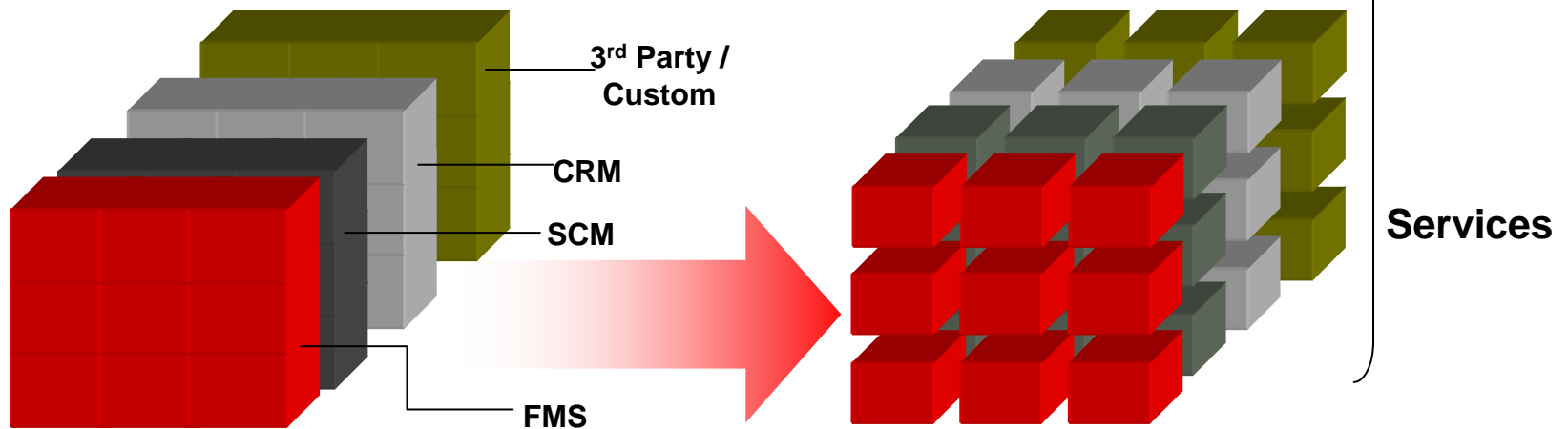


# Enterprise Business Services & Processes Challenge

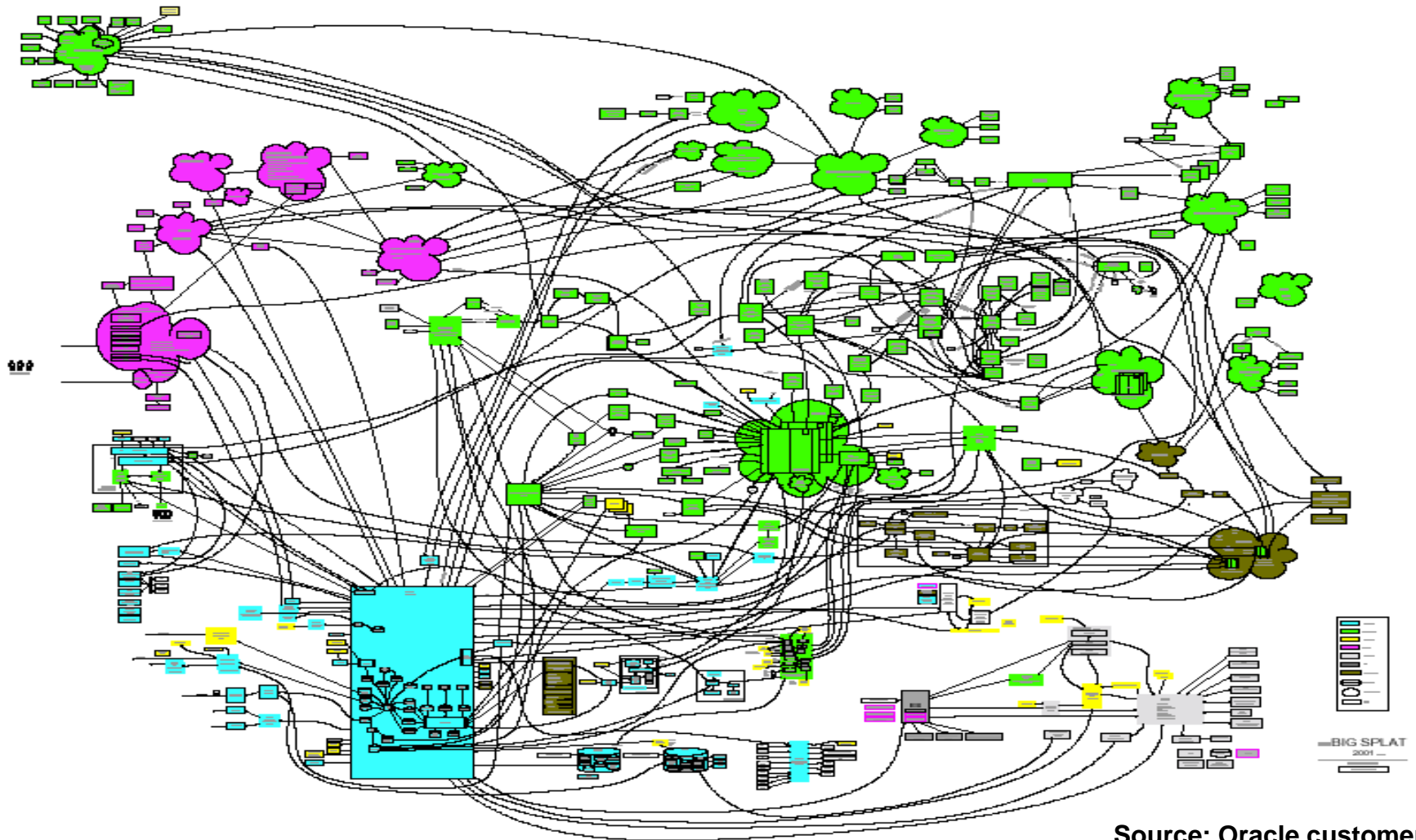
## Example: Order Fulfillment



## Monolithic Applications



# Enterprise Information Challenge



Source: Oracle customer

ORACLE

# Data Explosion Challenge

Interactions & Communications

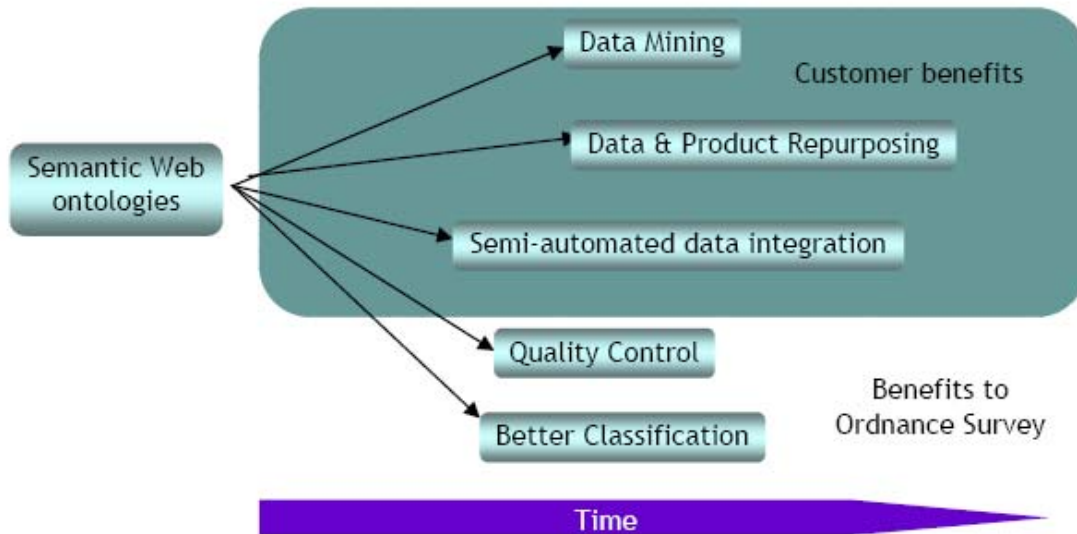
- 17 billion devices on the Internet by 2012
- Sensors everywhere = huge data volumes
- IP and telephony convergence
  - “SIP is the 3<sup>rd</sup> great protocol of the Internet” – Vint Cerf
  - Nearly 4 billion communication devices
- Networked content continues to increase dramatically
- Social computing techniques maturing
  - Wikis, blogs, RSS are just the start

2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015

