Why Harmonized Data Matters
Semantics & Inference Processing in Finance

14th International Semantic Web Conference
October 13, 2015 • Bethlehem, PA
Michael Atkin, Managing Director, EDM Council
SESSION AGENDA

Status of Activity in the Financial Industry

Financial Industry Business Ontology

Requirements for Adoption
• Lessons from scientific, technical, legal, medical and scholarly publishing

• Three discoveries that changed the knowledge representation equation
Banks are the New Publishers

and must adopt the principles of the information industry in order to meet regulatory reporting obligations, support model-based strategies and automate linked processes
THREE TENETS OF THE BCBS CONTROL ENVIRONMENT

Tenet 1: The Governance Mandate

Objective Direction: Establish a fully documented, appropriately resourced data framework, with no organizational barriers and top-of-the-house engagement

BCBS Observations

- Formal and documented frameworks missing
- No defined ownership or clear demarcation of responsibility
- Lack of coordination among business, IT and risk
- Decentralized and undocumented data policies
- SLAs and measurement criteria for risk data processes needed
- Higher standard for audit of risk data aggregation needed

You Ain’t Got No Choice
Tenet 2: The Data Infrastructure Mandate

**Objective Direction:** Build integrated data architecture (taxonomies, metadata, identifiers). Establish controls across the full data lifecycle; create flexible classification and aggregation schemes; support on-demand, ad-hoc reporting and scenario-based analysis.

**BCBS Observations:**
- Inconsistent taxonomies
- Inconsistent metadata, identifiers and dictionaries
- Inability to harmonize, integrate and compare among repositories
- Data identification and CDEs needed
- Failure to understand interdependencies between linked processes

Data Harmonization is Mandatory
Tenet 3: The Data Quality Mandate

Implications: [Risk] data must be timely, accurate and comprehensive; must adopt authoritative sources and the creation of a data control environment; need to align data to “concepts” for consistency of meaning across the organization; must be able to generate timely [risk] reporting across all dimensions of quality and all risk categories.

BCBS Observations:
- Too much reliance on manual processes
- Insufficient data reconciliation (root cause analysis and executable business rules)
- A need for better control across lifecycle of data (data inventory, transformation mapping, cross-referencing, authoritative sources)

Data must be Fit for Purpose
## Why Harmonized (Common Language) Data Matters

<table>
<thead>
<tr>
<th>Instrument ID</th>
<th>Classification Risk</th>
<th>Interest rate risk</th>
<th>FX risk</th>
<th>Spread risk</th>
<th>Leverage risk</th>
<th>Depreciation risk</th>
<th>Correlation risk</th>
<th>Redemption risk</th>
<th>Liquidity risk</th>
<th>Volatility risk</th>
<th>Counterparty risk</th>
<th>Valuation risk</th>
<th>Execution risk</th>
<th>Settlement risk</th>
<th>Custody risk</th>
<th>Exposure risk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Issue versus instrument identification, instruments without codes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Security Classification</td>
<td>Convertible, subordinated, guaranteed, unsecured, preferred, senior, rate-based, backed, index-linked, floating, etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instrument Descriptors</td>
<td>Coupon rates, payment dates, effective dates, termination dates, call/put feature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fundamentals</td>
<td>Earnings ratios, market capitalization, dividend rates, capital distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terms &amp; Conditions</td>
<td>Conversion rates and dates, sinking fund provisions, maturity dates, reset dates, redemption schemes and dates, restrictions, eligibility, benchmarks, day count conventions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prices</td>
<td>Historical, close, time series, volatilities, correlations, NAV, VWAP, yields, durations, OAS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legal Entities</td>
<td>Ownership, control, affiliates, role performed, identification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research</td>
<td>Analyst reports, reporting dates, credit ratings, categorization, credit enhancement support</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obligations</td>
<td>Priority, exercise, debenture, guarantee, obligation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructions</td>
<td>Standing settlement instructions, transfer agents, tax instructions, regulatory jurisdictions, calendar, currency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corporate Actions</td>
<td>Dividends, mergers, acquisitions, tenders, IPOs, corporate restructure, class actions, consolidations, payments, splits, repayments, liquidation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
WHY HARMONIZED (COMMON LANGUAGE) DATA MATTERS

