Using Data Fabrics to Drive Regulatory Compliances, Optimize Business Performance, & Increase Profitability in Retail Banking

A conversation with

Steve Wyer
Founder & CEO
UNIFYDATA

Jinal Shah
Enterprise Architect
Pentacube, Inc.

Michael Bartlett
Dir., Financial Services
Cambridge Semantics, Inc.
Moderated by Mike Meriton
Co-Founder & COO, EDM Council

• Joined EDM Council full-time 2015 to lead Industry Engagement
• EDM Council Co-Founder & First Chairman (2005-2007)
• EDM Council Finance Board Chair (2007-2015)
• Former CEO GoldenSource (2002-2015)
• Former Executive for D&B Software and Oracle
• FinTech Innovation Lab – Executive Mentor (2011 – Present)
Jinal Shah
Business & Data Transformation Consultant

- Award-winning enterprise architect & change agent with 29+ years experience
- Front-to-back experience across major global & regional banks, including: Sales & Trading (cross asset), Relationship Banking, Coverage & Sectors, Corporate Finance, Financial Crime Compliance, Risk Management, Regulatory reporting, Finance & Operations
- Work Experience include: NatWest Markets, Deutsche, ABN AMRO, RBS, Lloyds, Nordea, HSBC
- Startup include Pentacube, Regulativ.ai
Data fabrics – why are they an important component of any data strategy?

Data Challenges and Drivers:

• Silo’d data
• Approved and unapproved copies of data
• No defined data lineage & data cataloguing
• Poor understanding of data and all its relationships
• Too many localised transformations taking place through the data journey
• Plethora of EUC’s that managed offline local data mappings
• Poor application & systems integrations
• Inconsistent data formats different across asset classes and within classes
• Inconsistent reference data being used

Takeaways:

• Focus on breaking/bridging data silos, adopt tools like Graph DBs
• Connect business users to the context of data with Semantic description
• The data fabric enables an organisation to transform the data journey, remove EUCs & localised transformations
• Introduce use of industry standard models to help you understand your data, add context and consistency
• Re-enforce the need for and use of data governance tools and processes.

And finally, connect your data to your organisations business landscape.
UNIFY your data – Data Connected Organisations

What is a Data Connected Organisation?

1. Senior Management have ‘line of sight’ to business outcomes..
2. .. By using industry business models (e.g. BIAN, FIBO) that describes the ‘domain’ specific business landscape..
3. .. which provides an understanding of the causal dependencies between data, processes, technology applications and change activities
4. .. to enable the implementation of the appropriate analytics & operational systems to achieve your strategic objectives

The value of a Data Connected Organisation:

- Connecting of the data within an organisation to deliver accurate reliable information, provides transparency, consistency, clarity, control and achieves strategic business outcomes
- Understanding the “role of data” within a “connected organisation”, vital to ensure business change programmes deliver the value you need
- A data connected organisation will enable effective portfolio management & unlock the value of your data
Steve Weyer  
Founder & CEO, UNIFYDATA

• Founder & CEO of UNIFYDATA & Capital Markets Consulting.
• 30+ years of experience delivering business & data change related solutions & systems to the financial services industry.
• Focused on global banking & capital markets.
• Provides leading techniques, tools, & platforms to deliver regulatory change imperatives whilst streamlining business processes & increasing operational Alpha.
• Previously 17 years with IBM & Andersen Consulting
• University of Oxford, Worcester College, Mechanical Engineering - 1st (Hons)
Connecting the data within an organisation is critical to deliver accurate reliable information, provide clarity and control and meet reporting requirements of the regulators.

- Business challenge – exponential growth in business data, complexity of products, multiple locations
- Regulatory challenge – ever increasing demands for information: complexity, granularity and consistency
- Cost challenge – continual demand to “do more with less” creates pressure on headcount and costs

Source: McKinsey Institute ‘Big data: The next frontier for innovation, competition and productivity’
Technology issues & ineffective data management

Data Challenges

• Large disparate data sets on multiple applications create complexity and lack of understanding of linkages
• Performance issues can be caused by high data volumes
• Improvements in infrastructure can be slow to deliver and extremely costly
• Projects often fail as data related issues may not surface until late in the development cycle

