Chevron

Enterprise Architecture

Chevron Position Paper for W3C Workshop on Semantic Web in Oil & Gas Industry

Frank Chum, ITC EA
Mario Casetta, ETC IM
Roger Cutler, ITC EA

9 December 2008
Houston, Texas

Agenda

- Background & Key Business Drivers
- Problem Domains
- Some Semantic Web Technology Activities in Chevron
- Some Technical Lessons Learned
- Key Challenges
- Conclusions
- Q&A

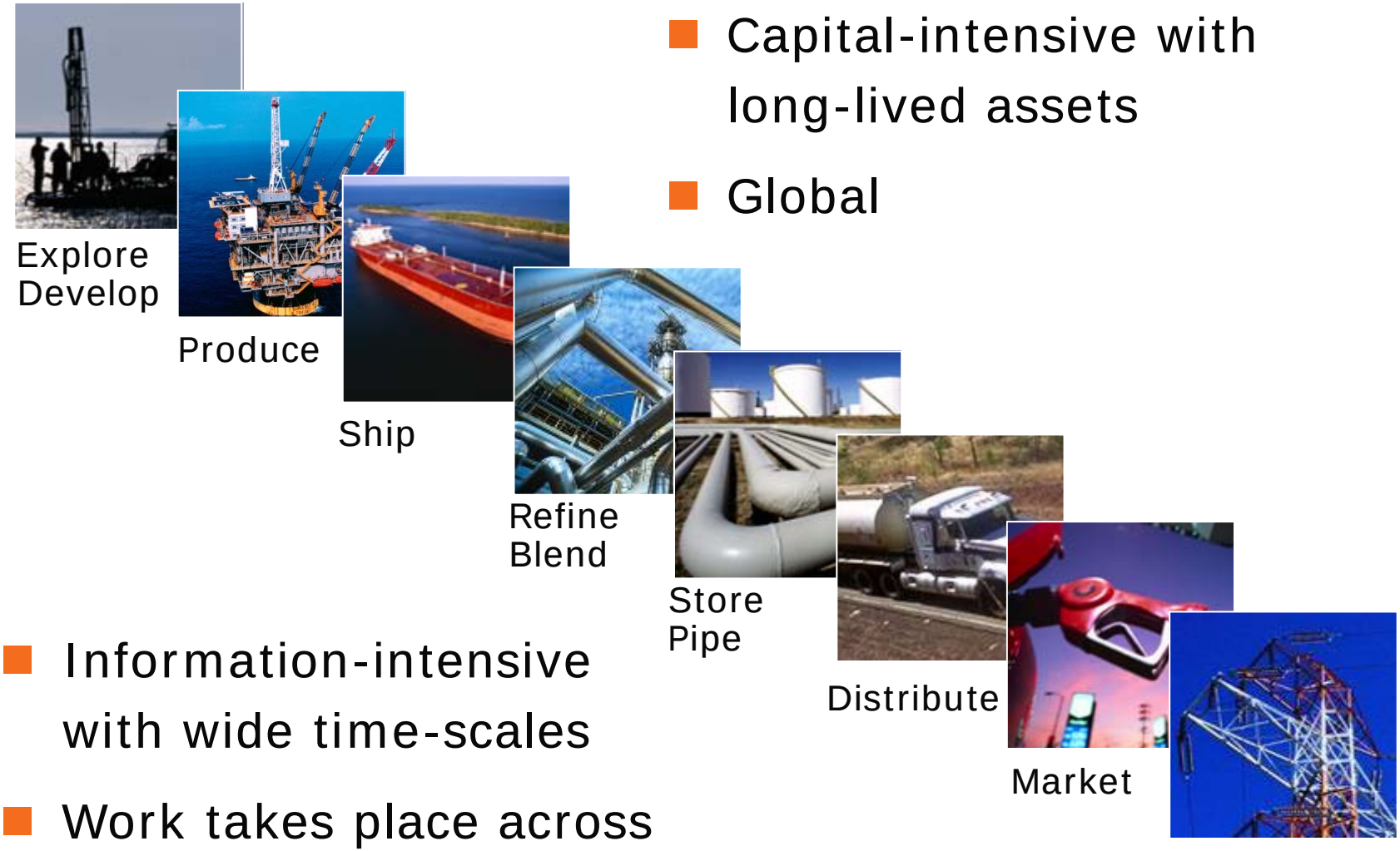
Background

- Semantic Web Technologies Scan and Assess since 2004

- Joined W3C in 2005
