Semantic Interoperability

The Next Integration Challenge

David S. Frankel
Lead Standards Architect, Model-Driven Systems
Global Ecosystem and Partner Group
Standards Management and Strategy
SAP Labs, LLC

THE BEST-RUN BUSINESSES RUN SAP™

Agenda

A Decade of MDA® – An Assessment

- Accomplishments & Gaps
- Emphasis on Integration

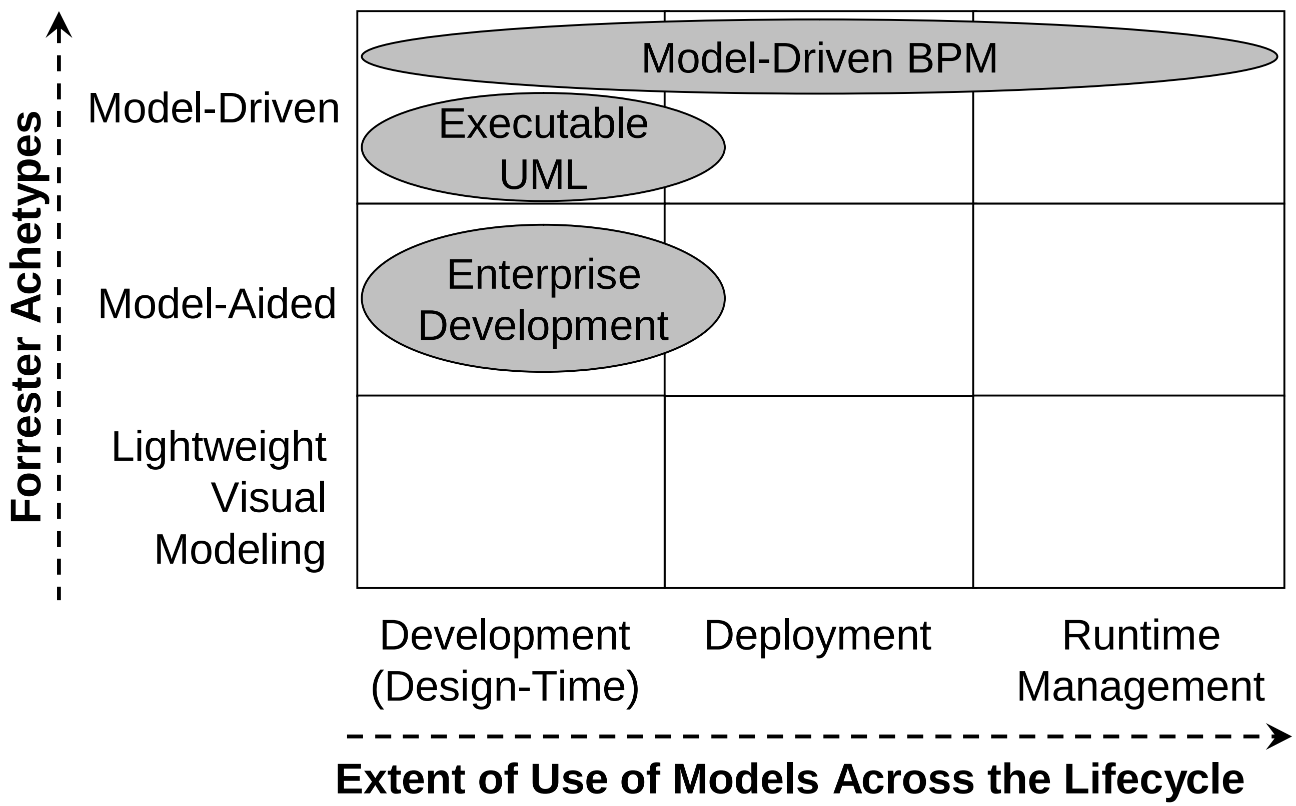
Data Integration: Macro-Economic and Technical Drivers

- Business Network Transformation
- Industrial Convergence
- Standards Silos
- The Persistence of Legacy

Semantic Interoperability

- The Semantic Integration Barrier
- Overcoming the Barrier with Semantic Metadata
- Industrial Trends

Sophistication of Model-Based Systems
Two Dimensional Taxonomy



Model-Based Systems Additional Points

The Model-Driven Architecture® (MDA) stack

- UML®
  - Much better tools available than earlier in the decade
  - UML tools' model management facilities are primitive
    - a long way from scaling to the requirements of enterprise-class model/metadata repositories
- XMI®
  - Interoperability among tools is not where it should be
  - UML tool ventors are finally cooperating with OMG efforts to fix this
- MOF®
  - Little adoption of pure MOF
  - But Eclipse's Ecore is close to MOF and has broad adoption as the core of the Eclipse Modeling Framework (EMF)
  - QVT (Query View Transformation) has little traction
- BPMN®
  - Inclusion in the MDA stack is a milestone
  - Integration with UML is an open question

Most growth in model-based systems is in model-driven integration

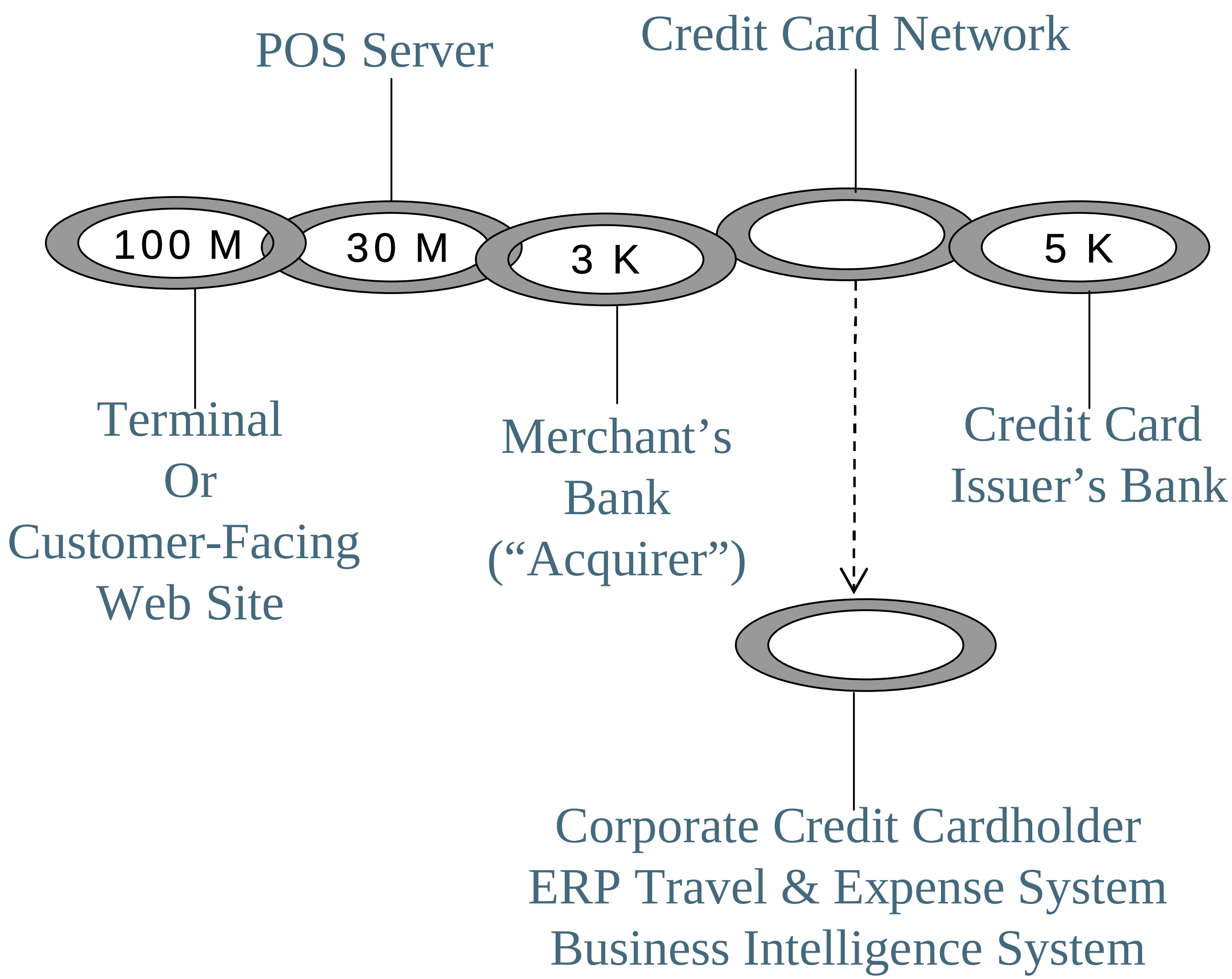
- The trend is for IT development budgets to fund integrations over classical application building

