ENTERPRISE DATA MANAGEMENT

EDM COUNCIL

MIT Information Quality Industry Symposium

"The Building Blocks of Data Management"

Michael Atkin, Managing Director, EDM Council

Massachusetts Institute of Technology

July 14-16, 2010

The Data Benefit of Crisis

- The credit crisis, economic meltdown and the rise of the concept of systemic risk analysis is great news for data management
- Forced participants to think about data content as the foundational building blocks for complex analysis
- Data content management is a relatively new discipline for the financial industry
- We're good at processing data, but not experienced at managing the meaning of data where granularity, precision and comparability matter
- The "financial industry" meets the "information industry"

Systemic Analysis is Not a Thing

- There is no standard report for identifying market vulnerabilities
  - Assess the interrelationships and interdependencies of factors to determine the impact of specific events on the overall financial system
  - Look at the nature of risk concentrations, monitor unusual escalations in asset prices or variances in credit spreads, etc.
- Understand the legal characteristics of instruments and portfolios
  - Understand all the obligations and restrictions related to trading, managing, clearing, settlement and custody
  - Comply with the terms of indenture, redemption, maturity, priority and exercise
  - Understand the terms and parameters for payment and reinvestment
- Determine the links and relationships between issuers, issues, instruments and obligations in order to know what is being traded with whom and under what conditions

Necessary Conditions (building blocks)

- Instrument Reference Data (factual)
  - Identifiers at the instrument level
  - Standard language associated with the legal and contractual structure of financial instruments (including corporate actions/maintenance)
  - Classification of instruments types and cash flow processing
- Entity Reference Data (factual)
  - Identifiers for all involved in the transactions chain and where due diligence is required under any jurisdiction
  - Ownership structures and business relationships throughout the transactions process
- Pricing and Valuation Data (interpretative)
  - No problem for exchange traded
  - Exposure and comparability of calculation methodologies and all derived inputs for off exchange and thinly traded
- Positions and Transactions Data (sensitive)
  - Summary of settled and unsettled contractual agreements where financial entity has a legally binding role
  - The goal is comparability among factors of input into positions and transactions

This Is What Is Required

- Identify financial instruments and legal entities with precision (ISO-based standard identifiers)
- Establish a common language for all contractual obligations that drive our industry (ontology and semantic precision)
- Manage our fragmented chain of data supply and shift from ‘data scrubbing’ to ‘data manufacturing’ mindset (legal compulsion)
- Overcome ignorance, arrogance, obsolescence and power (the four horsemen of the EDM apocalypse)

The Objectives are Achievable

- Instrument identification and the historical lesson from the 1968 paper reconciliation crisis (we've done this before)
- Entity identification is a direct parallel except only 800,000 entities versus 8 million financial instruments (we can do this again)
- Stop thinking about semantics as an IT problem. Start thinking about it as a simply a formal and factual representation of reality (remember Aristotle)
- Majority of reference data needed for business processing and regulatory oversight are based in legal contracts (tag them at the point of issuance using industry standards)
- The one aspect of regulatory reform that is common to everyone around the world is DATA. Global harmonization of data standards is logical and inevitable (and the financial industry would be forever grateful)

Semantics in Context

Tagging of meaning is the key to knowledge management

- (1963) GML – describe format, organizational structure and content components
- (1986) SGML – ISO standard for describing a documents structure and attributes = rise of online information industry
- (1991) HTML – Tim Berners-Lee for document sharing = rise of dynamic web formatting
- (1996) XML – abbreviated version of SGML allowed for separation of components of a document (logical/physical structures, elements/attributes, processing instructions) = communication of precise meaning and structural relationships

Data Attribute Tagging

- First rule of complex systems is to get the language straight
  - Prime directive – data that users have confidence in to be fit for purpose without reconciliation
  - Different ‘data dialects’ are the enemy of automation
  - Our industry is dominated by common words with different meaning, common meaning using different words, important nuances not captured
- Semantic precision is essential in order to compare data, set precedence rules, achieve STP, promote confidence in analytical models, structure portfolio strategies and perform cross asset risk assessment

Semantics Repository

- Two year project to standardize terms and definitions for all reference data attributes (common ontology)
  - Formal and factual representation of reality (things, facts about things and relationships among things)
  - 5000 terms and definitions (created and verified by industry practitioners)
  - Covers common instrument categories, equities, bonds, structured finance, money markets, entitlement rights instruments, options, futures, CIV, indices, OTC derivatives, dated terms (market data) and terms related to issuance
- Relationship to other semantic activity
  - ISO 20022 (financial information business model
  - XBRL (accounting and business reporting)
  - DTCC (corporate actions), EFAMA (funds), FpML (derivatives)

Bottom Line

- Rare opportunity to implement a functional (standards-based) data infrastructure and truly fix the fragmented chain of supply for reference data
- Pathway is clear and the objectives are viable
- Pieces of puzzle coming together – let's not waste the power of the "perfect storm"
- **Pay attention**. The principles of data and foundations of data management are in the spotlight with a growing sense of urgency – but not fully sorted out. The goal is practical and additive to the way business operates in reality

Contact Information

Michael Atkin

Managing Director, EDM Council

301.933.2945

atkin@edmcouncil.org

web: www.edmcouncil.org

Semantics Repository

www.hypercube.co.uk/edmcouncil