  - "Observer" to the Health Care and Life Sciences Interest Group (HCLS-IG) in its formation Nov. 2005 with ~40-50 members
  - Chevron representative member to the W3C Advisory Committee
  - Oct. 2006 - active member of the Semantic Web Education and Outreach Interest Group (SWEO-IG)
  - April. 2007 - Published an Ontology-based Information Integration and Delivery Use Case for O&G Industry on W3C SWEO-IG site

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

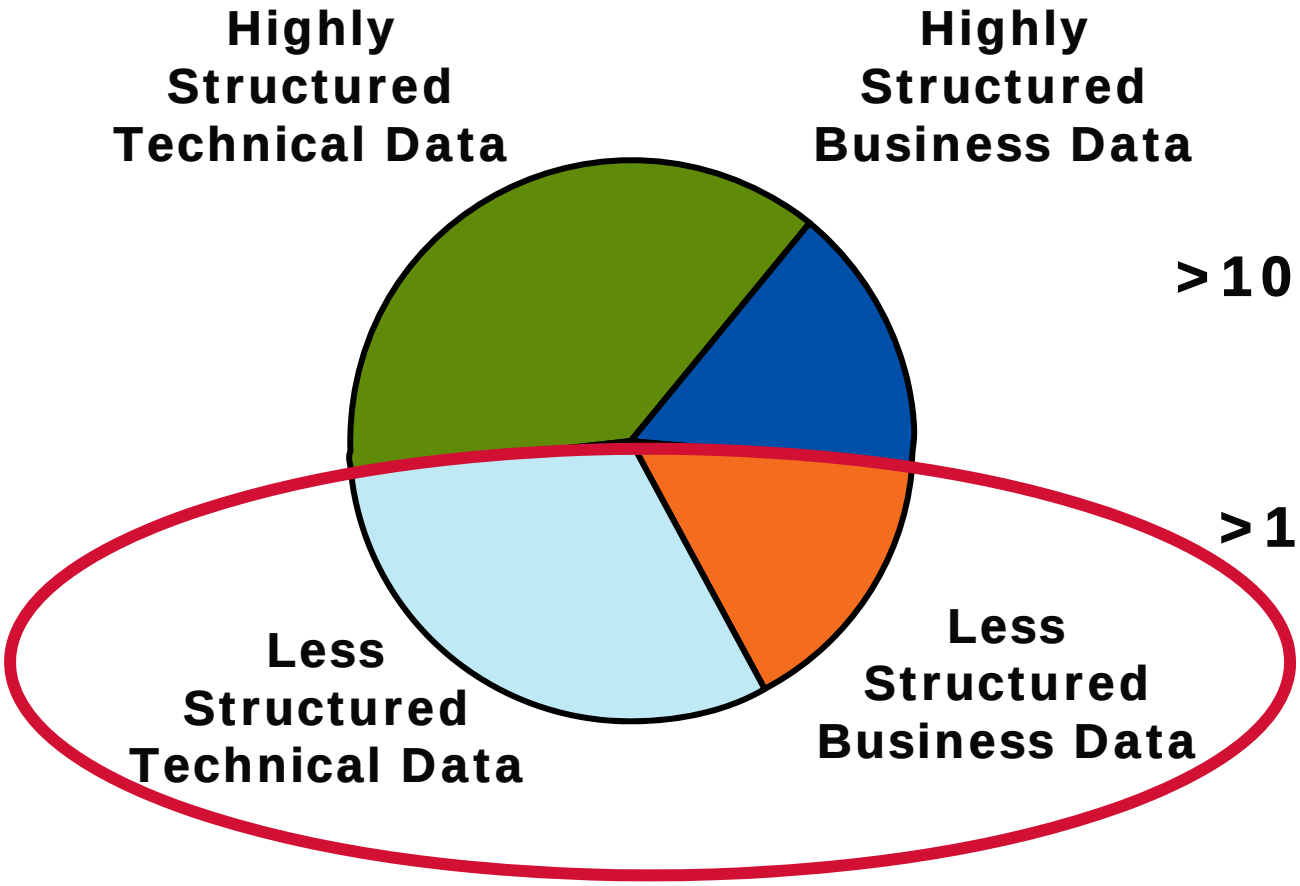
Key Business Drivers: Information is so Critical