Macro-Economic Trend:
Business Network Transformation

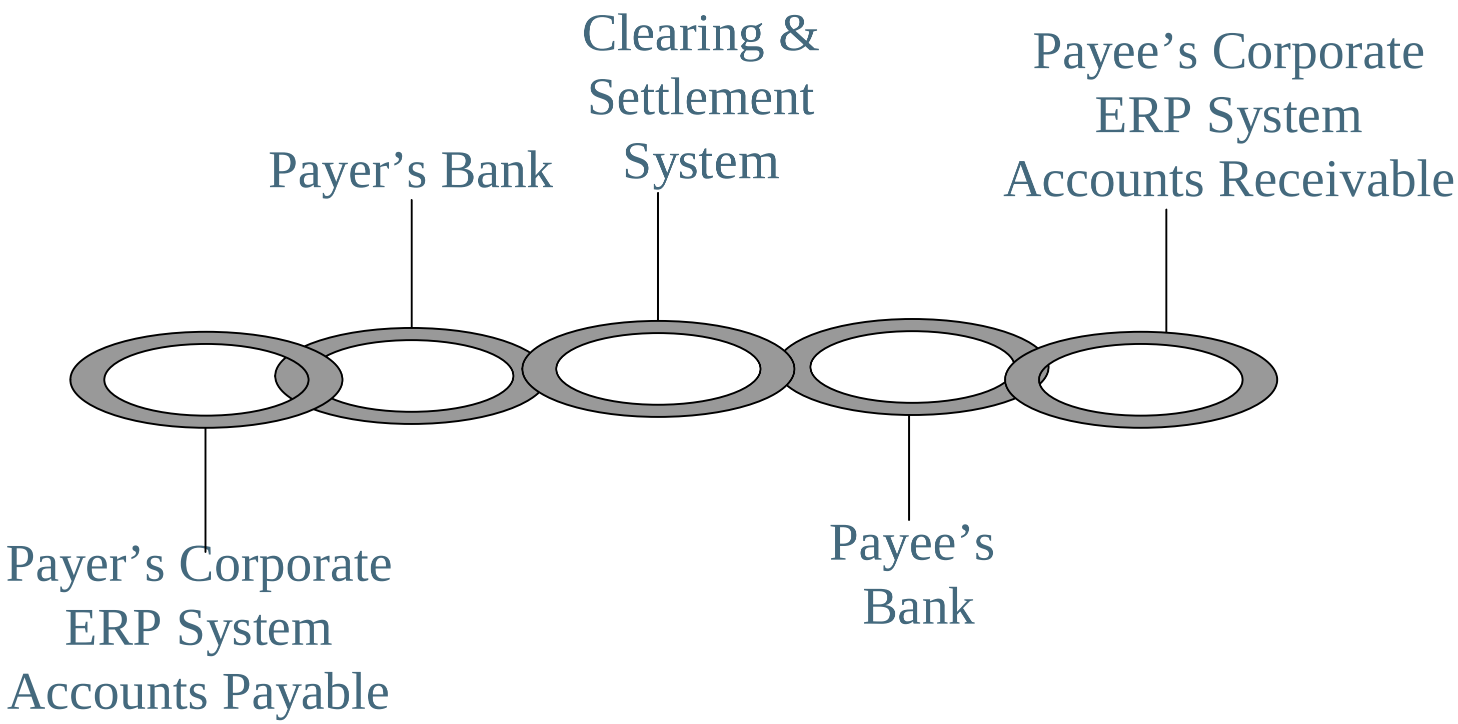
- A company's network of suppliers, distributors, customers, partners, and employees is becoming an increasingly important source of competitive advantage.
- Digitization of business networks is enabling new collaborative business models.
- This dynamic increases demand for integration.
- SOA and BPM initiatives both address and require integration.

Credit/Debit Card Value Chain

Crossing Multiple Subsystem Boundaries



Corporate-Bank Payments Value Chain



Corporate-Bank Integration Is Messy

Global corporates average 8 to 15 different bank relationships

Largest corporates have to deal with:

- 100+ banks
- 35-40 proprietary software systems
- 100 file formats
- 40 EDI connections

Requires costly data integration work up front

Payment orders for each bank have to be handled separately

Bank statements have to be collected from each bank separately

Fading Industrial Boundaries
Adding to the Integration Burden

Today, ***33 percent of the components in an automobile are high tech components***.

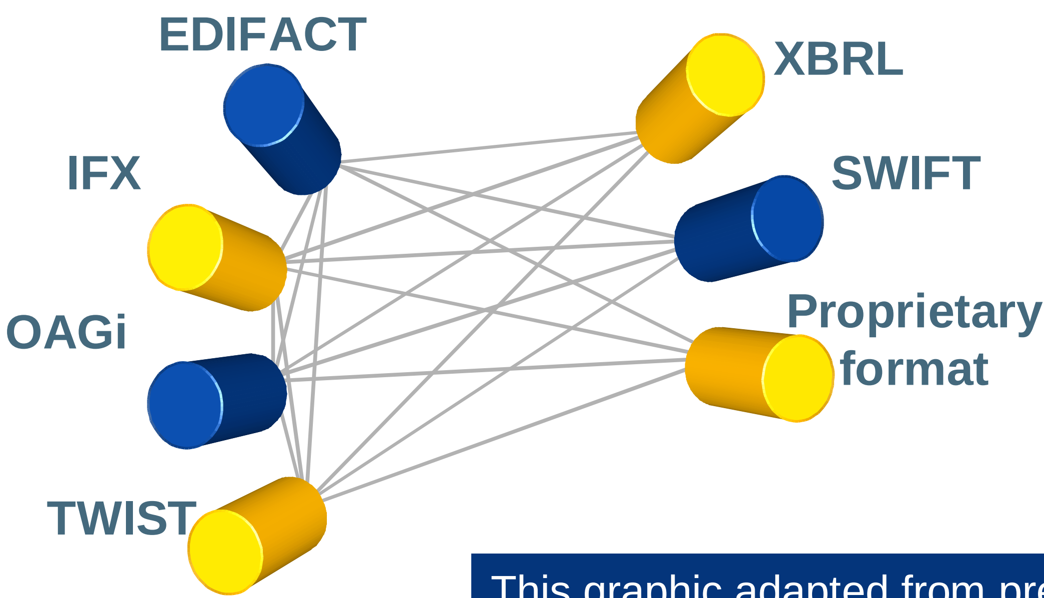
- Thus, messages defined by standards bodies for the high tech vertical market, such as RosettaNet messages, must interoperate to some extent with messages designed by standards groups that target the automotive industry.

