

# Data Harmonization

Enabled by Semantic Data Standards and FIBO

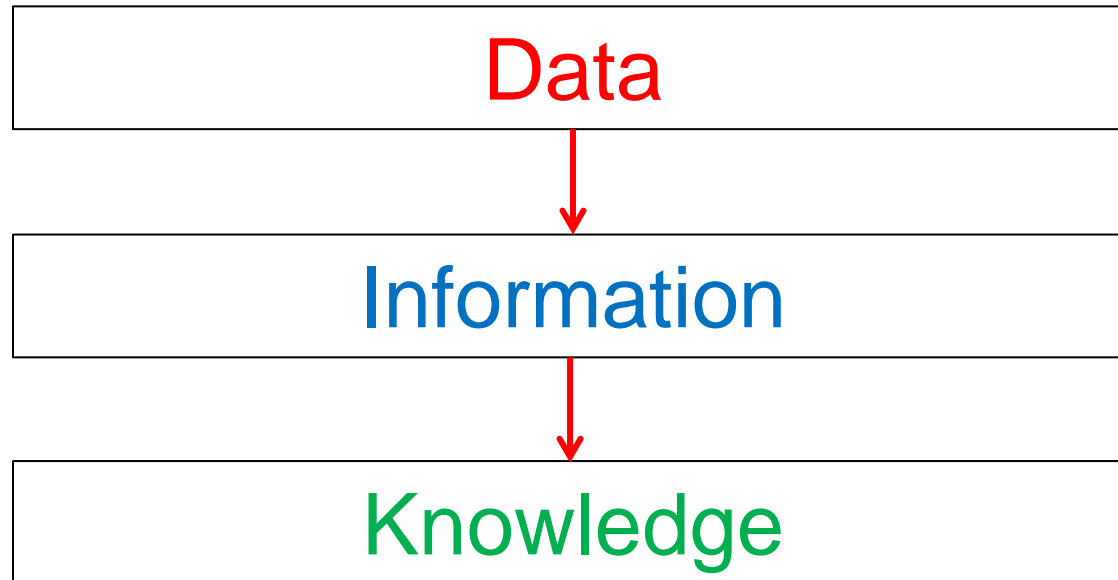
Financial Intelligence and Strategic Decision-Making  
06 March 2020