Source: Adapted from IDC data 2005/6



# Use Case: Information Integration using Geographic Data



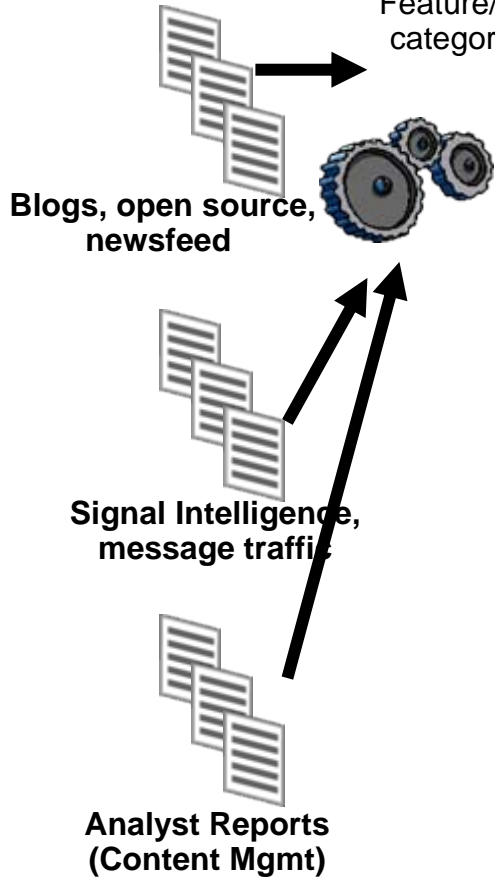
- Ordnance Survey maintains definitive mapping data of Great Britain, the world's largest and most detailed Geo DB
- Semantic Web is used to integrate different, semantically diverse sources of data

- **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**

# Use Case: Knowledge Discovery

## - National Intelligence

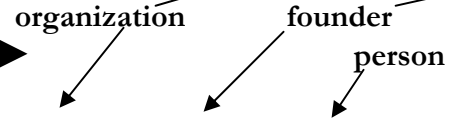
### 1. Unstructured Data (Text)



**Entity Extraction Engine**  
 Feature/term/relation Extraction, categorization (Insight, Lymba, Calais, Gate)



### 2. Model



### 3. Structured Data

“Oracle’s founder Larry Ellison wins 2010 America’s Cup Race ...”