Automobile also companies play role of health insurer and financial institution

Walls are coming down between the embedded systems and enterprise systems worlds

- ERP systems interface with manufacturing process control and fabrication

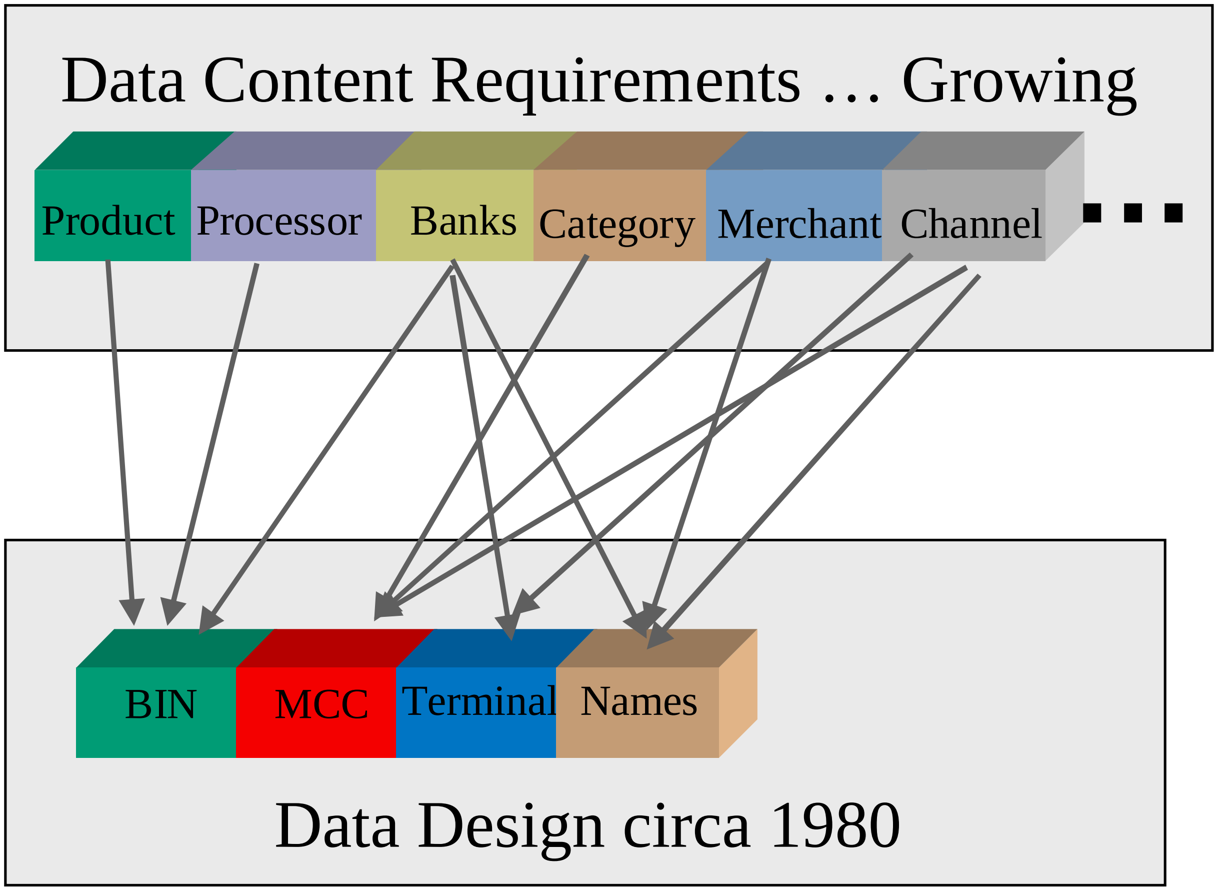
Standards Silos



This graphic adapted from presentation at www.iso20022.org

- **Incompatible standards in the financial services sector inhibits interoperability needed to build efficient business networks**
- **Contributes to enormous costs of integration**

Legacy Doesn’t Go Away
So we stretch it and stuff it and…



This graphic courtesy of Visa Inc.

Integration Problem in the Large

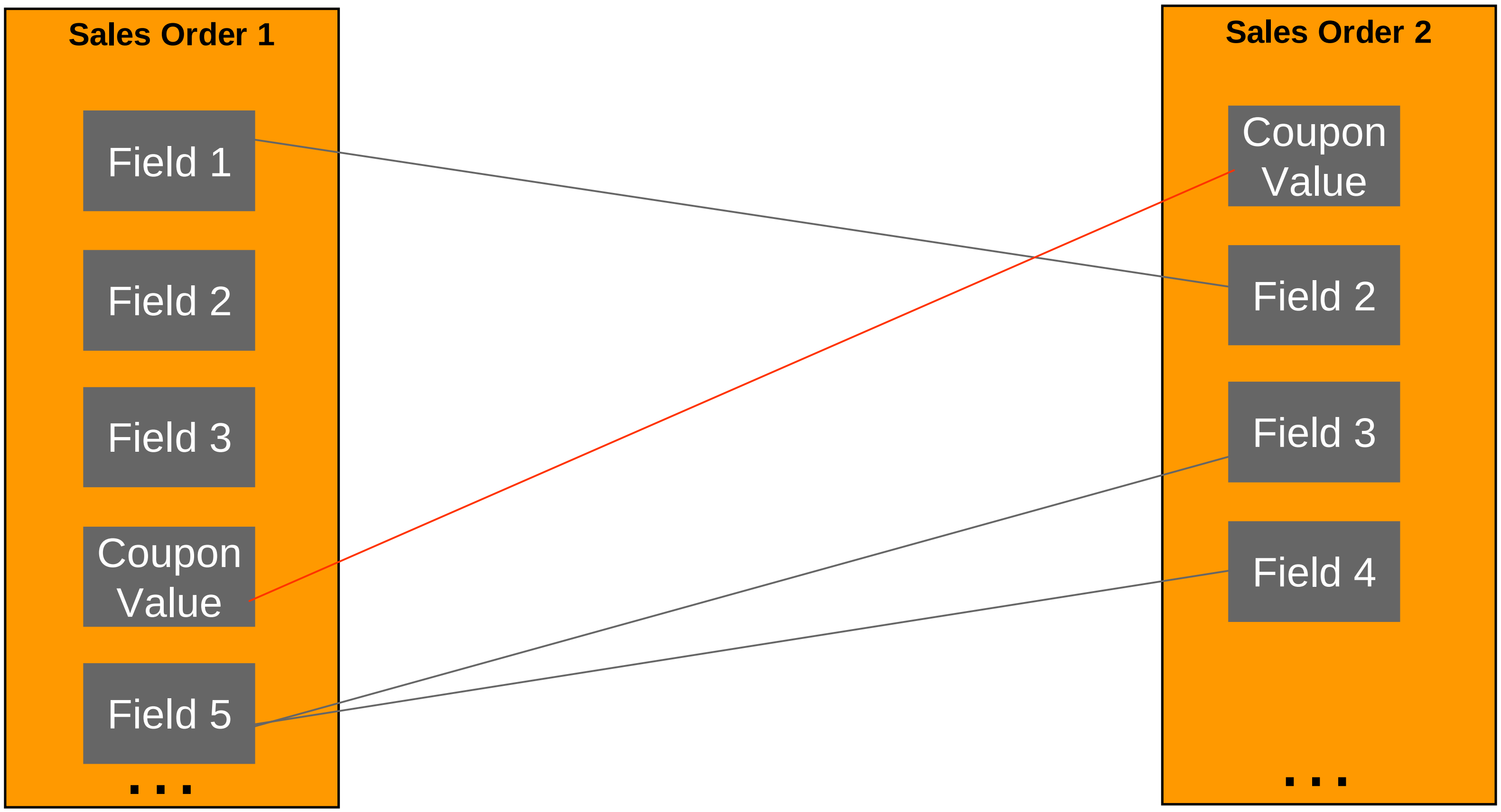
Integration costs consume 30-40% of IT budgets or more

Common occurrence: A business opportunities cannot be pursued because integration costs are too high

- Entering a new market segment or sub-segment carries high up-front costs

The State of the Art Data Integration Tools
But How to Decide What Maps to What?