### Instruments
- Identification
- Classification
- Description (rates, dates, features, schemes, provisions)
- Value (i.e. price, date, time)
- Calculate (volatility, correlation, yield, duration)
- Maintain (corporate actions)

### Entities
- Entity type (legal persons, formal organizations, corporations, partnerships, affiliates, trusts, functional, etc.)
- Ownership structures
- Controlling relationships

### Obligations
- Issuance process
- Trade and execution
- Guarantee
- Allocate and administer
- Clear and settle
- Transfer

### Holdings
- Firm portfolio (individual entity risk)
- Corporate structure (organizational risk)
- Industry wide (systemic risk)

Degree of interconnectedness
- Transitive relationship
- State contingent cash flow
- Collateral flow
- Degree of centricity
- Funding durability
- Leverage & liquidity
- Guarantee & transmission of risk
- Degree of diversification
Core Regulatory Initiatives

- **Dodd-Frank Act Title IV, VII, X, XIV** = US framework for regulation of swaps markets, hedge funds, CFPB, mortgage, Volker
- **EMIR** – European Market Infrastructure Regulation (EU version of Dodd-Frank Title VII on derivatives transparency).
- **Regulation AB2** = regulations on asset backed securities (unravel links between loan, tranches, pool, etc.).
- **FATCA** = individual reporting of foreign accounts and FSI reporting of foreign financial accounts about US clients
- **UCITS** = Undertakings for Collective Investment in Transferable Securities (EU Directive on simplification of prospectus and their expression using clear, accessible and standardized data).
- **AIFMD** – Alternative Investment Fund Managers Directive (EU proposed law to provide more oversight and transparency to hedge funds and private equity).

- **Dodd-Frank Act Title I** = the financial stability component (creates Financial Stability Oversight Council and OFR)
- **EU System of Financial Supervision** = establishment of the European Systemic Risk Board (and ESFS)
- **Basel Principles for Effective Risk Data Aggregation and Reporting** = implementation of a “data control environment” and healthy “risk appetite framework” within systemically important financial institutions
- **Basel III** – global regulatory standard on bank capital adequacy, stress testing and market liquidity risk.
- **CCAR** = Comprehensive Capital Analysis and Review (stress test methodology in the US; CCAR reporting is putting lots of pressure on data alignment and comparability. This includes the FR Y-9C (Bank Holding Company Capital Report) and FR Y-14Q (detailed ‘show your calculation methodology work for BHC). This is the US version of Basel III.
- **Solvency II** – EU Directive that harmonizes insurance regulation (requirements for capital reserve and reduction of risk of insolvency) – to be implemented January 2014.

- **MiFID II** – Revised Markets in Financial Instruments Directive (mostly about trading, but does require common instrument identification for consolidated pricing).
- **ACORD** – Insurance standards development body (UK) likely to be mandated as the format for reporting.
- **Regulation SCI** – SEC proposed Regulation Systems Compliance and Integrity (to ensure that core infrastructure is functional)
- **COREP** = Common Reporting requirements (developed by Committee of European Banking Supervisors (CEBS) with the goal of developing a supervisory reporting framework based on common data standards and formats.
- **FSB Templates** = Common Data Template for G-SIB’s seeking to harmonize the data compounding methodology for reporting.
THE DAUNTING REGULATORY CONNECTION

**BCBS 239**
Mandatory implementation of control environment to support scenario-based analysis

**Systemic Risk**
We aren't getting what we need (i.e. derivatives) and requirements are clear

**Financial Instrument DB**
Our conclusion is “ontology” and “identification” to support analysis of state contingent cash flow