David N. Saul  
Executive in Residence  
Foisie Business School  
WPI

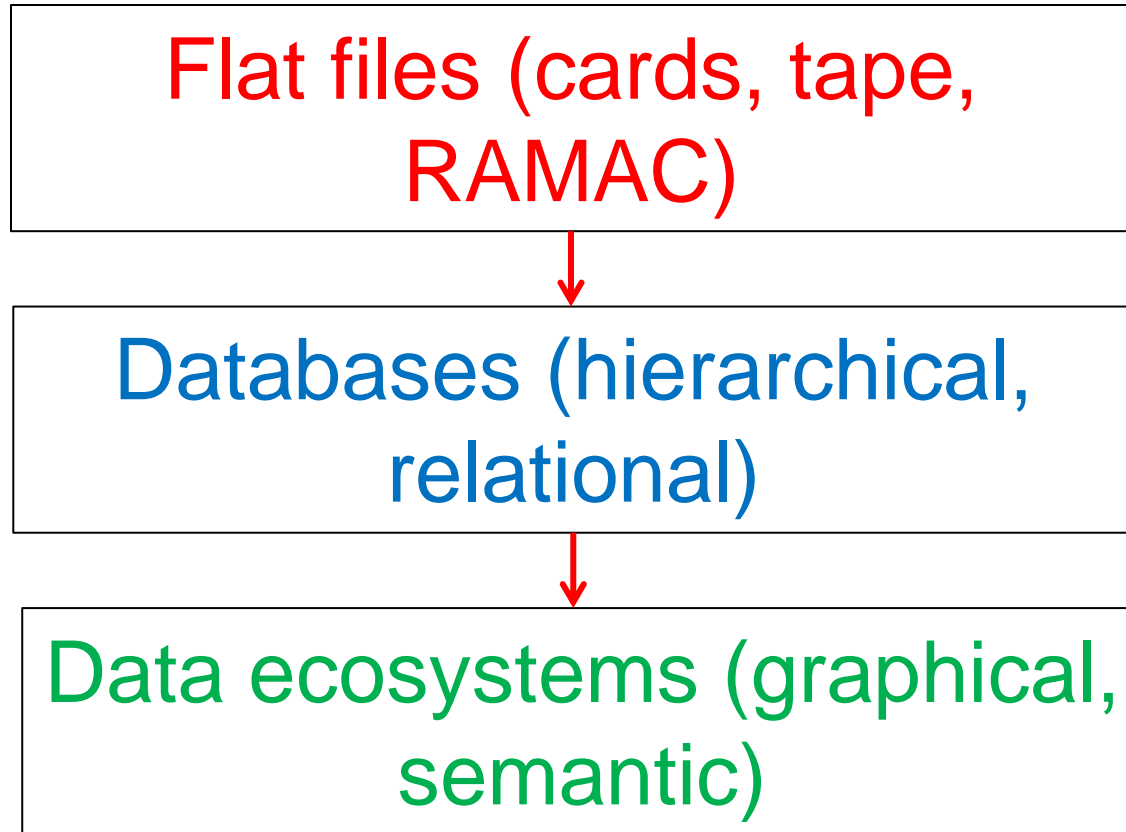
# Agenda

- Evolution of data management
- Dual conflicting requirements for financial firms
- Public/private collaboration
- Data harmonization initiatives
- Trust as a product of cross financial sector collaboration
- Semantic data is smart data
- W<sup>3</sup>C semantic or linked data standards
- What is FIBO and what can it do?
- FIBO concepts
- FIBO semantically defines core financial industry concepts and relationships
- How is FIBO used in the financial industry
- Fund reporting Proof of Concept