Impact

• Business can’t wait - strategic solutions often have lengthy lead times
• Inefficient processes built around incomplete or inaccurate data sets
• Reliance on key “data experts” to plug the gap
• Risks implicit in dependency on EUCs
FIBO framework and terminology – What purpose?
• To clearly, completely & consistently describe entities

FIBO – What are the applications & business drivers?
• Customer management/decisioning
• Fraud management, financial crime, financial & regulatory reporting
• Increase business performance, reduce business risk
• Improve regulatory oversight, compliance and transparency

FIBO and graph - Why?
• Visibility & connectivity
• Data lineage
• Standard common language
• Business context surrounding data
• Rigorous reporting and regulatory compliance

FIBO and graph - What are the benefits?
• Data alignment & consistency
• Understanding lineage - traced back to its original, authoritative sources
• Streamlining & cost reduction – data usage, business process & reporting
• Data Connectivity across graph - detection and prevention of fraud and crime
• 360-degree visibility of your customers
• Underpins “what-if” scenario analysis
**A real world example in BCBS 239 – Risk reporting**

**What elements require consideration**

- Governance and infrastructure
- Risk data aggregation
- Risk reporting

**Conclusions**

- Mandates such as BCBS 239 - significant new demands on data architectures
- Fast access to real-time data lineage/financial risk information requires architectural change
- Organisations have solid justification for revisiting the old, relational reporting systems they've struggled with for years
- New architectures utilising FIBO/BIAN & Graph – not only deliver infrastructure for reporting but also assist in increasing business performance

Building robust/agile models & infrastructure for risk reporting means tackling serious data management issues.

- Data lineage
- Data silos
- Terminology differences
- Legal Entity Identifiers
- Data consistency and latency
Absence of business context in data

Challenge:
No business-context added to data for self-serviced exploration

Resolution:
Graph/Semantics approach:

Business meanings of file and column labels are not natively captured

<table>
<thead>
<tr>
<th>New Session ID</th>
<th>Process Stage Code</th>
<th>Count</th>
<th>Date</th>
<th>Application ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>10000001</td>
<td>600</td>
<td>1 20180603</td>
<td>20000001.0</td>
</tr>
<tr>
<td>51</td>
<td>10000001</td>
<td>610</td>
<td>1 20180603</td>
<td>20000001.0</td>
</tr>
<tr>
<td>52</td>
<td>10000001</td>
<td>620</td>
<td>NaN</td>
<td>20000001.0</td>
</tr>
<tr>
<td>53</td>
<td>10000001</td>
<td>630</td>
<td>NaN</td>
<td>20000001.0</td>
</tr>
<tr>
<td>54</td>
<td>10000001</td>
<td>640</td>
<td>NaN</td>
<td>20000001.0</td>
</tr>
<tr>
<td>55</td>
<td>10000001</td>
<td>650</td>
<td>NaN</td>
<td>20000001.0</td>
</tr>
<tr>
<td>56</td>
<td>10000001</td>
<td>660</td>
<td>1 20180603</td>
<td>20000001.0</td>
</tr>
</tbody>
</table>

Natively represents business context in data models
Challenge:
Opaque relational schema cannot be used by business users for data exploration

Resolution:
Graph/Semantics approach:
Challenge: Rigid schema need major changes to accommodate new information

Resolution: Graph/Semantics approach:

P&L data received separately cannot be easily added to the schema

P&L integrated automatically when onboarded
Business Performance Suite & R-PaaS (Regulatory – Platform as a Service)

- Need for **cost effective** solution for **complex regulatory reporting** requirements
- Need for reporting to be **responsive and flexible** to meet increasing and ever changing regulatory demands
- By using standard **regulatory data models**, **consistent flexible reporting** is delivered across multiple regulations

- Business process and business logic under **strict control** improves & maintains **data quality**
- Specialist artefacts **accelerate delivery** of complex regulatory programmes
- R-PaaS offers the opportunity for organisations to **outsource** their regulatory reporting requirements to **industry experts**

**GRAPH & FIBO**

**INDUSTRY SEMANTICS & GRAPH MODELS**

- Contextually integrated policy, operations and data modelling
- Automated policy-based business rules processing
- Generate Once
- Analytic models
- Store Once
- Control Models
- Use many

**ANALYTICS & DASHBOARDS**

- Operational Objectives
- Compliance Objectives
- Reporting Objectives

**EXTERNAL DISCLOSURES**

- R&P, CRD IV
- BCBS 239,
- Dodd Frank,
- EMiR etc.,

**DATA SOURCES**

- Risk Data
- Reference Data
- Trading Data
- Finance Data
- Customer Data
- Market Data
Michael Bartlett
Director of Sales, Financial Services, Cambridge Semantics, Inc.