**Legal Entity Identification**
The identifier is underway but the goal is understanding counterparty risk and the degrees of interconnectedness
STANDARDS-BASED INFRASTRUCTURE

Pathway to all Objectives Starts with Implementation of the Standards-Based Infrastructure

Copyright © 2015 EDM Council Inc.
DATA MANAGEMENT REALITY…

• This stuff is Hard to Do
  Lots of repositories, lots of functions, difficult to get people to play nice in the sandbox

• Hard to Sell
  Short-term/spreadsheet-based mentality, expensive, disruptive, not well understood, hard to measure, few serious wins, ROI hard to define

• Hard to Implement
  Execution capability lagging, legacy spaghetti and silo barriers, project-based mentality

• Hard to Orchestrate
  Social and political culture, interconnected activity, IT barriers, data factory complexity

• Takes a Long Time
  “Unless you are helping me with something real, you are either an annoyance or a problem”

• Skill Sets Are Missing
  Need combination of IT knowledge, product and process expertise, modeling capability, legal obligation understanding, diplomacy, project management skills

• Hard to Sustain
  “Big hairy process,” ownership and accountability missing, little executive air cover, lack of sustainable funding, broad scope of data quality problems to manage
Legacy Environment
Existing technical and operational environments
- Disparate data sets
- Proprietary interfaces and point-to-point links
- Multiple repositories managed independently
- Inconsistent formats and data definitions

Compared to

Simple
Reconciliation of complex environments
- Designation of “authorized data domains”
- Definition of critical data elements (CDEs)
- Documentation of end-to-end data flows (compounding process, derived calculations, risk and business formulas)

Simplification
①
②
③
④

Controlled Data Environment
Governed by policy, sanctioned by executive management, based on standards, harmonized across the lifecycle, with clear accountability and monitored by audit

Manage Data Quality
Fit-for-purpose data without reconciliation and transformation
- Data quality criteria (all relevant dimensions)
- Establish data quality control points
- Quality assessment and remediation (current state analysis)
- Define business rules, thresholds and tolerances
- Root cause analysis (trace to source)
- Management of data manufacturing chain of supply

Alignment to Meaning
Harmonized and precise data based on contractual precision
- Attribute level business glossary/business conceptual ontology
- Unique identification and flexible classification scheme
- Harmonization and transformation processes (cross-referencing and mapping across systems)
- Metadata repository (administrative, structural, descriptive)

Technical Implementation
Integrate data into operational and production environments
- Platform and authorized tool stack
- Messaging and distribution infrastructure
- Semantic to logical data model
- Logical to physical instantiation
- Location identifiers and namespace management

Scope of Work Required for Control
BUSINESS VALUE TO BE REALIZED

Operational Efficiency
- Reconciliation & manual processes
- Trade repairs & fails
- Settlement instruction mismatches
- Corporate actions processing
- Redundant systems
- Capital reserve/unproductive capital
- Transformation & integration challenges

Flexible Data Infrastructure
- Classification and aggregation capabilities
- Flexible queries
- Data harmonization
- Adaptability
- Extensibility
- Structural data quality

Sphere of Value

Commercial Insight
- Upselling opportunities
- Enhanced customer service
- New product engineering
- Product & client ROI
- Sentiment and contextual analysis

Trust & Confidence
- Risk analysis & reporting
- Compliance requirements
- Derived data, calculations and analytical models
- Business process automation
- Management & financial reporting
Addressing the Challenges of Data Management must be built upon a foundation of industry best practices.
2015 BENCHMARKING METHODOLOGY

Global Participation


Control Group of Data Management Professionals and CDOs
Foundational Levels of Governance are in the Process of Being Established

Seed funding in place with executive air cover and authority

Data owners hired and tasked to address gaps and challenges

Office of Data Management is an official function with defined processes
Changing Organizational Behavior is Difficult