# Evolution of data management - 1



## Evolution of data management - 2



# Dual Conflicting Business Requirements for Financial Firms



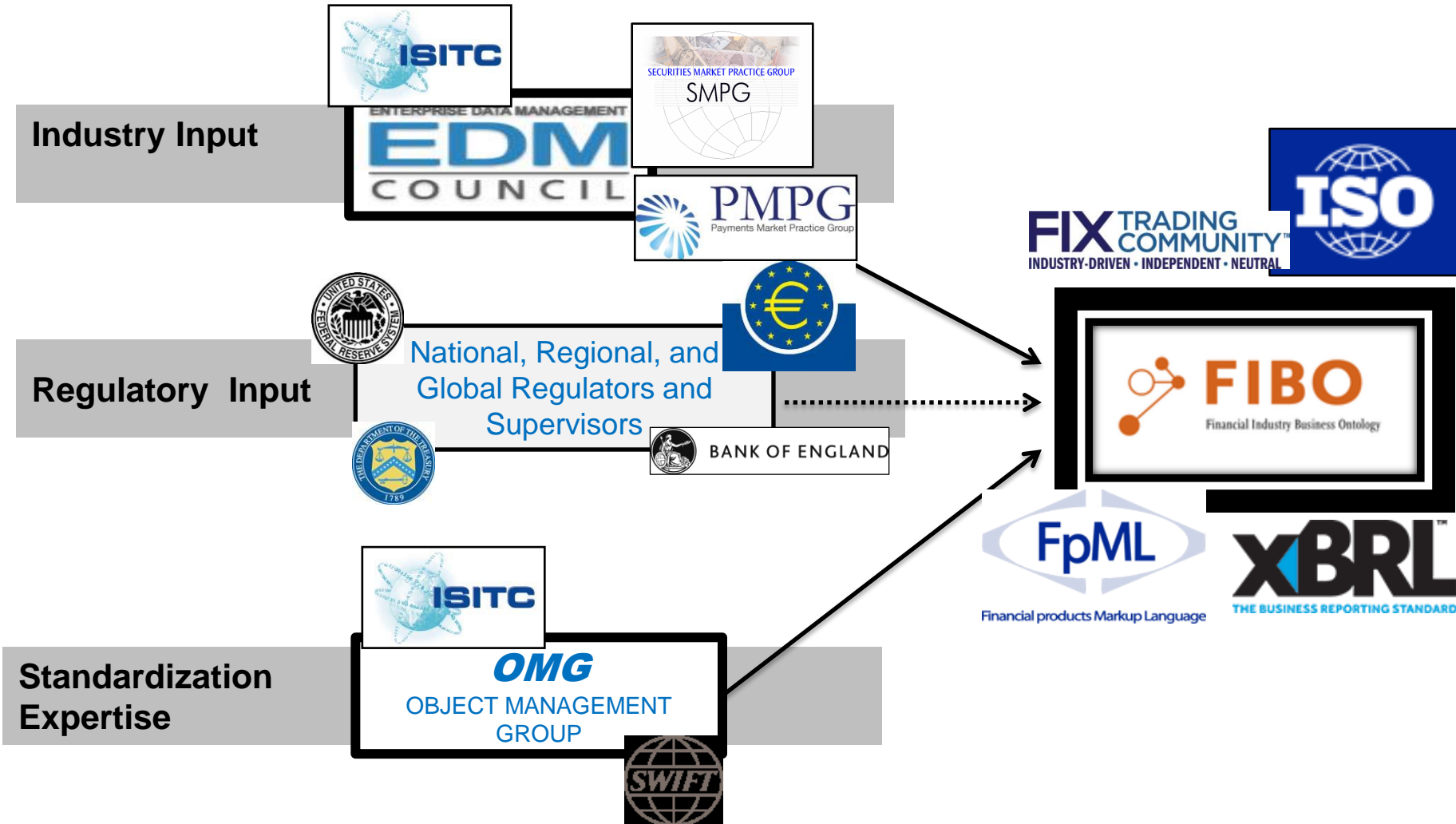
- Improve capabilities to exploit analytical tools and techniques
- Extend better quality to clients, as well as new products and services
- Exploit better and broader data integration for deeper insights and enhanced management

- Reduce burden in meeting existing and new requirements
- Find commercial benefits from investments in regulatory compliance
- Reduce perceived and actual risks

***Semantic data standards enable data synergy for both purposes.***

# Getting to Industry Standards and Shared Meaning

Public/Private Collaboration among many key contributors



*The information contained above is for illustrative purposes only*

# Data Harmonization Initiatives - 1

- **Legal Entity Identifier (LEI):** The Legal Entity Identifier (LEI) is a 20-digit, alphanumeric code based on the ISO 17442 standard. developed by the International Organization for Standardization (ISO). It connects key reference information that enables clear and unique identification of legal entities participating in financial transactions
- **Basel Committee on Banking Supervision:** Data harmonization and cross-industry data comparability are key components of BCBS Standard 239. Eleven of the principles outline ways for financial institutions to develop a bird's-eye view of the risks they face across their businesses and legal entities and three specify the role of regulators in monitoring and encouraging compliance.
- **Enterprise Data Management Council (EDMC):** The Financial Industry Business Ontology (FIBO) is a conceptual model of financial instruments, legal entities and financial processes. It is used as a common language for harmonization of data across disparate repositories; as a mechanism for validating data quality and as a graph structure for flexible financial analysis.
- **Object Management Group (OMG):** The mission of the Object Management Group (OMG) is to develop technology standards that provide real-world value for thousands of vertical industries. OMG's modeling standards, including the Unified Modeling Language® (UML®) and Model Driven Architecture® (MDA®), enable powerful visual design, execution and maintenance of software and other processes.