Semantic interoperability requirements consume 80% of the integration effort



Expression Editor: ______________________________

Standardizing the Message Formats Does Not Solve the Problem

Semantics of data flowing through legacy systems is often poorly understood

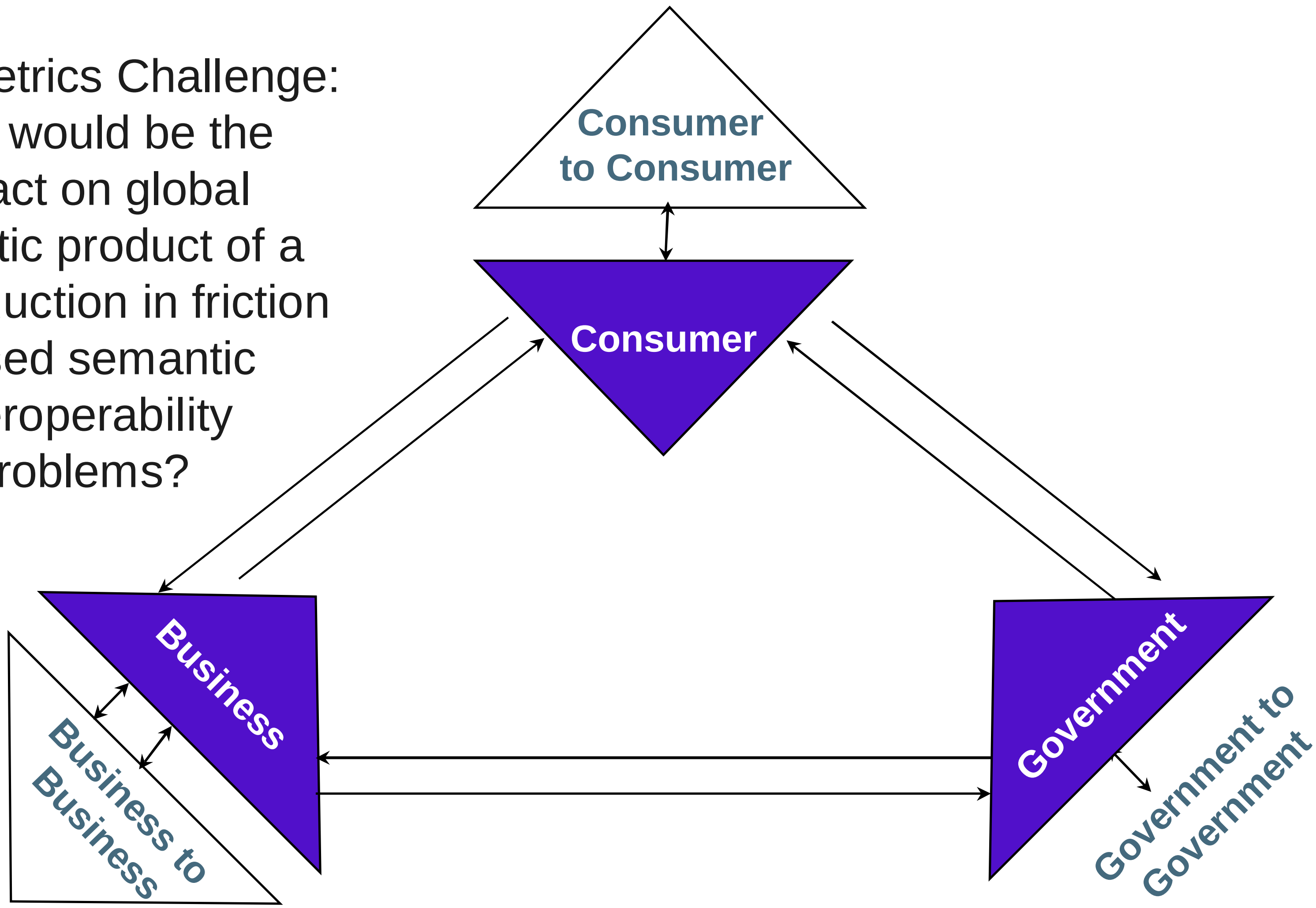
- Overloading of fields without traceability
- No common reference point for semantics

Persistence of legacy formats means analysts have to design many transformations at subsystem boundaries across the value chain

- Poor visibility of semantics of data results in subtle transformation errors and outright loss of data
- Causes friction in the system
  - Errors require expensive human intervention
  - Poor visibility makes its more difficult to detect and trace fraudulent transactions

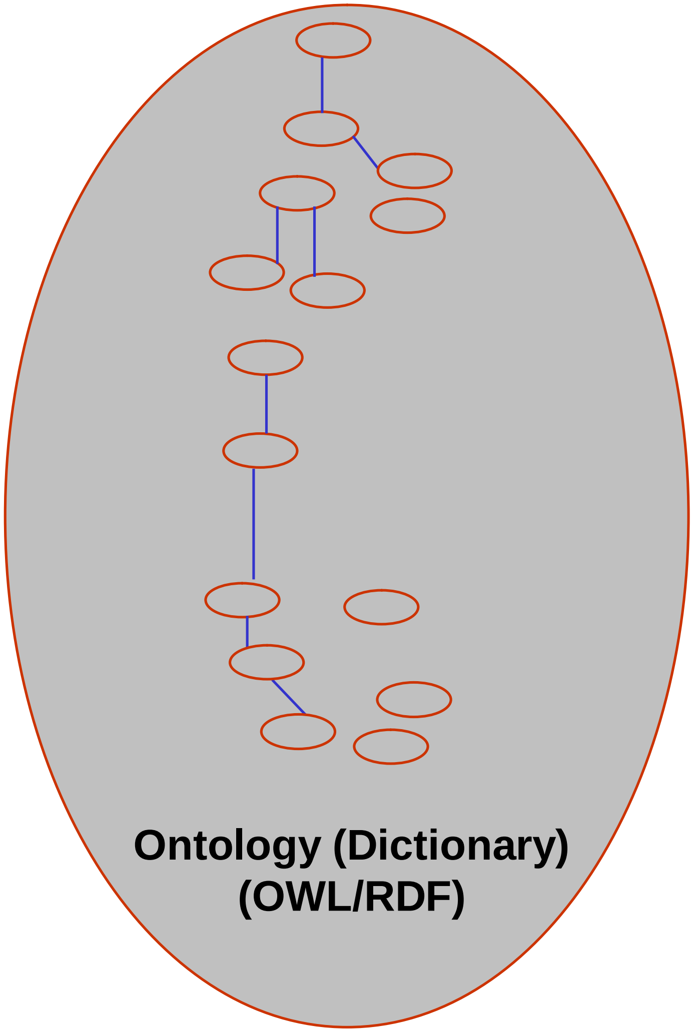
Economic Potential of Friction Reductions
Magic not required! Modest improvements help a lot

Econometrics Challenge:
What would be the impact on global domestic product of a 20% reduction in friction caused semantic interoperability problems?



This slide adapted with permission from Visa, Inc.

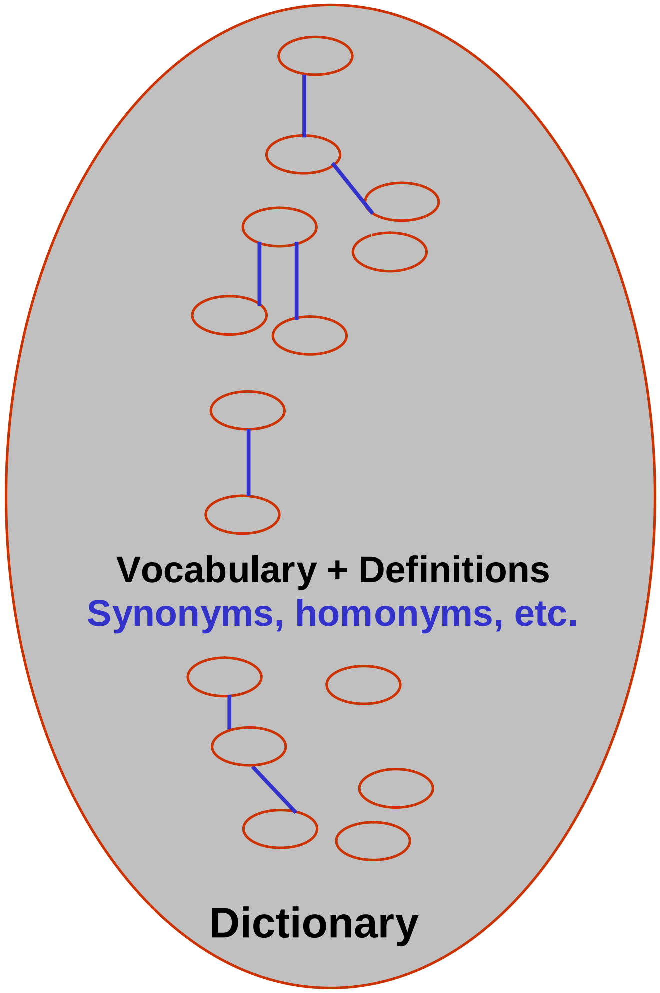
Semantic Ontologies and IT Data Structures
What is the Link?