• Director of Sales for Financial Services at Cambridge Semantics, where he helps customers realize the transformational power of graph-based, ontology-driven Enterprise Data Fabric technologies in their daily operations.

• Prior to joining Cambridge Semantics, focused on data management and analytics sales and strategy within the financial services vertical enabling organizations to access and trust the quality of their data and analytics.

• Michael has a Bachelor of Science in Science Business from the University of Notre Dame.
We apply semantics and graph to a data fabric – so anyone can find, understand, blend, and use enterprise data.

- Based in Boston
- 110+ Employees
- Origins in IBM and Netezza
- Anzo 5.0 GA Feb 2020
Cambridge Semantics Technology Component Overview

**ANZO®**

A discovery and integration platform for an enterprise data fabric

- Applies graph and semantics to modern data management
- Multi-cloud speed, scale, and flexibility
- Enterprise-grade governance, security and metadata
- Business user ready: GUI, application, no-code
- Embeds AnzoGraph engine
- Integrated multi-cloud automation

**ANZO®GRAPH™**

A graph data warehouse for high performance integration and analytics and scale

- OLAP warehouse-like system
- MPP design for extreme scale/performance
- Multi-modal analytics capability
  - Warehouse-style queries competitive with Redshift, Snowflake and SparkSQL
  - Graph analytics and graph algorithms competitive with Neo4J and TigerGraph
  - Embedded data science and machine learning algorithms
- Open APIs for 3rd party extensions, algorithms and connectors
A modern data discovery and data integration as a service platform for your enterprise data fabric.

Anzo lets business users find, connect, and blend enterprise data into analytic ready datasets.
Catalog and map your existing data assets – structured or unstructured.

Translate dataset into graph models. Add business definitions, object types, and relationships with semantics.

Create blended analytic ready datasets. Connect graph models. Transform data. Harmonize into canonical models.

Analyze data using semantic and graph models. Export data and provide services for use with BI, analytics, and machine learning tools.
Automated Deployment and Operations with Kubernetes

Storage and Compute Integration

Graph Data Model
- Lift Data into Data Fabric
- Design Ontologies
- Connect Data Models

MODEL

BLEND
- Combine and Align Related Data Sets
- In-memory MPP OLAP Query Engine
- Data Layers

ONBOARD
- Automated ETL
- Collaborative Mapping
- Metadata Capture

Ingest & Map

ACCESS
- Analyze All Data Together
- Fast, Iterative Queries Ad Hoc, What if
- Code-Free or API

Hi-Res Analytics

ONBOARD

Machine Learning and AI

“Last Mile” Analytics Tools

Enterprise Search

Cloud or On-Prem Data Storage Infrastructure

Ingest
The Data Fabric for Compliance and Surveillance

Example Graph Model
- 400 million events
- 50 million reference records
- 321 million EComms
- 45 billion triples

Unstructured Data - Text-based Communications
- 321 million messages
- 36 billion triples
- Indexed for search using ElasticSearch

Structured Data - Transactions and Reference data records
- 400 million transactions
- 50 million reference records
- 9 billion triples

Data Model, Blend, and Access
- 30 classes
- 32 data layers
- 7 dashboards
- 20+ visualizations
*DEMO*
UNIFYDATA – offering a complete solution

- UNIFYDATA focussed upon delivering lean data solutions within the regulated markets of global financial markets, insurance and asset management
- Working closely with technology partners we are able to offer a full range of solutions, consulting & delivering upon process re-design and streamlining or fully outsourced data management & regulatory compliance solutions
- Our global client base includes a number of major institutions within banking and insurance

Global presence via our partnering network, operating with native speakers in a number of jurisdictions. Outstanding track record delivering complex programmes of data change and transformation within challenging regulatory environments
Questions?
FOR MORE INFORMATION:

Michael Bartlett  
Director, Financial Services  
Cambridge Semantics, Inc.  
michael.bartlett@cambridgesemantics.com