## Data Harmonization Initiatives - 2

- **International Organization for Standardization (ISO):** ISO 20022 standard is an international financial industry message scheme (which also used to be called "UNIFI") that defines the ISO platform for the development of financial message standards. Its business modelling approach allows users and developers to represent financial business processes and underlying transactions in a formal but syntax-independent notation.
- **International Swaps & Derivatives Association (ISDA):** The ISDA Common Domain Model (CDM) is intended to provide a standard representation of data, events and actions that occur during the life of a derivatives trade. Establishing a common set of data and processing standards that all participants can access and deploy will enhance consistency and interoperability across firms and platforms. This is necessary in order to realize the full potential of new technologies like distributed ledger and smart contracts.
- **Bank for International Settlements (CPMI-IOSCO):** The Principles for financial market infrastructures (PFMI) are the international standards for financial market infrastructures, i.e. payment systems, central securities depositories, securities settlement systems, central counterparties and trade repositories. Issued by the Committee on Payments & Market Infrastructures (CPMI) and the International Organization of Securities Commissions (IOSCO), the PFMI are part of a set of twelve key standards that the international community considers essential to strengthening and preserving financial stability.

– Unique Trade Identifier (UTI) & Unique Product Identifier (UPI)



## Data Harmonization Initiatives - 3

- **Bloomberg:** OpenFIGI is a 12 character, alphanumeric, randomly generated ID covering hundreds of millions of active and inactive instruments. In total, there are over 300 trillion potential identifiers available. The identifier itself acts as a Uniform Resource Identifier (URI) to link to a set of metadata that uniquely and clearly describes the instrument.
- **International Securities Association for Institutional Trade Communication (ISITC):** The Reference Data & Standards Working Group advances critical discussion on issues impacting financial instrument and market participant reference data within the trade lifecycle and works with other industry groups and participants in developing and promoting market practice standards.
- **Governance Risk and Compliance Technology Centre (GRCTC):** Semantic technology-based solutions being developed at the GRCTC help organizations:
  - Perform enhanced Regulatory Compliance Change Management;
  - Create regulatory and business vocabularies and rules;
  - Develop enterprise-wide knowledge bases to enable better risk data management and aggregation;
  - Perform regulatory compliance knowledge management;
  - Enhance regulatory compliance capabilities.

# Trust as the Product of Cross-Financial System Collaboration on Semantic Standards

Potential benefits to all constituencies

- *Products and services opportunities*
- *Path to improved efficiency and reliability*
- *Clients and public transparency*
- *More confidence by counterparties*
- *Less complex risk management*
- *Way to lower cost of compliance*
- *Improved information security*

Financial Services Organizations

Standardization Processes



- *Alignment with statutory missions*
- *Regulatory and supervisory/prudential surveillance improved*
- *Ability to detect and respond to systemic and internal risks*
- *Potentially lower costs and burdens*

Product/Service Providers

Regulators

- *Revenue possibilities*
- *Opportunity for innovation and development*
- *Cost avoidance*
- *IP innovation with protection*
- *Lower risk investment*

Standards Organizations

Standardization Processes

- Standards**  
Benefits All Might See:
- *Confidence*
  - *Efficiency*
  - *Growth*
  - *Effectiveness*
  - *Lower Risk*
  - *Simplicity*
  - *Transparency*
  - *Precision*

# Semantic Data is Smart Data

- **Semantics:** The same concept that World Wide Web creator Tim Berners-Lee used to develop hyperlinks
  - Mathematical foundation in Graph Theory (Euler, 1739)
- **Semantics** gives meaning to data
  - Structured, Semi-Structured and Unstructured data
- **Semantic technologies** are used to map and attach (or tag) precise meaning onto each piece of data and its interrelationships
  - Built on *Dictionaries* and *Taxonomies*
  - *Ontologies* extend semantic meaning to describe the relationships between data elements
- **Semantics and semantic technologies** together enable transparent data analysis to create actionable, timely business insights
- Standards are required for the “meaning” to be held “in common” across institutions, processes and jurisdictions.