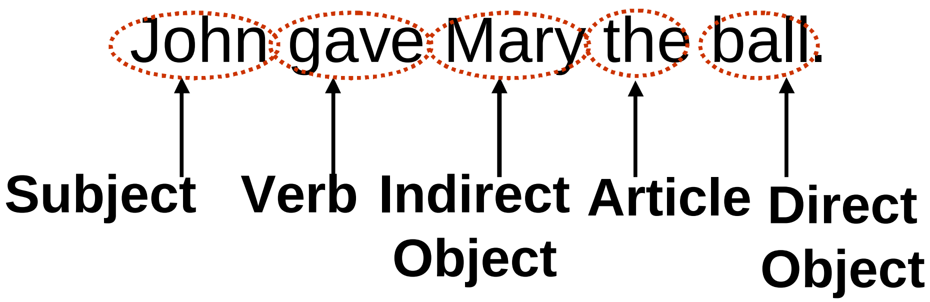
Transfer_Identification

+ ...
+ ...
+ ...
+ ...
+ ...
+ ...
+ Local_Instrument : SEC_Code [0..1]
+ ...
+ ...
+ ...

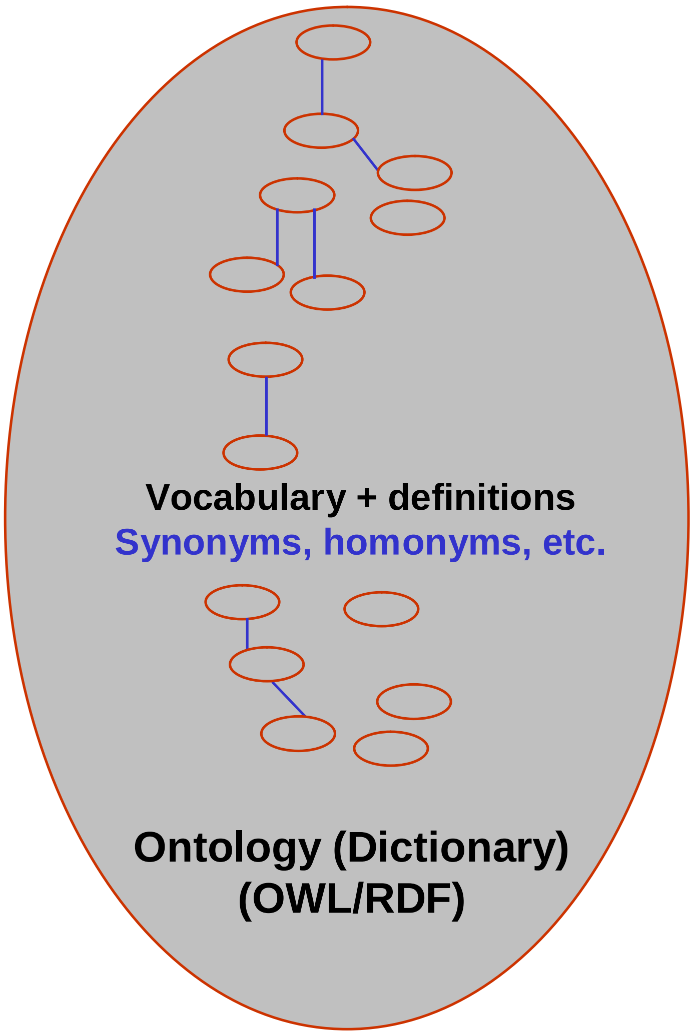
Basics of Semantic Metadata
Based on Principles of Language

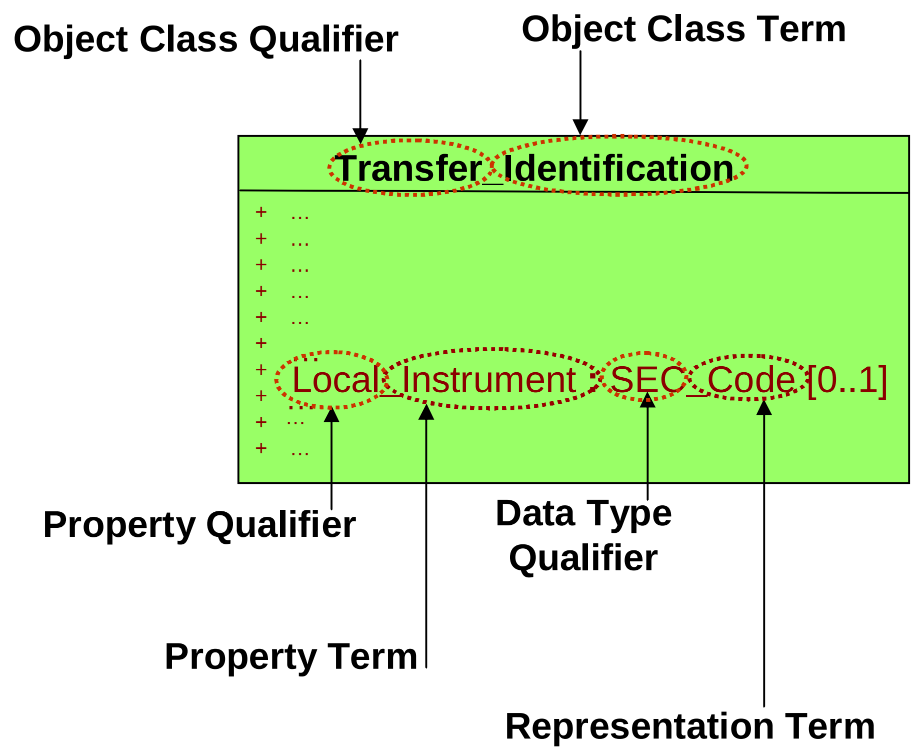


Words play roles in sentences.