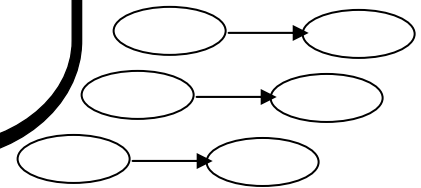


founder



XML/OWL/N3

Triple Structure:  
 Subj – Pred - Obj



10’s of billions of triples



SPARQL/SQL

### 4. Mining & Discovery



Explore

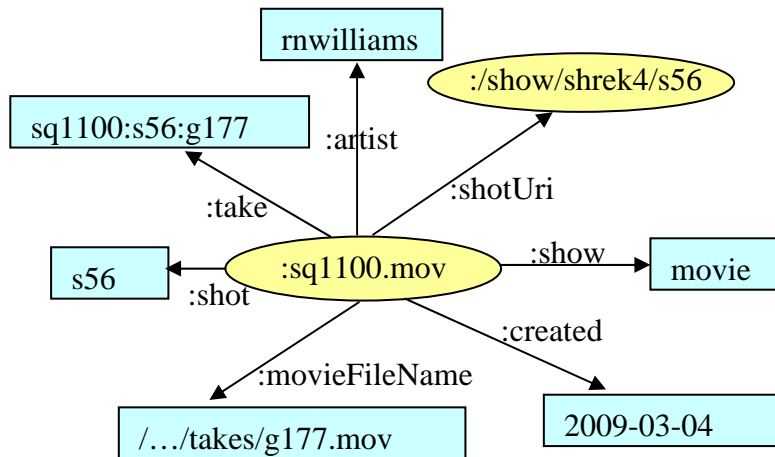


Analyst

Browsing, Presentation, Reporting, Visualization, Query Tools (e.g. i2, Centrifuge, Visual Analytics)

# Use Case: Movie Image Discovery - Dreamworks

- An example RDF graph describing a movie shot



- One RDF based application uses SPARQL to find movie information

Search

Play List

Search

Basic

Show: guardians

Sequence: sqp10

Action: a000

Artist: Search artist...

Department: Search dept...

On: MM/DD/YYYY

Before: MM/DD/YYYY

After: MM/DD/YYYY

Description: Search description...

Advanced

# Use Case: Cisco Enterprise Collaboration Platform



- Cisco's first-ever enterprise software solution
- RDF serves as metadata model for blogs, wikis, calendar, IM, WebEX, voice, and video

**Integrates vital communications and collaboration tools with everyday business applications into a single interface**

# Use Case: Content Search for Collaborative Communities

ORACLE®

The top screenshot shows the Oracle Technology Network (OTN) homepage. It features a search bar at the top right and navigation tabs for Home, Headlines, Downloads, Forums, Articles, Podcasts, Blogs, and More. A sidebar on the left lists various categories like DATABASE, MIDDLEWARE, APPLICATIONS, TECHNOLOGIES, and DATE, each with a count of items. The main content area displays a 'Welcome to Oracle Blogs' message and a list of authors.

The bottom screenshot shows a specific blog post titled 'Ready to say "Hello" to E-Bus R12'. It includes a 'Blogs Posted Over Time' bar chart showing the number of posts per day from Dec 07 to Jan 09. The post content discusses the R12 release and mentions the event 'Jan 21st 1st Feb is the day for release'. The author list includes Pankaj Chandiramani, Phil Hunt, Robert Smyth, Shay Strmelzer, Stephane Marcink, Steven Chan, Steven Dawalar, Siva, Thomas Kyle, and Tim Dexter.

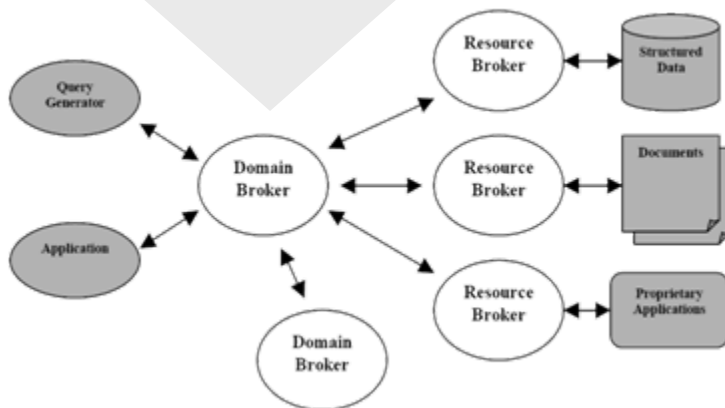
- Oracle Technology Network (OTN) aggregates many sources of content through a single portal
- Oracle's taxonomy is used for annotation of news feeds
- Semantic Web allows dynamic re-aggregation of results and leads to more comprehensible search results
- Advantages include enhanced search and navigation, and more powerful user interface