# W<sup>3</sup>C Semantic or Linked Data Standards

- **RDF** (Resource Description Framework): The data modeling language for the Semantic Web. All Semantic Web information is stored and represented within the RDF.
- **OWL** (Web Ontology Language): The schema language, or knowledge representation (KR) language, of the Semantic Web. OWL enables you to precisely define concepts as components so that these concepts can be reassembled and reused as needed for many different applications and purposes.
- **SPARQL** (SPARQL Protocol and RDF Query Language): The query language for the Semantic Web. It is specifically designed to query data across multiple different systems.

# What is the Financial Industry Business Ontology (FIBO)?

- The concept of ontology was developed as early as the 3<sup>rd</sup> century BC by ancient Greek philosophers. The objective was to establish common meaning to promote collaboration on the fundamental nature of reality. FIBO continues that tradition as it is a formal and factual model (i.e. an ontology) of the reality of how financial instruments, legal entities and processes work. FIBO is expressed in the W<sup>3</sup>C standard Web Ontology Language (OWL) for machine readable inference processing.
- FIBO has been structured into three core modules –
  - Foundational components (the contractual scaffolding of FIBO).
  - Financial instruments (product domains).
  - Process areas (including issuance, corporate actions, pricing, loans and derivatives).
- FIBO is expressed in the triple store language of the Web (RDF/OWL) for machine readable inference processing and UML for people reading analysis.

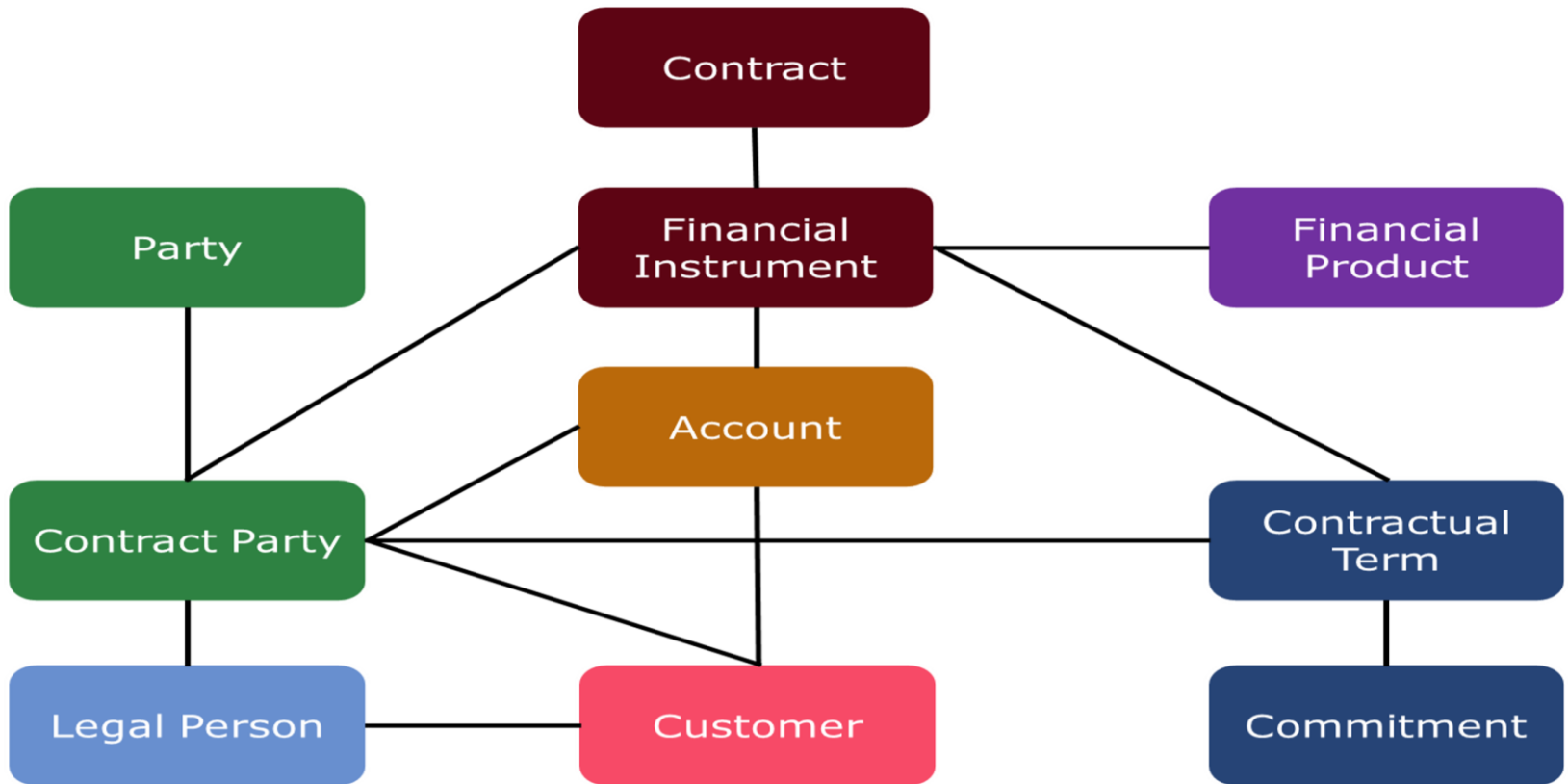
# What can FIBO do?