- Capital-intensive with long-lived assets
- Global

- Information-intensive with wide time-scales
- Work takes place across geographies

Key Business Drivers: 6000 TB of Data



>100 thousand items

>100 million items

- Data volumes are doubling every 6-12 months (50 HIS clients survey, Berkley research)
- Growing impact of information volumes on employee productivity
- Growing external compliance and risk elements associated with information

Key Business Drivers: The “Big Waves”

- Huge challenge in retaining and transferring knowledge due to the “crew change”
- Globalization continues – service suppliers, consumer market growth in Asia, etc.
- People and data continuously on the move

Problem Domains

- Semantic reconciliation of enterprise metadata
  - Provides an organized approach to metadata management by resolving differences in meaning in order to enhance metadata shareability and interoperability.
- Standardization for information exchange between enterprise and business partners
  - Standardizations are needed for semantic reconciliation of definitions and specifications across corporate boundary.
- Information integration and delivery
  - Provides application interoperability by connecting information from highly diverse sources and having a shared, common understanding of the data to facilitate enterprise application integration.

Some Semantic Web technology activities in Chevron

- ETC Reservoir Management and Production Engineering – Integrated Asset Management (IAM)
  - Ram Soma, Amol Bakshi, et. al. presentation
- Drilling & Production knowledge management/data exchange
  - Lee Feigenbaum, Cambridge Semantics presentation
- Major Capital Projects Operational Systems (MCPOS) – ISO 15926 Ontology and Reference Data Library – MCP Facilities Engineering
  - Project with Bentley Systems and Fiatech
- ETC exploratory pilot on the Unix file system
  - See next slides

A Semantic Web Exploratory Pilot

ITC/ETC Collaborative Strategic Research Project
ITC Enterprise Architecture and ETC Technical Computing

Are Semantic Web Technologies mature enough to be useful in CVX at this time or in the near future?

Proof of concept:
Link technical data and document data using semantic technologies.

ETC Sub-projects:
Can a Semantic data store be used to collect and store metadata for our technical data?

Can we develop a semantic model of our technical data?

ETC Exploratory Pilot

Can we create and put metadata concerning our technical data in an RDF store?

Can we create a semantic model of our technical data?

Process:

1. What questions would we like to answer?
2. What metadata can we gather?

Available Metadata
File Paths and Permissions
SeisWorks groups
Project names

Subset of Unix data for testing:
Unix directories
SeisWorks projects
Gocad++ projects