Semantic Metadata (ISO 11179-5)
Linking Data Structures to Ontologies





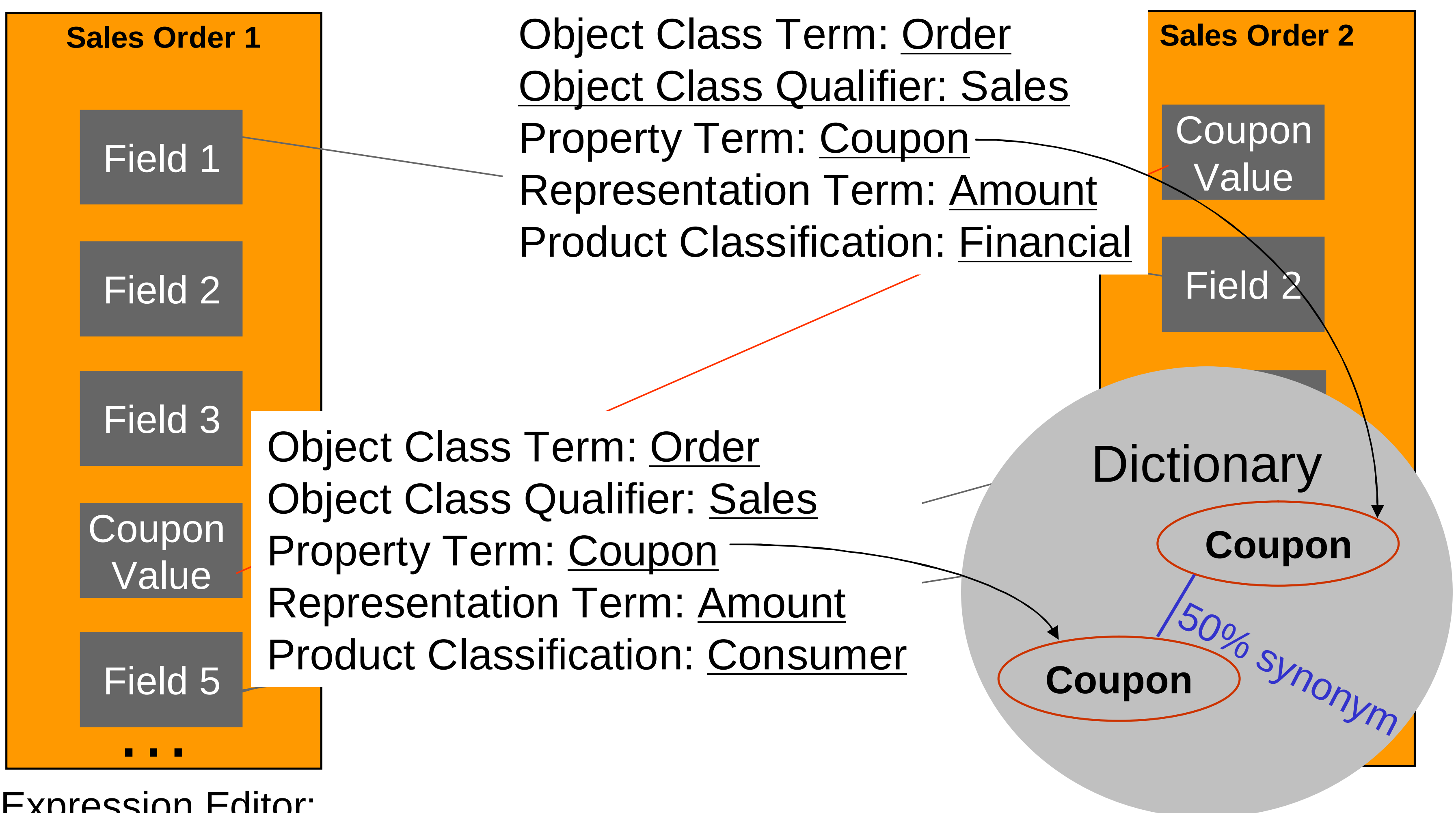
More Semantic Metdata
Business Context

Every data element is associated with a *Business Context*, which designates a value for one or more of the following properties:

- Business process: The business process(es) with which the element is associated
- Business process role: The role(s) directly involved in the business process(es) that are relevant to the element, e.g. shipper
- Supporting role: The role(s) indirectly involved in the business process; e.g., a data element in an order response from seller to buyer could be required by a third-party shipper
- Geopolitical scope: The continent, economic region, country, or region for which the element is relevant
- Industry classification: The industry or industries for which the element is relevant
- Product classficiation: The product class(es) for which the element is relevant
- Official constraint: A legal requirement(s) associated with the element; e.g., an element could be required by Sarbanes-Oxley
- System capability: Indicates that the element is there in order to address the limitations of a computing system(s)

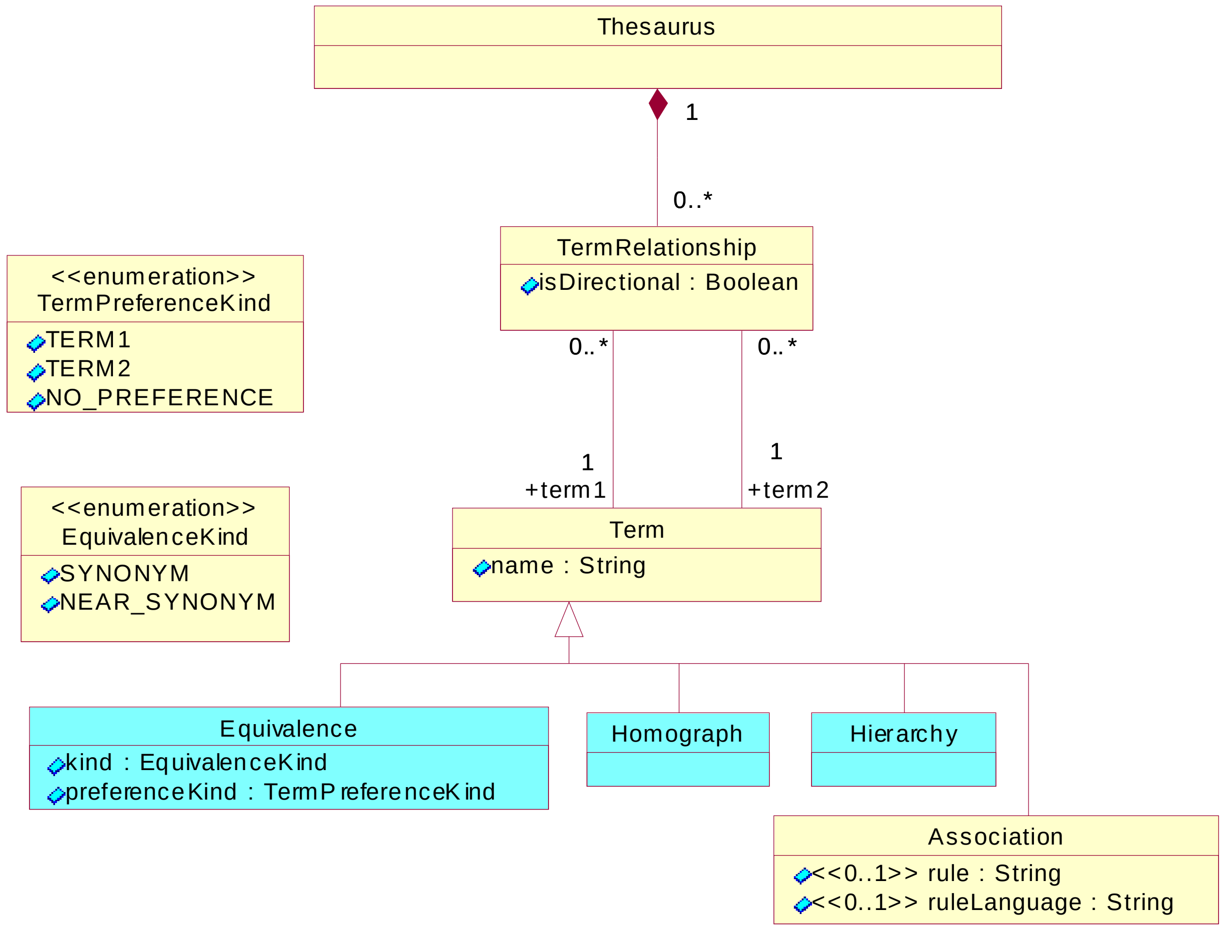
Enables software to calculate the divergence between two elements, and to trace how the difference in context accounts for the divergence.