- The financial system is a massive global network of interdependent entities, instruments, transactions and processes. Individual firms manage risk and support operations in this environment by pulling data from vertically aligned systems that are linked to consuming applications and from independently managed data repositories that are compounded by inconsistent data meaning.
- FIBO establishes an industry-wide baseline and common language for rationalizing data to its precise legal meaning. This contractual certainty promotes trust in data and minimizes the need for costly reconciliation across disparate repositories. FIBO also establishes the semantic structure associated with financial contracts and expresses them in the Web Ontology Language (OWL). By defining these semantic concepts in OWL, FIBO takes advantage of machine intelligence and network graph capabilities to express, classify and link data. FIBO enables users to create and explore connections and relationships between instruments, business entities and counterparties to determine expected cash flow, unravel risk and identify transitive relationships that were not previously possible.

# FIBO Concepts

- **Ontology:** The FIBO ontology is a description of the concepts and relationships that exist in the financial industry. At its essence, FIBO is a formal model of precise legal obligations as defined in the contracts and agreements that govern our industry. These concepts and relationships have been verified by subject matter experts and constructed using widely accepted standards. The FIBO ontology is a stand-alone deliverable. It is used for data harmonization across repositories – regardless of format.
- **Triple Store:** FIBO is based on the W<sup>3</sup>C standard “triple store.” A triple store is a way of organizing information in groups of precisely defined “subjects” and “objects” linked together by standard “predicates.” This structure was invented by the US Department of Defense to analyze complex environments with rapidly changing sets of data. Triples are linguistically aligned with how we think and speak, can be understood by both humans and machines and are much more flexible than traditional relational databases for complex (ad hoc) queries.
- **Inference Processing:** The process of discovering new facts about your data based on a set of rules that render the meaning of concepts in an expressive way. FIBO is based on a set of verified facts and rules that can be used to unravel logical connections and gain new insights. FIBO does this without the need for physical models, new lines of application code or complicated schemas. Inference processing allows us to better understand relationships across legal entities, identify risks within our portfolios and gain insights into customers.

