Semantics in the Financial Industry: The EDM Council Semantics Repository Progress Report

Mike Bennett
Hypercube Limited
89 Worship Street, London EC2A 2BF
www.hypercube.co.uk
Summary

• The Challenge
• The Framework
• Extending the Framework
• History of the Project
• Current Activities
• The Future
EDM Council Repository requirements

• The EDM Council
  “A non-profit trade association focussed on managing and leveraging enterprise data as a strategic asset to enable financial institutions to increase efficiency, minimize risk, and create competitive advantage”

• Industry Requirements
  – Consistent terminology
  – Mapping of terms in disparate systems
  – Standard agreed set of terms
  – Focus on meaning not words

• Business management of data
  = business ownership of data semantics
Ontology modeling Framework

- **UML Tool**: Enterprise Architect from Sparx Systems
  - Reports - processed into spreadsheets
  - Diagrams direct from tool
- **ODM (early draft)**
  - Adapted to natural language
  - Some tool specific adaptations
- **Spreadsheet framework**
  - Features for (most) OWL concepts
  - But in English
    - `owl:ObjectProperty = Relationship Fact`
    - `owl:DatatypeProperty = Simple Fact`
- **Diagrams**
  - “Boxes and lines” (no visible language)
  - Various levels of detail
Financial Instruments Semantic Model

Overview

- Financial Instruments
  - Common Instruments
  - Equities
  - Debt
  - Rights
  - Traded Options
  - Futures
  - OTC Derivatives
  - CIV (Funds)
  - Indices and Indicators

- Components
- Dated Terms
- Process Terms
- Basic Types
- Global Terms
  - Time
  - Geographical
  - Mathematical
  - Information
  - Business
  - Financial
  - Legal
  - Activity
  - Risk

- Prepared by:

- Classes of security which are Bonds a sub-class of Bond.
- Thing

- "Is A" relation

- Object Property
  (Relationship Fact in English)

- Multiple Inheritance
The Semantics Repository

Top level: KR Lattice hierarchy
philosophical order
continuant v occurrent
concrete v abstract

Mid level: Global terms
Accounting
Legal
Math etc.

EDMC: Financial Instruments
Semantics
Securities
Components
Market data
Processes

Common types and selection lists

Instruments
Date Varying Terms
process etc.

KR Lattice
Global terms
How the meaning gets in

- Written definitions against each term
  - Reviewed and agreed by domain SMEs
- Hierarchy (AKA taxonomy) showing where each thing fits into the world
  - What kind of thing is it?
- Relationships and other facts about those things
  - What facts distinguish this from other things?
- Explained to SMEs in set theory terms
Archetypes

• The simplest of kind of thing that each financial securities concept is a specialized type of
  – Contracts, Events, Accounting constructs
  – These have common necessary facts
• Represented using UML stereotypes
  – Color coded
  – Graphics for some
• These form common models such as accounting, legal, countries
  – Ultimately these would be replaced with industry-derived sets of standard terms
  – Accounting: XBRL
The Story So Far

• Initial design: May ’08 – Sept ’08
• Format review panel – finalized formats of spreadsheets, diagrams and web layout
• Roadshow
• Initial draft: Jan ’09 based on ISO terms
  – Tradable Securities: Static reference terms
  – Weekly SME Reviews to July ’09
• Draft released July ’09
  – Terms marked up with review status
  – Out for further review and validation
Current Status

- **Static Terms Draft**
  - Need final validation / QA of model
  - Validate in a real world context

- **Proof of Concept Exercises**
  - Securitization (MBS issuance)
  - Data Vendor
  - Risk analytics
  - Banks etc.

- **Date-dependent terms**
  - SME Reviews ongoing since July ‘09

- **OTC Derivatives, corporate events to follow**
Proof of Concept Findings

- Securitization (MBS Issuance) PoC
  - ECB, NY Fed, IBM Research
  - Demonstrate ability to tag new instruments semantically at issue
    - Plans to make this mandatory
    - Basis for systemic risk regulation

- Findings
  - The domain experts get it
  - Many terms not in ISO standards
    - Will feed these into ISO 20022
    - Refining these with domain experts
  - Complete view of poorly understood securities and missing data linkages (sub prime etc.)
Possible Applications

• Exploring in Proof of Concept exercises
• Simplest: Replace “ad hoc” business spreadsheets in business conceptual “layer”
• Model driven development
  – In UML
  – Can we derive semantic LDM equivalent?
  – How to relate to database models?
• Derive native OWL/RDFS
  – Not included in project
  – Would enable OWL apps / SPARQL etc.
• Metadata repository deployment?
Ontology Format Extension

- Things that are not in scope of OWL itself
  - Synonym
  - Archetypes
  - Classification facets
- Hope to collaborate on standardized use of OWL Annotation Properties
- Can classification facets be used to derive single-taxonomy data models?
The Future

• Finalize content and release formal version
• Formalize governance
• Track semantics standards evolution (OWL, ODM)
• Align upper ontology with semantics industry standards
• ISO Alignment
  – Alignment of content with ISO 20022 Logical Data Model
  – ISO 20022 version 2 semantics layer
    • work with WG4 on model standard
    • update the core modeling concepts in line with this

• **Objective:** Move from a working prototype model framework to something more standard while contributing our model concepts to industry
Thanks

Mike Bennett: mbennett [at] hypercube.co.uk

EDM Council:

www.edmcouncil.org

Semantics Repository:

www.hypercube.co.uk/edmcouncil