ORACLE®

# Use Case: Manufacturing and Design Knowledge Exchange - Boeing

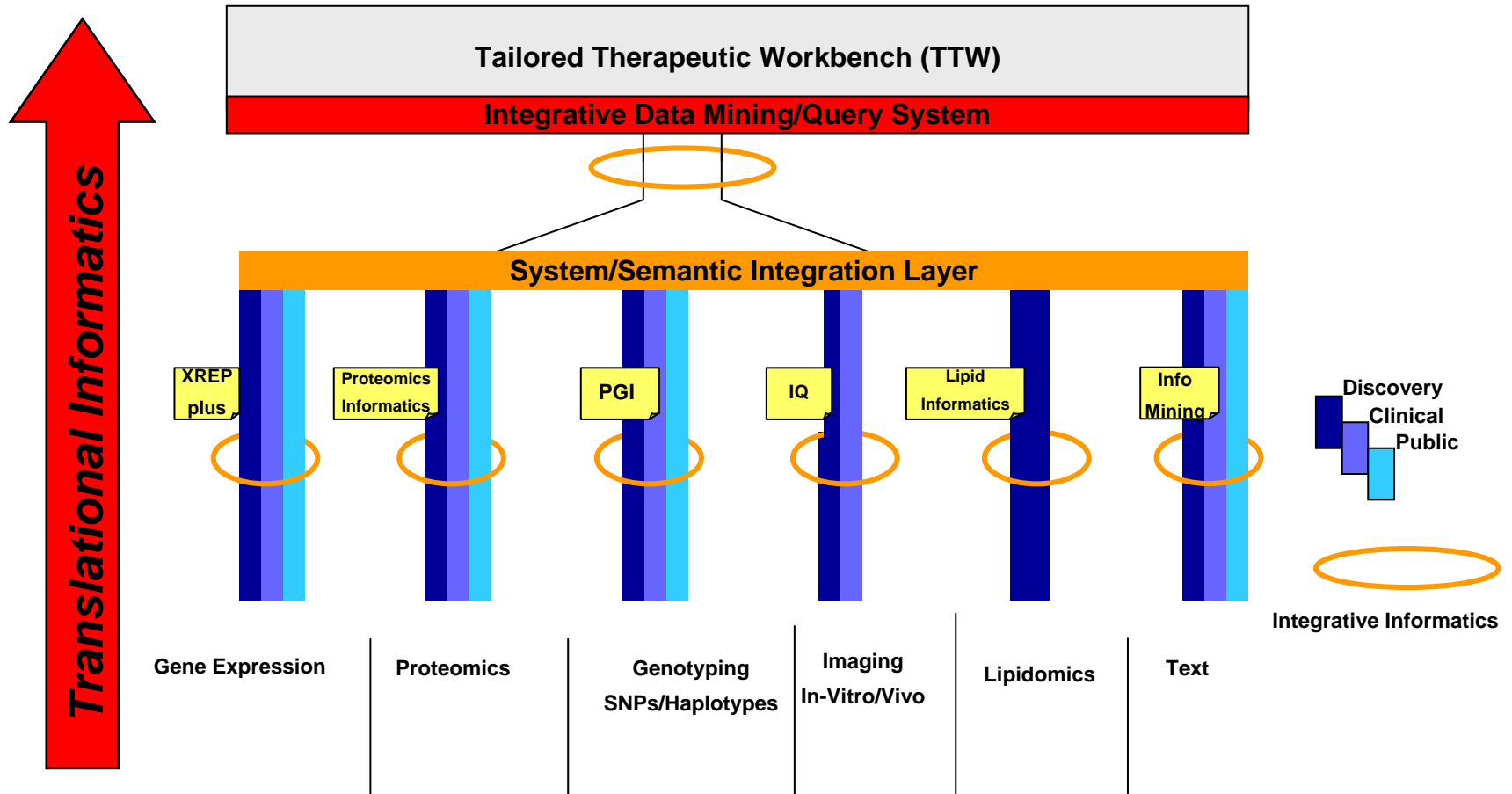
```
<db:DatabaseTable rdf:about="&LDMDEV;AIRCRAFT"  
  db:Label="AIRCRAFT">  
  <db:PartOfDB rdf:resource="&LDMDEV;LDMDEV"/>  
  <db:AssociatedClass rdf:resource="&am;CH-47"/>  
</db:DatabaseTable>  
<db:DatabaseColumn  
  rdf:about="&LDMDEV;AIRCRAFT.AC_SERIAL_NUMBER"  
  db:Label="AC_SERIAL_NUMBER">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

# Use Case: Integrative Data Mining and Ad Hoc Query

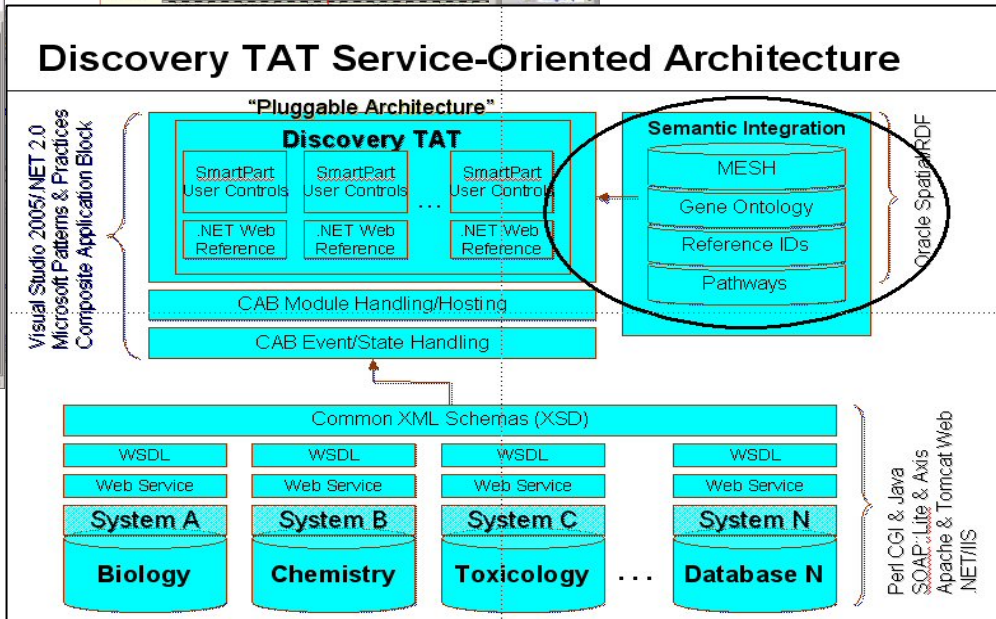