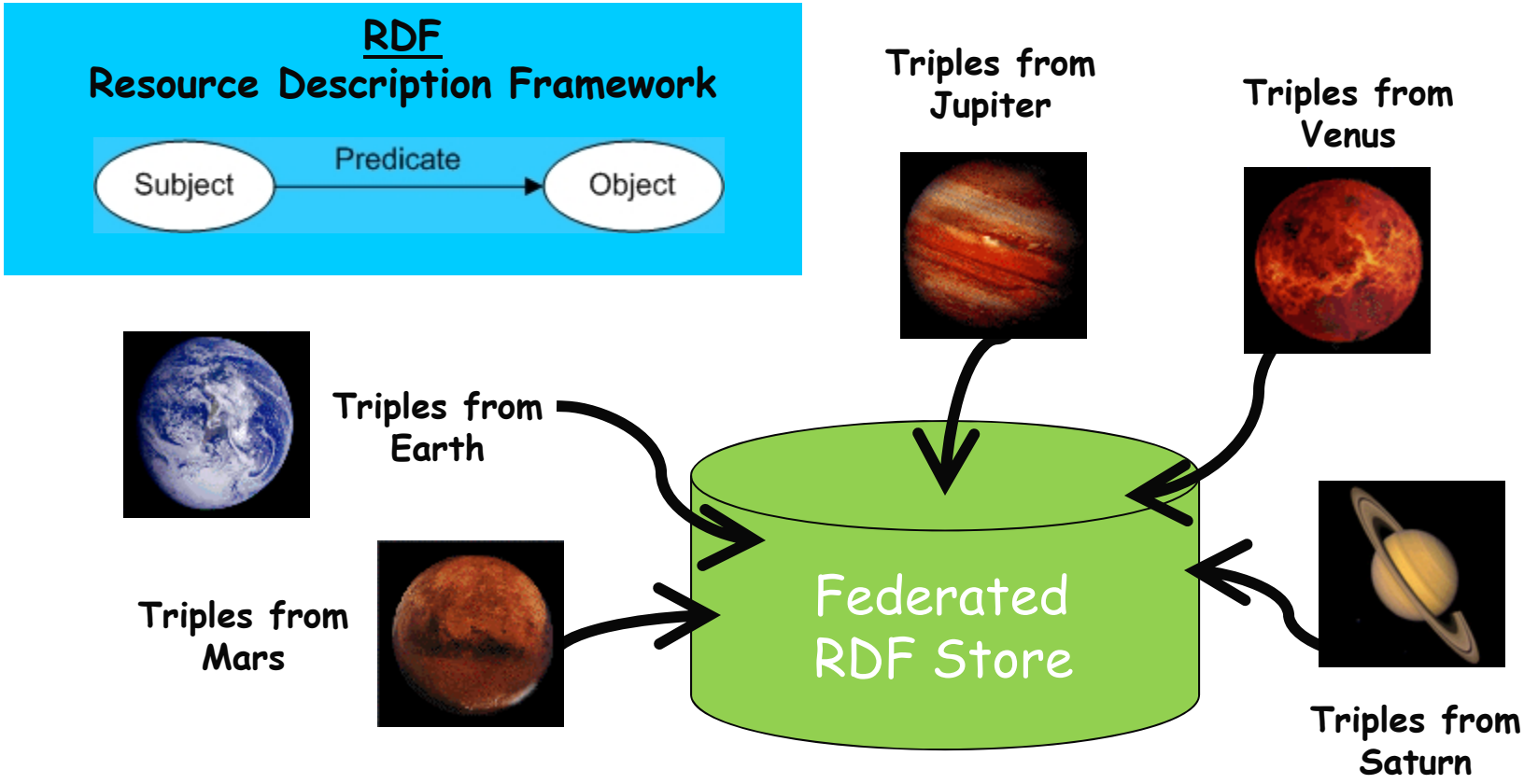
Identify Classes - "subject"

Identify Properties - "predicates"