Integration Tools That Use Semantic Metadata
Helping the Analyst Decide What Maps to What



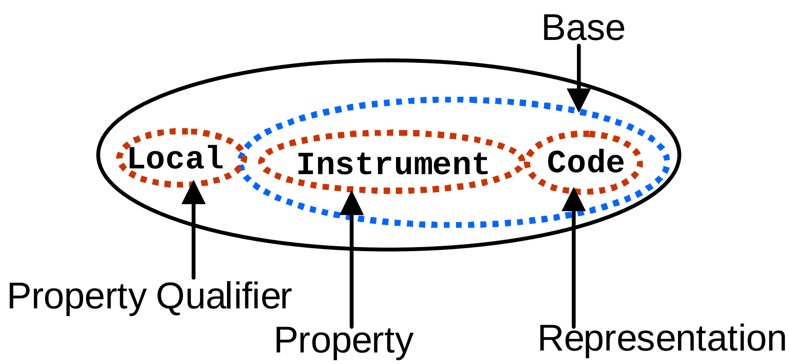
Expression Editor: ______________________

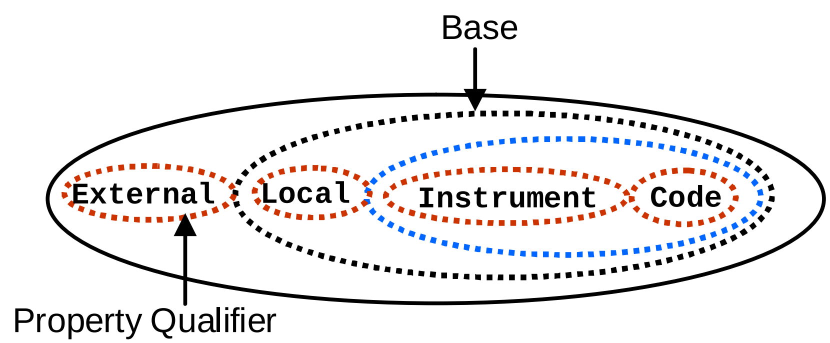
ISO 11179-5 Thesaurus Metamodel

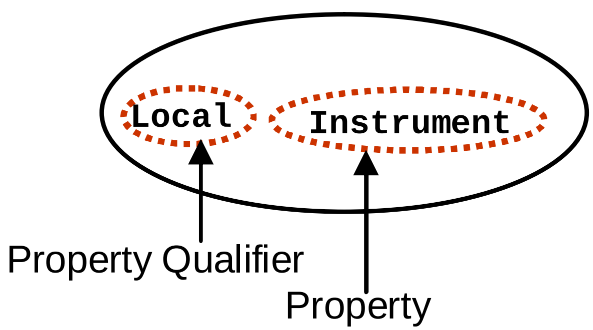


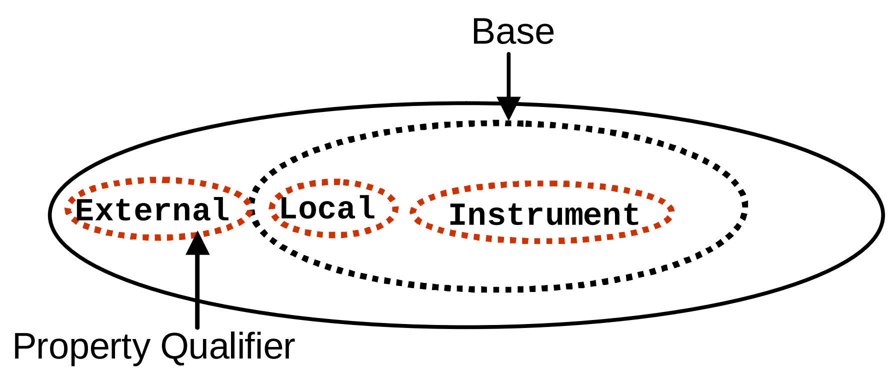
This slide adapted with permission from Visa, Inc.