# Use Case: Drug Discovery / Target Assessment Tool

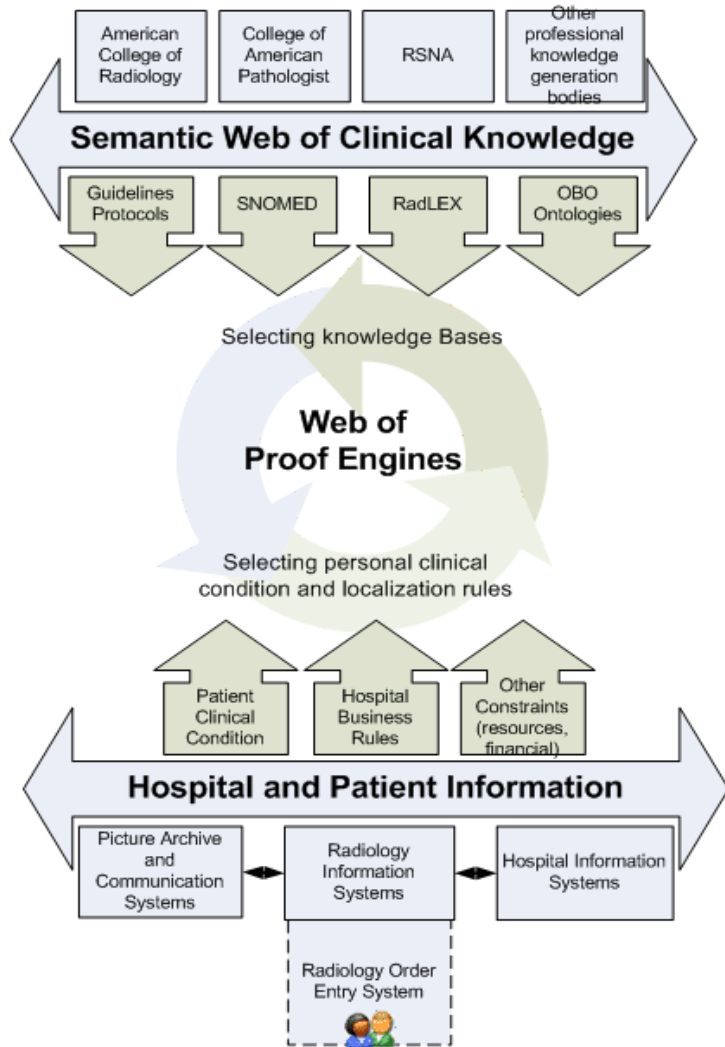


The screenshots show the Lilly Science Grid application interface. The top-left window displays search results for 'SCN5A'. The top-right window shows a detailed view of the 'SCN5A' gene, including its description, summary, and a diagram of its structure. The bottom-left window shows an ontology tree for 'SCN5A'. The bottom-right window shows a detailed view of the 'CDK1/CDC2' gene, including its description, summary, and a diagram of its structure.



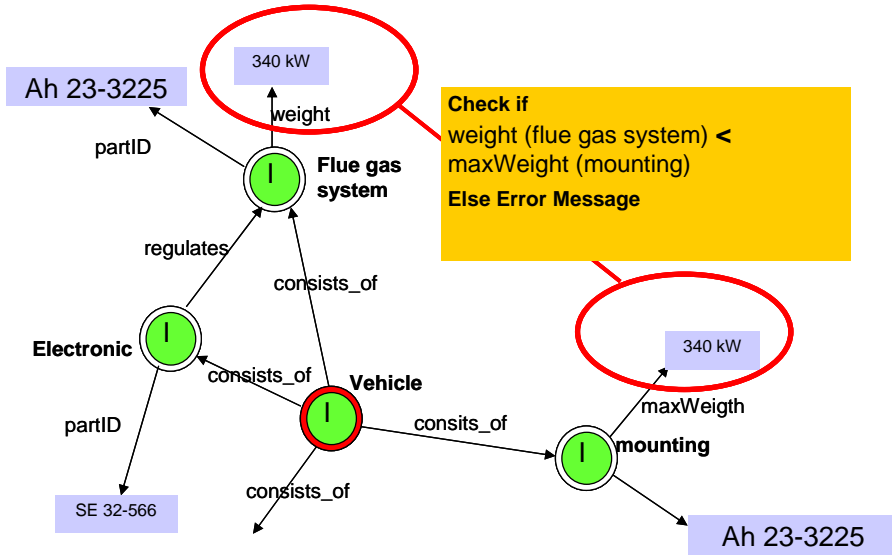