Develop Ontology

Generate RDF Store

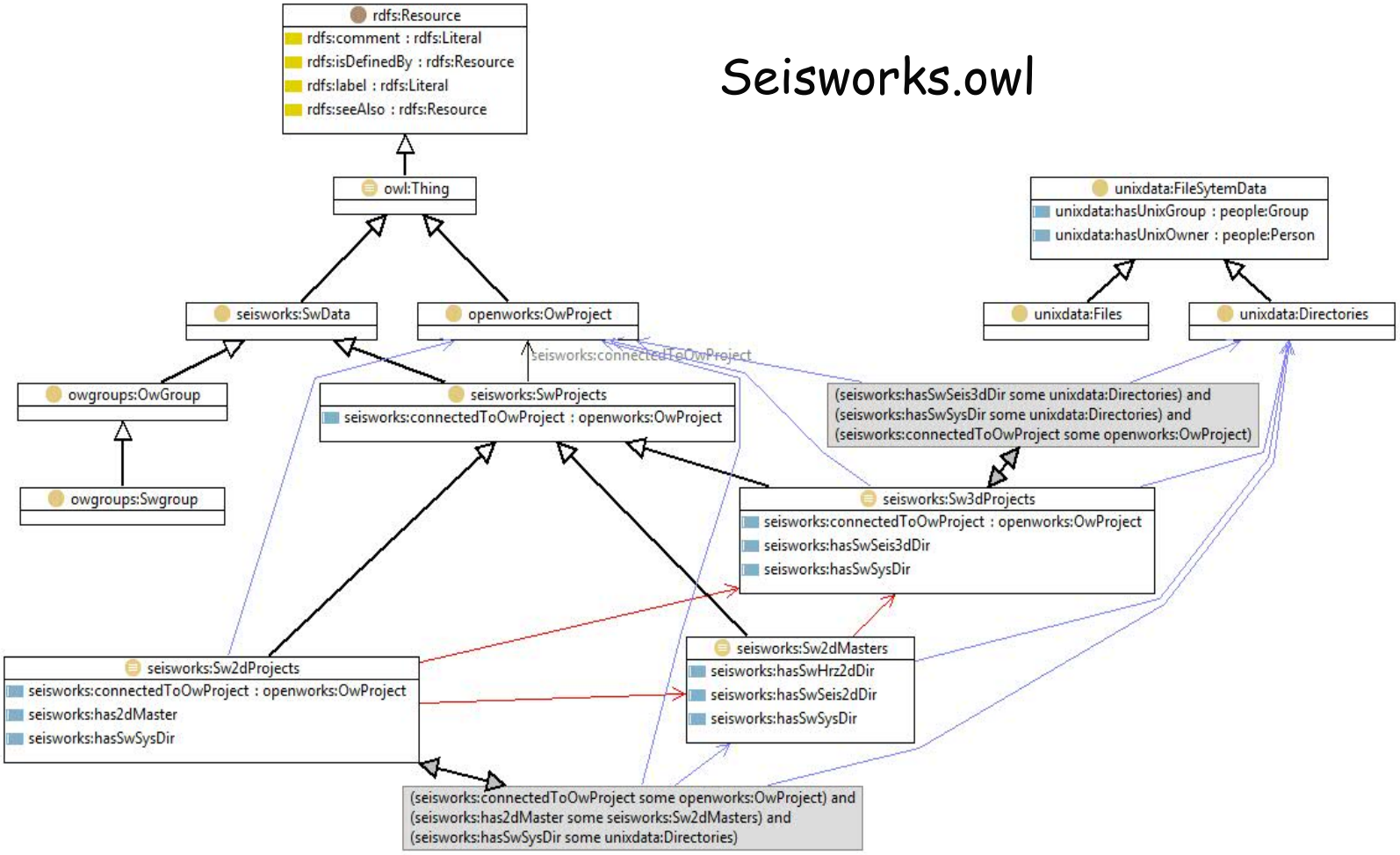
Resource Description Framework



Assertion: Standardized core parts of the Semantic Web can be leveraged to enhance Information Management Functions.