# *FIBO Semantically Defines Core Financial Industry Concepts and Relationships*



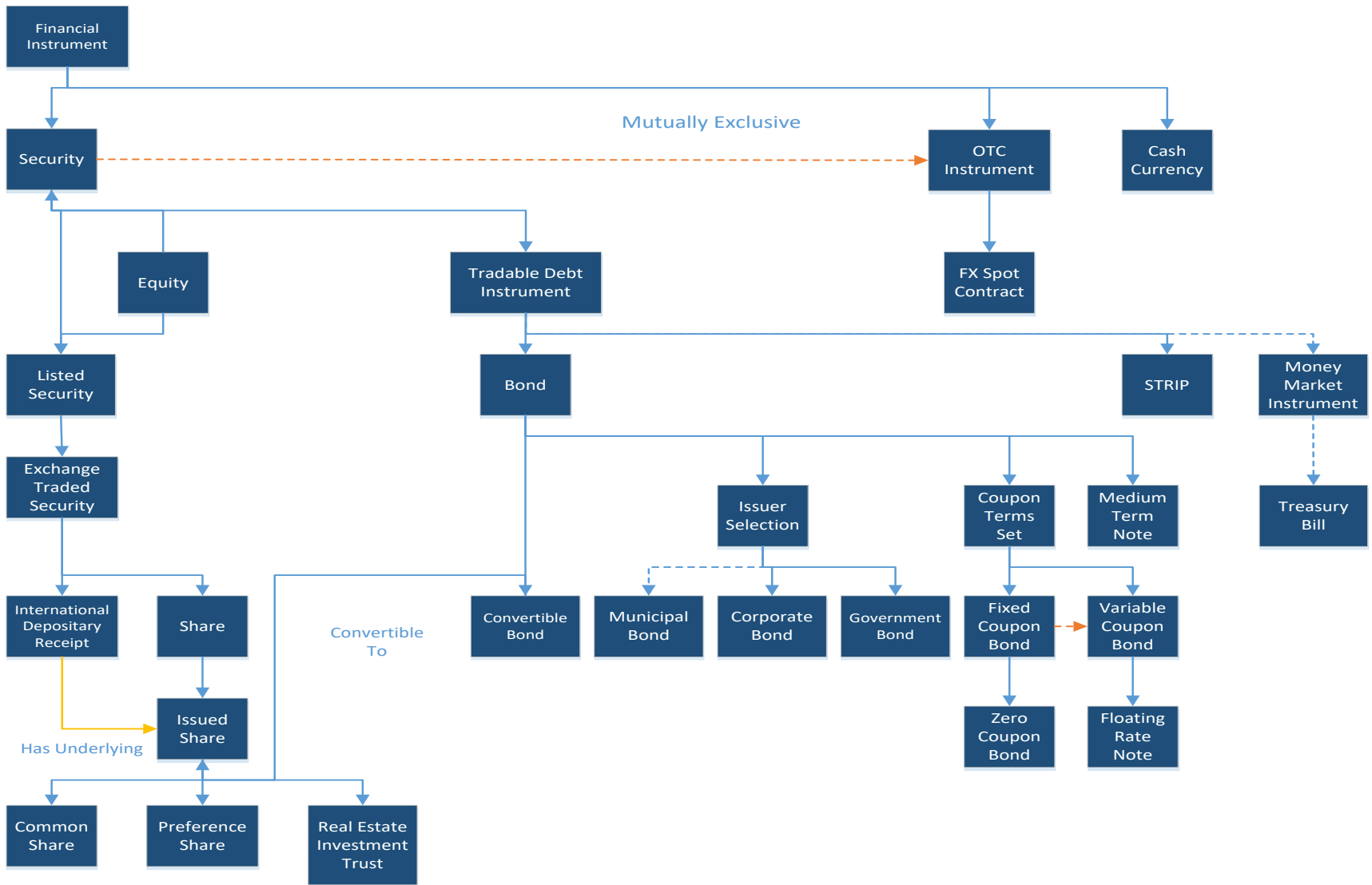
\* Slide courtesy of David Newman, Wells Fargo Bank



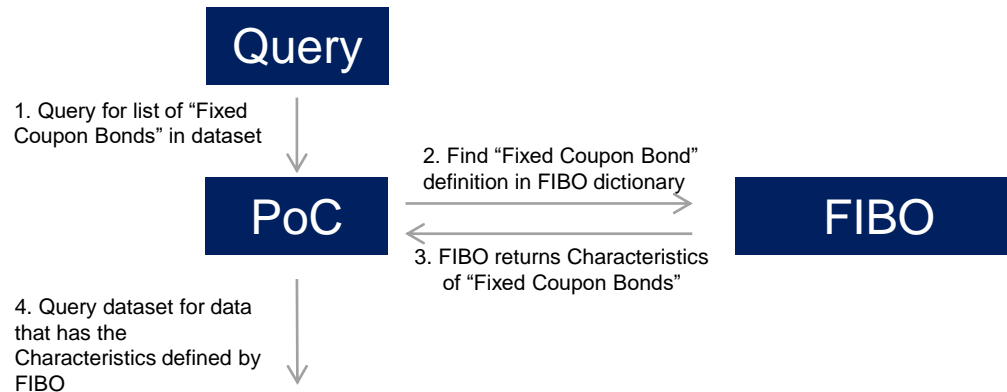
# How is FIBO Used in the Financial Industry