- Data stewardship and accountability being defined
- Data policy in process (policy mandates compliance)
- Business buy-in tentative (potential haircut looming)
Unique Identification and Precise Meaning are the Keys to Data Management

Unraveling lineage, locating data and managing complexity overwhelming

CDEs defined - but not fully inventoried or aligned with compounding processes

Harmonization to meaning across thousands of repositories a daunting challenge
Data Quality Control Procedures are Needed to Ensure Trust and Confidence in Data

Data quality control processes (business rules, authorizations) in concept stage

Tactical “find” and “fix” is still the approach for data reconciliation

Profiling and current state assessment not very advanced
Key Areas of Data Management

1. Establish the Data Management Program (strategy, communication, authority)

2. Commitment from Stakeholders (alignment, funding, metrics, resources, ecosystem)

3. Implement Governance (organizational structure, stewards, policy, adherence)

4. Content Infrastructure (meaning, CDEs, logical domains, lineage)

5. Manage Data Quality (profiling, control points, root cause)

6. Collaborate with IT (alignment on goals, capabilities exist)
What a long, strange trip it's been

FIBO
Financial Industry Business Ontology
Closing Price is Dependent on Circumstance and Objective
Forget the Words – Focus on Meaning
WE HAD TO INVENT A LOT OF THINGS

How to Model Concepts with Business

Grounded in Law
Defined by Contract

How to Express Contractual Issues

How to Effectively Collaborate

[Logos for JIRA, Confluence, Jenkins, and GitHub]
Business Conceptual Ontology

- Formal and factual model of financial industry business/contractual concepts (ontology of the business structures and legal obligations)
- Expressed as a glossary of terms and definitions plus all business relationships (what “things” are and how they work in the real world)

Primary FIBO Domains

- **Market Data (time and date) Semantics** – Pricing and Yields, Analytics, Temporal Terms, Credit Rating
- **Process Related Semantics** – Securities Issuance, Corporate Actions, Transactions, Payment Processing, Portfolio and Holdings
EVOLUTION OF FIBO – SORTING OUT THE PIECES

Legacy FIBO
Baseline work and underlying model

Development Methodology
Governance and collaboration process

Underlying Framework
The essential building blocks (FIBO scaffolding)

FIBO Vocabulary
Unambiguous shared meaning (classification)

Standard Ontology
Inference processing in RDF/OWL

Messaging Collaboration
Alignment with XML schemas

Web Applications
Financial concepts for the Internet based on FIBO
STATUS OF FIBO DEVELOPMENT – UNDERLYING FRAMEWORK

FIBO-Foundations

- Basic building blocks (not unique to financial services)
- Concept of “contracts,” “transactions,” “processes,” “units of measure,” “accounting,” “schedules” (concepts continue to be added)

Financial Business & Commerce

- Common building blocks across all classes of instruments (things that are unique to financial services)
- Types of contracts, concepts of classification, types of participants in the financial system, payment terms, schedules, etc.

Business Entities

- Everything needed to unwind business relationships (legal perspective)
- Concept of “ownership,” “control,” “role,” “function,” “legal capacity,” etc.

Indices/Indicators

- Everything needed to understand based on rates and benchmarks
- Rates (interest, stock, foreign exchange) and economic indicators

Pricing, Analytics and Processes

- The concepts of “time” and “date” (including areas related to pricing, duration, yields, past and present value, volatility, etc.)
## STATUS OF FIBO INSTRUMENT VOCABULARY

Much of FIBO is modeled (with SME sign-off) and ready to be used as common language for classification, lineage remediation and harmonization to meaning