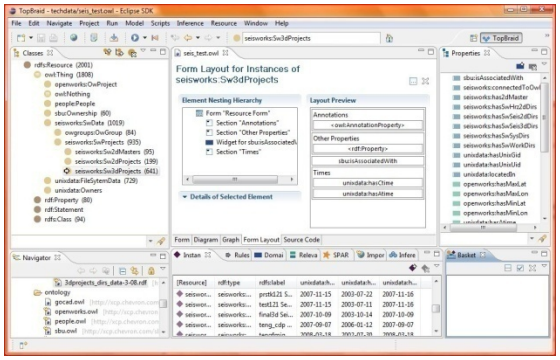
Ontology

An Ontology is a way of describing things in a domain, their properties, and their relationships.



Technical Data Metadata Store

Ontology

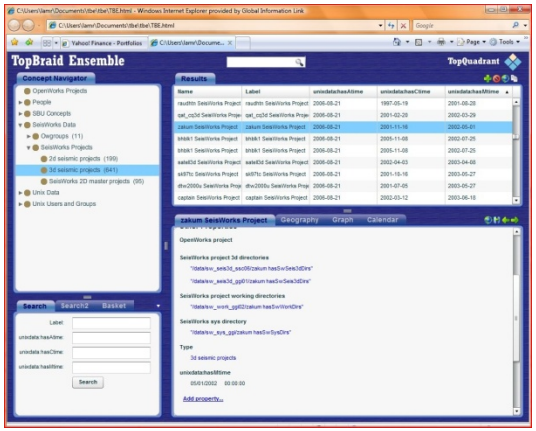


Analytic Scripts

- Technical Project Metadata
- Unix data directory metadata
- User, Business Unit, Geo-information

Oracle 11g Semantic Store with Inference

Browser User Interface



Technical Lessons Learned

- URI's (Uniform Resource Identifier) need to be unique and repeatable
  - Development of appropriate URI system is critical
  - Requires semantics in crawling scripts
  - Standards facilitate the process of knowledge discovery.

- Develop ontology in a logical modular fashion.
  - Get ontology training or consulting.
  - Test modules at each step.

- One tool doesn't fit all
  - Many Semantic Web technologies and tools available

Extending the use case?

Potential use cases for bridging data from different sources and formats for functions such as:

- Search of technical data
- Life Cycle management of technical data
- Archive systems
- Tie technical data to other project data such as financial, legal, reserves…

Next steps

- Form Partnerships to continue research and "proof-of-concept" projects.

- Continue to develop and refine ontologies and merge with other ontology development efforts.

- Develop demos and use cases of things we can do with the technology that are difficult or impossible otherwise.

Some Key Challenges

- Domain ontologies can be complex and require major commitments from SMEs to build by hand. Automated ontology construction approaches can be helpful.
- Ontology management can also be complex as the knowledge base continues to grow.
- How to use open ontologies? Where does intellectual property start and end?
- How to promote an information sharing mindset and rationalize to a common, shared ontology?
- Context based mining and automatic extraction of metadata from structured and unstructured data.

Conclusions…

- Ask the following questions before considering Semantic Web standards and technologies
  - What does the Semantic Web bring to the table that cannot be solved by traditional technologies?
  - How does automated inference help solve the business problems?
  - Where does the needed metadata come from?
    - Up front effort is usually needed to classify, categorize, tag, and extract meaningful metadata for semantic processing.

Conclusions, and some discussion questions…

- We think that semantic technologies are important but real progress is going to require an industry specific interest group (IG) for us to make tangible progress with critical mass in this space
- To that end, some discussion questions to consider at the wrap up of this Workshop
  - Will there be enough interest to form an IG?
  - What is required to form the IG?
  - What is the charter for the IG? What are some of the objectives we should work towards?
  - What would be the advantages/disadvantages of an IG?
  - What are the alternatives involving industry consortia (e.g., Engeristics, POSC/Caesar, etc.)
  - What will take us to come to consensus on IG formation?
  - What are the next steps?

Q & A