# Use Case: Clinical Health Care Decision Support - AGFA



- 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.

# Use Case: Diagnostic Tools for Complex Machinery



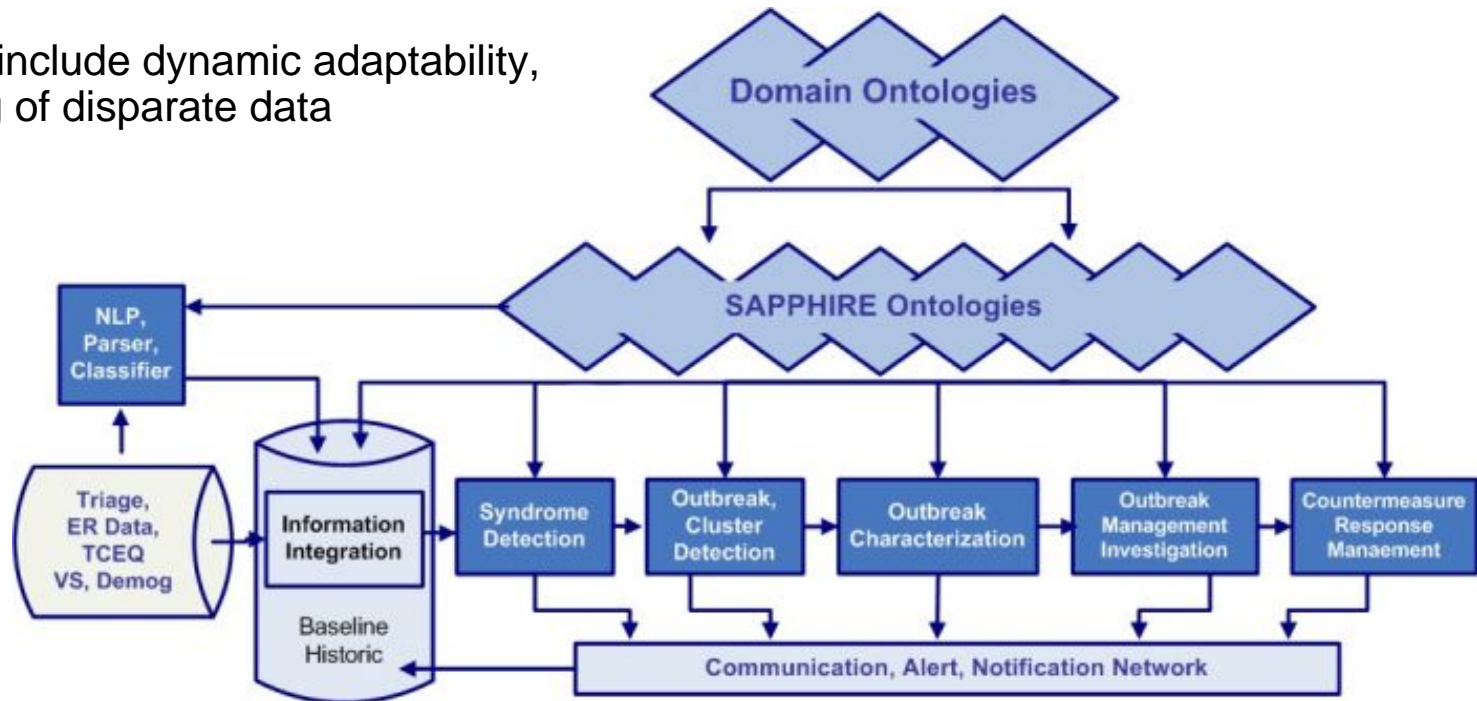
- Use of Semantic Web for generation of a diagnostic engine that can compute necessary procedures on the fly
- A question may be “What are the preliminary steps to test the resistor of the air conditioning engine on a Renault Clio 2006?”
- RDF is used as a flexible bus for exchange of information that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

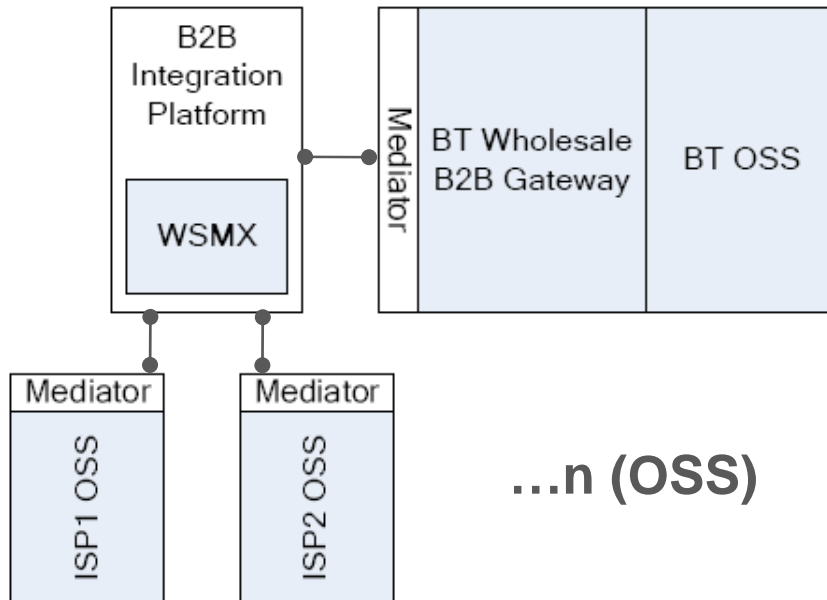