Semantic Elements of Differing Granularity





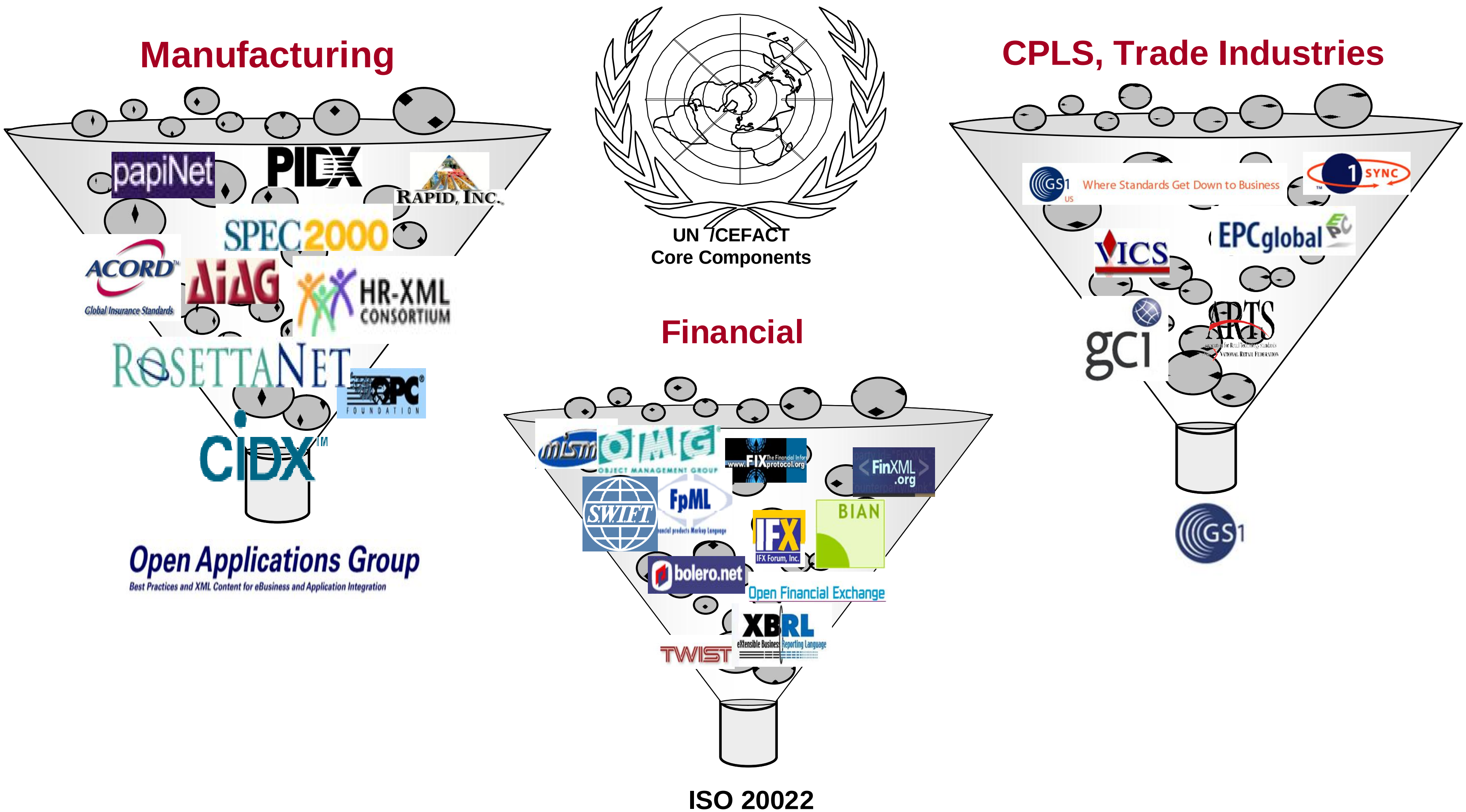




UN/CEFACT Representation Types

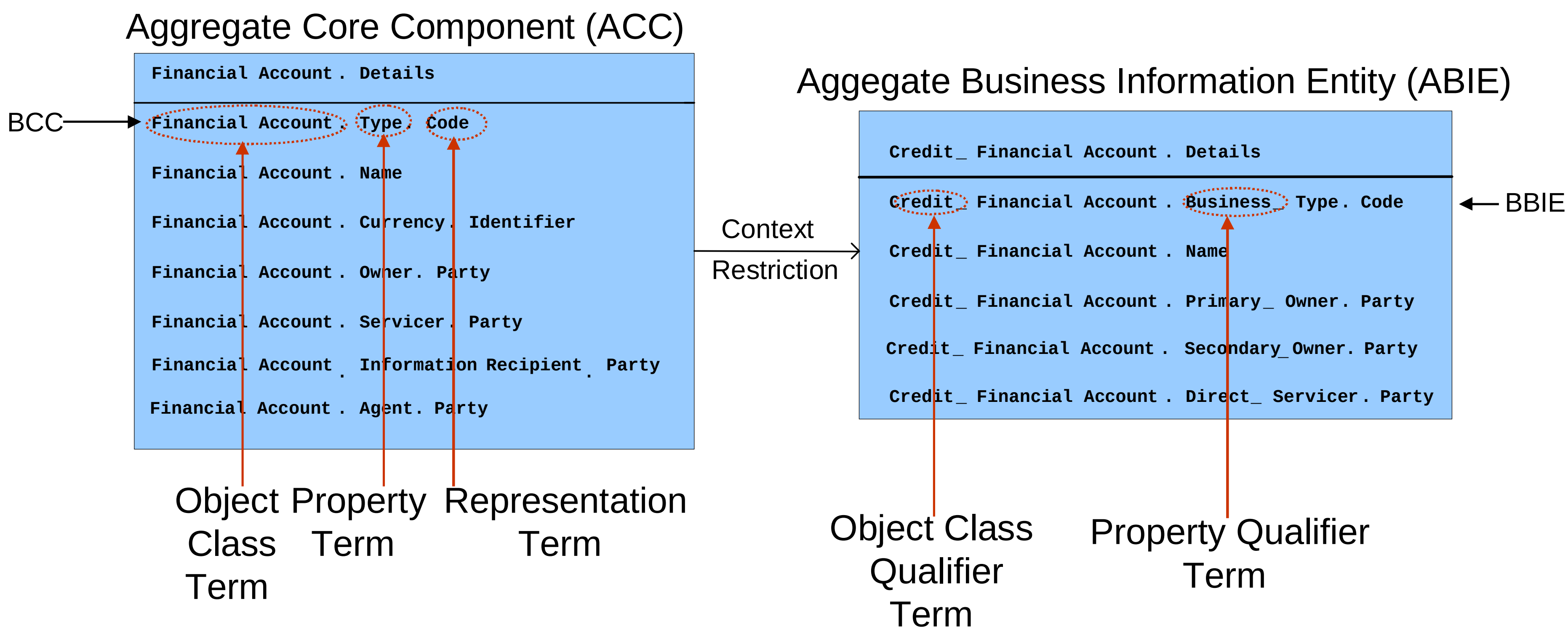
Amount	Ordinal
Binary Object	Percent
Code	Picture
Date	Quantity
Date Time	Rate
Duration	Ratio
Graphic	Sound
Identifier	Text
Indicator	Time
Measure	Value
Name	Video

Related Standards Projects



UN/CEFACT Core Components

Different Structural Modeling, Same Semantic Metadata



BCC = Basic Core Component

BBIE = Basic Business Information Entity

Related Standards Projects
and Prospects for More

Many different ways to structure data models

- But the same kind of semantic metadata can be used for all
- This provides new ways of federating and mapping the data models that these communities produce

Non-invasive approach

- No need to alter the approach to the structural modeling of data
- Add semantic metadata as an additional dimension
- E.g: Semantic metadata in XBRL could be in an extra linkbase

Retrospective markup of pre-existing format definitions

- Much of the work will consist of extracting and correlating domain dictionaries from existing data and message structures
- …and devising ways to markup the legacy data definitions with semantic metadata that specifies the object class term, property term, business context, etc.

Summary

Syntactic metadata has enabled model-driven data integration tools to generate data transformation code

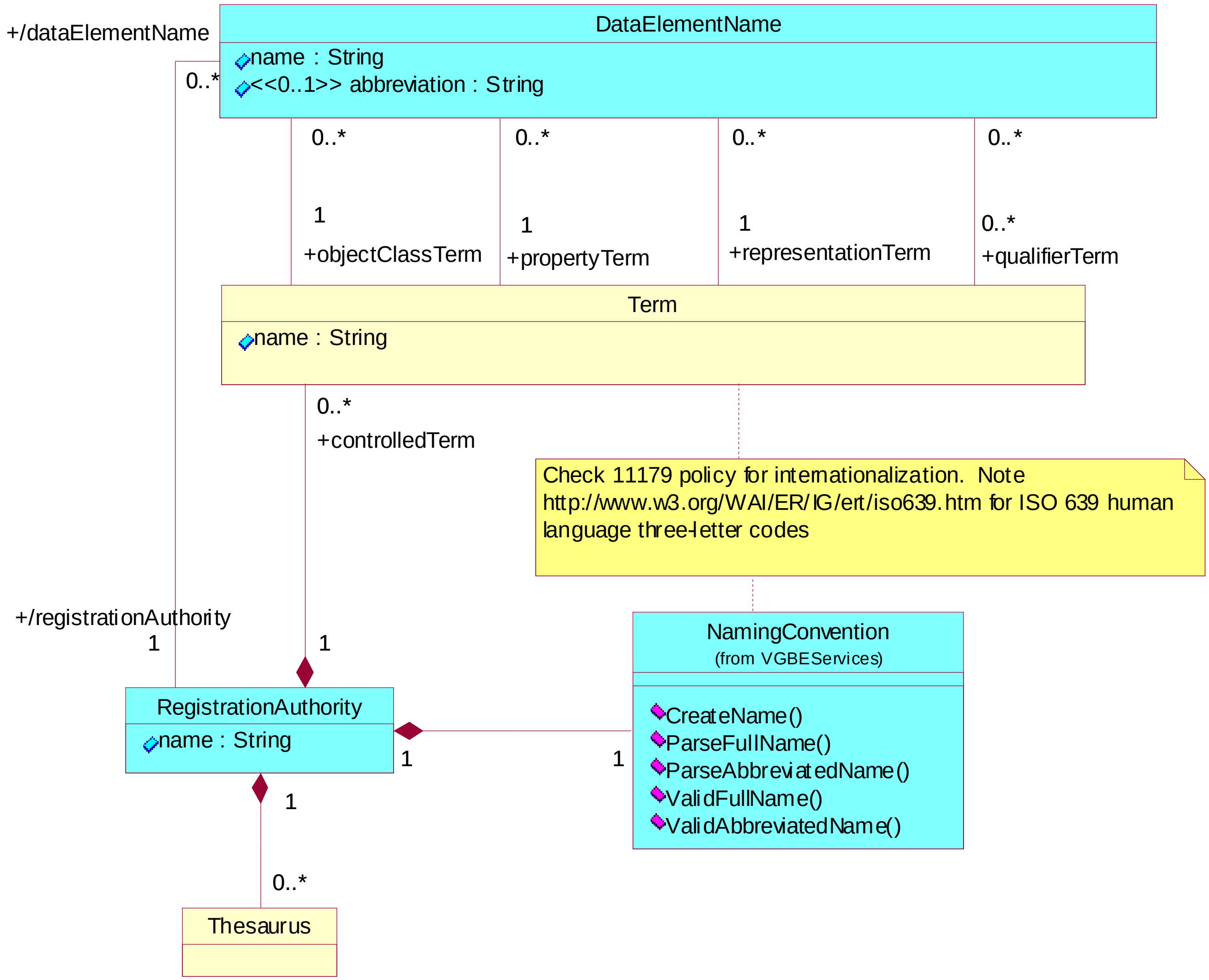
- One of the most significant achievements of model-based approaches
- But only solves 20% of the problem
- 80% of the problem is about semantic interoperability
- BPM and SOA face the same semantic interoperability issues

Semantic metadata

- Based on principles of language
- A link between semantic vocabularies/ontologies and IT data/message structures
- Modest improvements in semantic interoperability have large economic repercussions

Backup Slides

ISO 11179-5 Metamodel



This slide adapted with permission from Visa, Inc.

Business Association Elements

Navigable ends of associations are properties whose type is the Business Component on the navigable end

E.g., below, First Agent is a property of Transfer Instruction, and the type of the property is Financial Institution

These properties are *Business Association Elements*