<table>
<thead>
<tr>
<th>Status</th>
<th>Type</th>
<th>PRIORITY</th>
<th>FCT underway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substantive</td>
<td>Equities and securities</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Debt (including bonds, structured finance and money markets)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rights and warrants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete</td>
<td>Market data and pricing (including analytics and credit ratings/status)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Corporate actions and events</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>Loans (including mortgage, general purpose, construction, student)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Derivatives (including rate-based, credit default, FX, asset, commodity, CFD and exchange traded)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Collective investment vehicles and funds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initiated</td>
<td>Issuance (including equity, debt, asset-backed and mortgage-backed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transactions (includes trade, post trade, clearing and settlement)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Payment processing and holdings</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
30 Core Ontologies

Upper Ontology

FIBO Foundations: High level abstractions

FIBO Business Entities
FIBO Financial Business and Commerce
FIBO Indices and Indicators

FIBO Contract Ontologies
- Securities (Common, Equities)
- Derivatives
- Funds
- Securities (Debt)
- Loans, Mortgage Loans
- Rights and Warrants

FIBO Pricing and Analytics (time-sensitive concepts)
Pricing, Yields, Analytics per instrument class

FIBO Process
Corporate Actions, Securities Issuance and Securitization

Future FIBO: Portfolios, Positions etc.
Concepts relating to individual institutions, reporting requirements etc.
**Interest Rate / Business Entity POC**

- Full FIBO implementation using State Street interest rate and business entity data
- Mapping from State Street data repositories and FpML messaging to the FIBO-IR and FIBO-BE ontologies (with link to LEI)
- Provide inference processing for full scope of transactions (at scale)
- Link to data visualization software for analytical display

**POC Objectives**

- Demonstrate the mapping process
- Classify and aggregate based on any fact about instrument or entity
- Provide automated structural validation
- Demonstrate ownership and control relationships
- Query on country, size, notional amount for each counterparty, by currency, etc.
- Identify and display degree of centrality
- Identify all counterparty and transitive relationships
- Demonstrate ease of writing queries in Anzo and SPARQL
The Federal Reserve Board’s ("FRB") Regulation W (Transactions Between Member Banks and their Affiliates) implements Sections 23A and 23B of the Federal Reserve Act ("FRA").

The Objective of Regulation W
- Defines rules about transactions between banks and their affiliates
- Defines what is covered, capital and collateral requirements, limits, portfolio restrictions
- Focuses on ownership, affiliation, credit relationships and allocation amounts

The Tasks for Regulation W
- Identify whether counterparties are "affiliates"
- Determine whether the transaction is "covered" by Reg W
- Determine if the transaction is within specified "parameters"
- Determine "am I in compliance" (why/why not)
DEMYSTIFY DATA MANAGEMENT
COMPLETE THE VOCABULARY

To get started, please search for a term here:

contract

Term: contract

Term information:
a voluntary, deliberate, and legally binding agreement between two or more competent parties

* Screen shot SmartLogic Semaphore Ontology review tool
COMPLETE THE VOCABULARY

To get started, please search for a term here:

contract

Term: contract

* Screen shot SmartLogic Semaphore Ontology review tool
SEPARETE “VOCABULARY” FROM “ONTOLOGY OF THE CONTRACT”
ALIGNMENT ON APPROACH
“Rules are built on Facts, Facts are built on Terms.”

- Base Business Definitions & Rules on Fact Types
- Associate Concepts to define Fact Types
- Define Concepts

Develop Vocabularies and Rules (to represent them starting with terms for the concepts)

COLLABORATE

Public

Private

{JSON}
CONTACT POINTS

Michael Atkin  
Managing Director  
301.933.2945  
(m) 240.602.8390  
atkin@edmcouncil.org

John Bottega  
Senior Advisor  
CDO Forum | Data Management Practice  
908.510.3826  
jbottega@edmcouncil.org

Dennis Wisnosky  
Senior Advisor & FIBO Lead  
630.240.6910  
dwisnosky@edmcouncil.org

Mike Meriton  
Head of Commercial Services  
908.720.7044  
mmeriton@edmcouncil.org

Carole Mahoney  
Business Operations  
+44 208 399 7189  
cmahoney@edmcouncil.org