# Use Case: Public Health and Disaster Preparedness

University Of Texas

- 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



# Use Case: 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**

# Opportunity: Event Driven Architecture Enablement

## Real-time Event Processing Trend

### Real-time Data Sources



Directional Drilling



Equipment Monitoring



Flow Rates



Market Feeds

### Immediate Automatic Responses



Display on dashboard



SMS Message

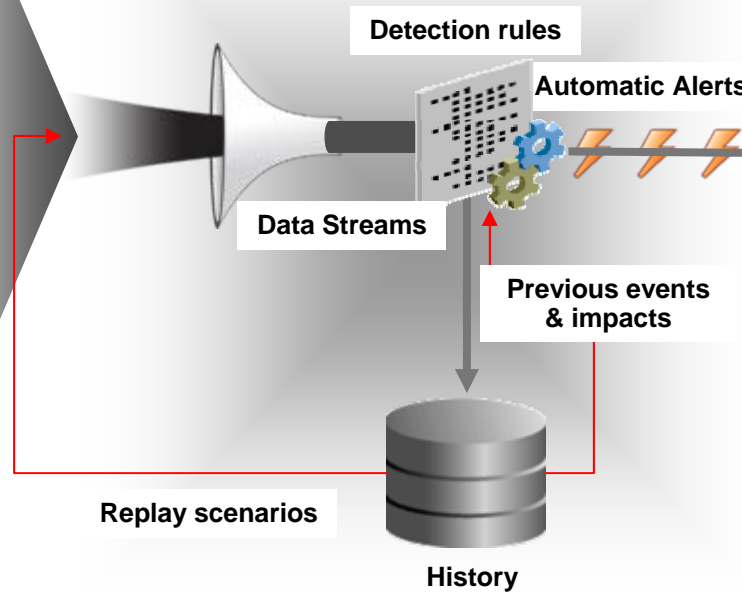


Initiate Workflow



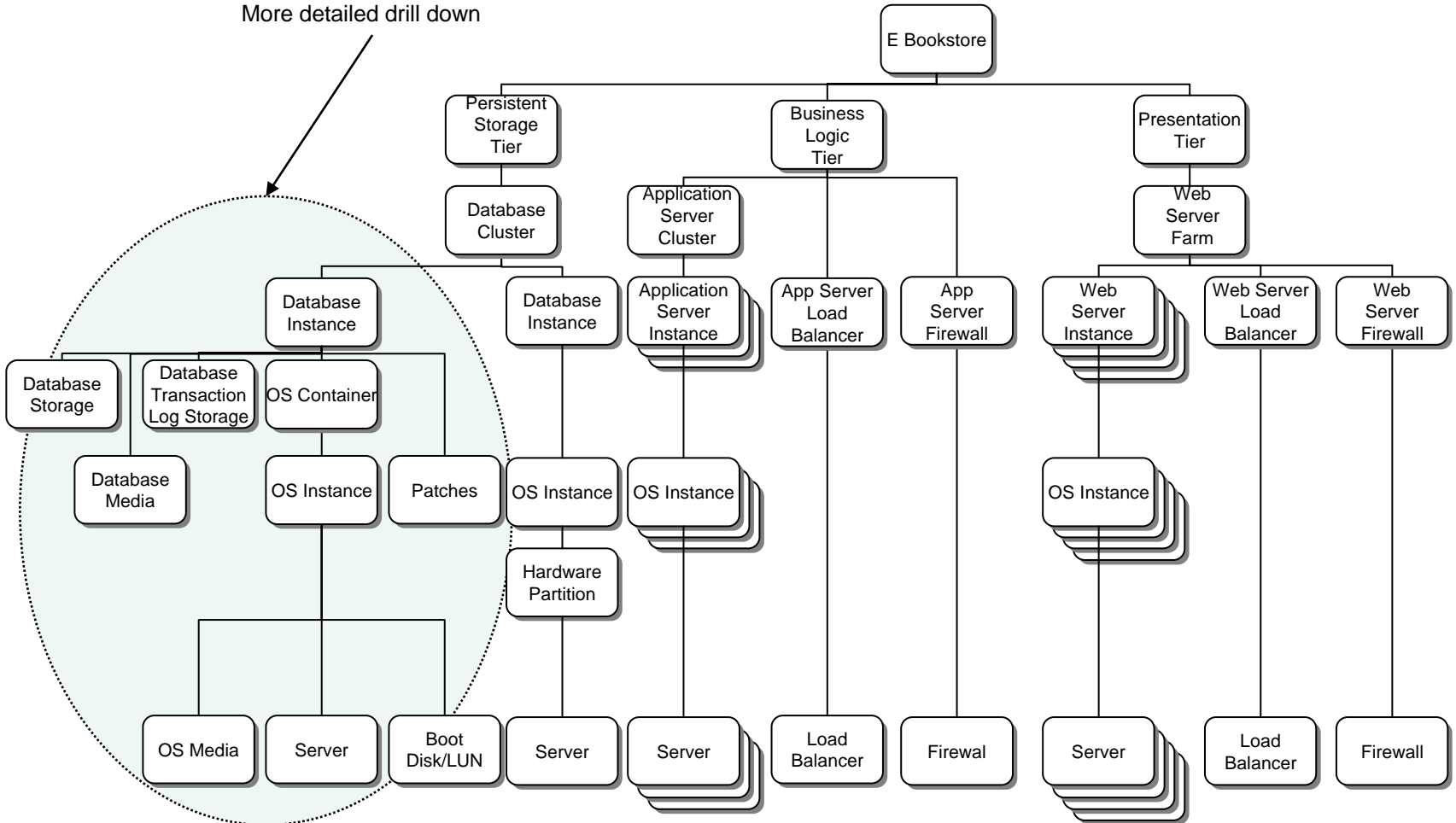
Buy/Sell