- **Data Harmonization:** The contractual certainty of FIBO replaces the ad hoc, spreadsheet-driven reconciliation processes that exist within many organizations and promotes confidence in data among business users.
- **Standardize Data Integration:** The alignment of content to its explicit meaning makes it easier to automatically process and integrate data from federated sources. FIBO reduces errors of interpretation, fosters reusability and facilitates response to changing conditions at lower cost.
- **Flexible Analysis:** FIBO supports the identification of complex relationships for assessing counterparty risk, classifying financial instrument types, flagging structural errors and calculating expected cash flow. FIBO delivers the BCBS 239 requirement for scenario-based reporting.
- **Blockchain:** FIBO creates a digital signature to ensure structural integrity of the data flowing across networks and allow users to take advantage of smart contract capabilities.

# Irish FIBO Fund Reporting Proof of Concept Taxonomy

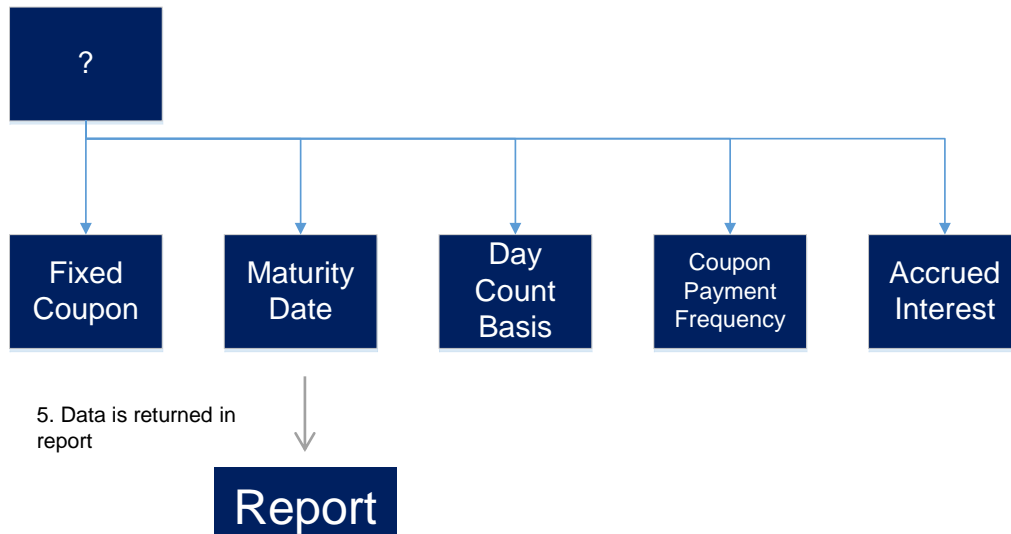


# Irish FIBO Fund Reporting PoC (Generalized Example)



- **Query:** Return data on all Fixed Coupon Bonds
- FIBO defines a Fixed Coupon Bond as having the following Characteristics:
  1. Fixed Coupon
  2. Maturity Date
  3. Day Count Basis
  4. Coupon Payment Frequency
  5. Accrued Interest

## FDR DATA



- The **PoC** queries the dataset returning data that has the Characteristics as described by FIBO
- Querying based upon **Characteristics** allows us to return all data that fits the FIBO definition, regardless of what database or system the data is stored on
- A traditional database **will not be able** to know this relationship unless it is explicitly stated in the database schema