### Complex Event Processing



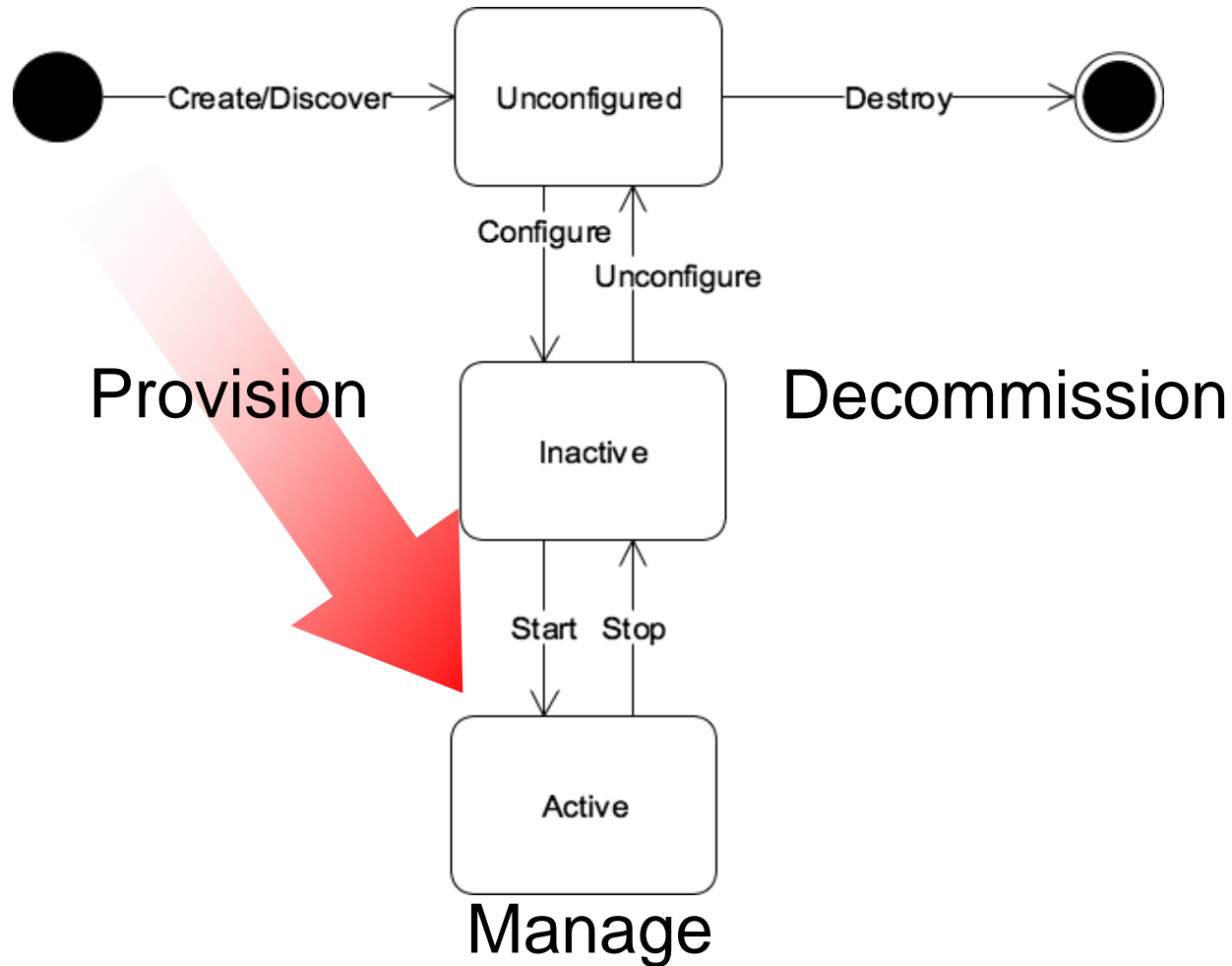
# Opportunity: Cloud Service Enablement

More detailed drill down



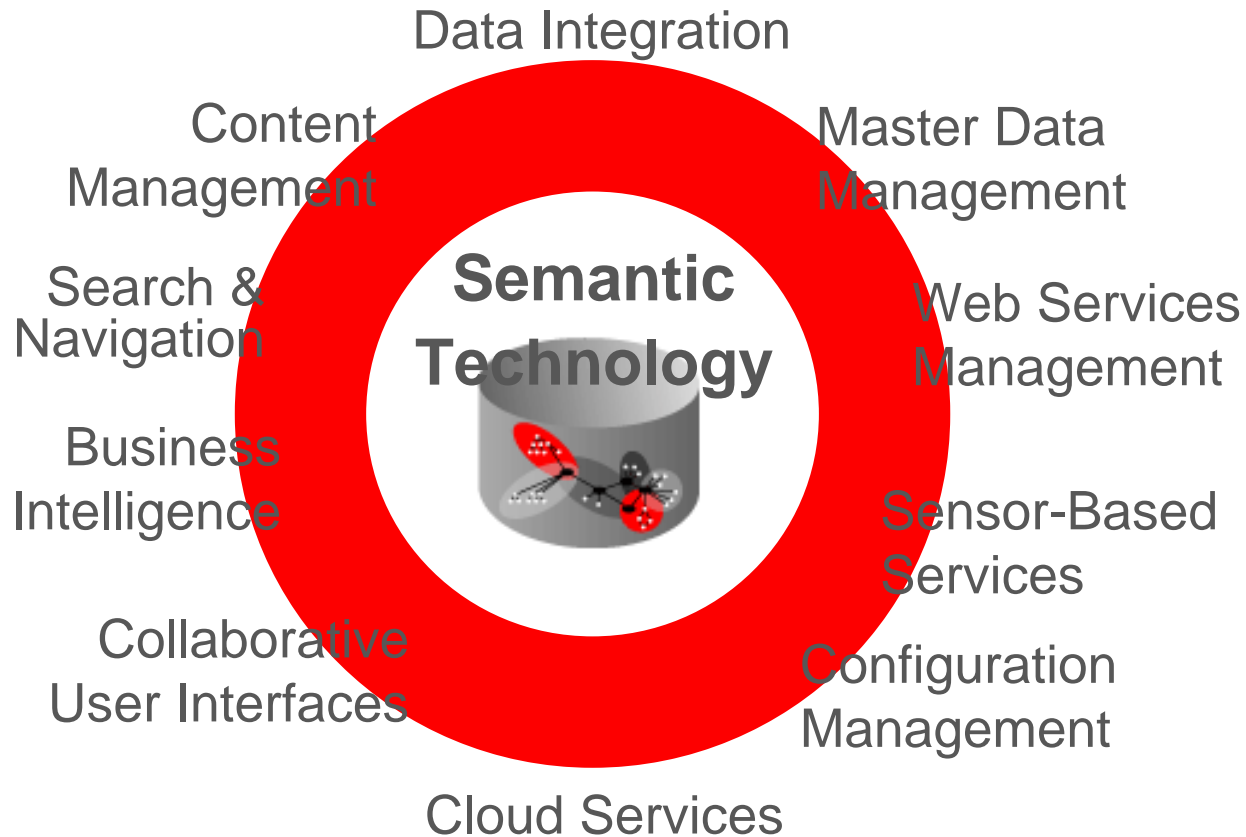
**EGA Reference Model**

# Opportunity: Cloud Service Enablement



EGA Reference Model

# Summary





# Conclusions

- Key semantic technologies are mature
- Semantic technologies are key enablers for Enterprise-wide solutions
- Useful data and ontologies exist but more are needed
- Model the real world rather than data artifacts
- A little semantics goes a long way
- Build upon infrastructure incrementally