Information Management Metamodel

Pete Rivett, CTO Adaptive
OMG Architecture Board

pete.rivett@adaptive.com
2011-05-11
The Information Management Conundrum

• We all have ‘Data’...do we have ‘Information’?
• Data is defined, modeled, stored in islands of technologies and exchanged via many incompatible formats
  – It is difficult to trace business Terms to models, XML Schemas, data stores (structured and semi/unstructured), applications, Services, processes...
  – Impact analysis of changes spanning across the life-cycle of Information and lines of business is even harder
  – Semantic nuances and geographically relevant variations of business Terms and rules is not often well documented and traceable
• Seamless integration across data life-cycle management tools and formats is flaky at the best
• All of the above issues result in poor quality Information leading to poor and costly business decisions
Common Warehouse Metamodel (CWM) - precursor of IMM

**Warehouse Management**
- Warehouse Process
- Warehouse Operation

**Analysis**
- Transformation
- OLAP
- Data Mining
- Information Visualization
- Business Nomenclature

**Resources**
- Object-Oriented (Object Model)
- Relational
- Record-Oriented
- Multi Dimensional
- XML

**Foundation**
- Business Information
- Data Types
- Expressions
- Keys
- Index
- Type Mapping
- Software Deployment

**Object Model**
(Core, Behavioral, Relationships, Instance)
What is IMM?

• UML has not been very data modeling friendly*!
  
  – Data modelers simply like the ‘Crow’s Feet’ even though a UML Class Diagram can be a good precursor of a ‘Conceptual Data Model’

• IMM, an OMG standard, is being developed to:
  
  – Facilitate modeling of structured and un/semi structured Information by standardizing traditional data modeling concepts and notations
  
  – Enable bi-directional traceability & lineage from business Terms to Object Oriented models, data models, XML Schemas, Ontologies, Services, Processes...
  
  – Exchange of Information life-cycle metadata across tools as XMI (an OMG standard)
  
  – Harmonization with other Information management related standards such as Semantics of Business Rules and Vocabulary (SBVR) and Ontology Definition Metamodel (ODM)

* As the data modeling community would like it to be!
Background

- OMG’s Common Warehouse Metamodel (CWM, http://www.omg.org/technology/cwm/) has been implemented by many vendors and used by ‘end users’ primarily for metadata interchange across tools
  - In use across data modeling, business intelligence, ETL tools and metadata repositories
  - The uptake has been somewhat hampered by CWM’s name – many of the potential uses of CWM have no connection with building or managing data warehouses. Hence the proposed name for the new standard is Information Management Metamodel instead of CWM 2.x.

- Over the years, many UML vendors and end users have expressed a desire to use UML for data and XML modeling and ended up defining their own tool-specific profiles for each
  - As a result, there is neither an accepted standard nor interoperability of models developed using such profiles/tools
  - IMM will become the bridge between UML, data and XML modeling ‘Islands’

- IMM will complement OMG’s business vocabulary and Ontology modeling standards by enabling consistent understanding, modeling and interchange of ‘Information’ across Business and IT
Business driven Information management: Nirvana State!

**Business**

- Define Business Concepts, Rules in *Natural Language*, targeted at the Business User
  - *Policy* is an *Agreement* between 2 Parties (*Seller*, *Buyer*)
  - *Agreement* is ‘*Policy*’ in Life and P&C Line of Business
    - For *Annuity* it is a ‘*Contract*’
    - *Policy* is ‘*inforce*’ if the *Policy* has not expired and......

**Information Architects, Modelers, DW/ETL, BI staff...**

- Import Business Vocabulary, Rules into data modeling tool that can Transform it into Conceptual model
  - Data modeler can further develop logical, physical, dimensional models... (automation of vocabulary to data models >> modeler productivity)
- Trace (bi-directionally) Business Concepts to data models, XML Schema elements, Core Components...
- Import Vocabulary/Rules, models into ETL, Data Quality, BI tools

**Application, Service development**

- Import Business Vocabulary, Rules to develop:
  - Use Cases
  - Class diagrams, ...
  - Services

---

**SBVR**

- Import into Ontology modeling tool
  - Develop Ontology, (output RDF, OWL...)
  - Vocabulary validation, reasoning...

**ODM**

- ODM

**IMM**

- UML, UPMS...
IMM Components Overview

- **InformationManagementMetamodel**
  - **Traceability**
    - Lineage(ETL)
    - Dependency
  - **ModelManagement**
    - Responsibility
    - Process
    - Deployment
    - Operations Recording

- **BusinessModeling**
  - Entity Relationship
  - OLAP
  - Information Visualization

- **TechnologyModeling**
  - Relational
  - ObjectOriented Database
  - XMLSchema
  - Multidimensional
  - Record
  - LDAP

- **IMMCore**
IMM for Entity Relationship Modeling (Illustrative)

Sample Logical Model

- ACCOUNT
  - account-id
  - account-type
  - account-open-date
  - account-review-date

- CHECKING-ACCOUNT
  - account-id (FK)
  - checking-balance
  - available-balance
  - per-check-change

- SAVINGS-ACCOUNT
  - account-id (FK)
  - savings-balance
  - interest-rate
  - interest-earned

- LOAN-ACCOUNT
  - account-id (FK)
  - original-loan-date
  - loan-interest-rate
  - current-loan-balance
IMM ER Metamodel - Attributes
IMM ER Metamodel - Relationships

[Diagram of IMM ER Metamodel showing relationships between EntityClass, Role, Relationship, Generalization, connectingRole, connectedEntity, role, relationship, specificEntity, generalEntity, ordered, and 0..* cardinality.]
IMM ER Metamodell - Constraints
IMM for Relational Database Design

Sample Model

- Student
  - StudentNumber: int <<PK>>
  - FirstName: char(20)
  - Surname: char(20)
  - AddressOId: int <<FK>>
  - PhoneNumbere: int
  - EmailAddress: char(50)
  - lives at 1

- Enrollment
  - 1 enrolled 1*
  - SeminarOId: int <<FK>>
  - Seminar: int <<FK>>
  - Assignment: int <<FK>>
  - MarkReceived: float

- WaitList
  - 0..1
  - SeminarOId: int <<FK>>
  - Seminar: int <<FK>>
  - Added: datetime

- Seminar
  - SeminarOId: int <<FK>>
  - CourseNumber: int <<FK>>
  - SeminarNumber: int
  - Term: int <<FK>>
  - LocationOId: int <<FK>>
  - TimeSlotOId: int <<FK>>
  - applies to 1
  - 0..1

- UniversityDB
  - stored procedures
  - averageMark(studentNumber: int): float
  - averageMarkInSeminar(seminarOId: int): float
  - studentsEnrolled(seminarOId: int): int
  - totalFees(studentNumber: int, term: int): float
  - waitListSize(seminarOId: int): int

- SeminarFee
  - SeminarOId: int <<FK>>
  - FeeNumber: int <<FK>>
  - FeeTypeCode: char(3) <<FK>>
  - Fee: float
IMM for XML Schemas

XML Schema Diagram Example
Mail Tracking System

Sample XML Schema
IMM Core

• Common concepts:
  – Attribute
  – Association
  – Thing Type

• Mix-in concept library:
  – Identifier specification
  – Reference
  – Reference Attribute Binding
  – Composition

• Technology-specific concepts:
  – Relational examples: Table, Row, Column, Primary Key, Foreign Key
  – XML Schema examples: Element, Complex Type, Sequence, Group
IMM Core - Fragment
IMM: Traceability and Lineage

• Traceability
  – The connection between information model elements.

• Lineage
  – Operational movement of data which may involve transformation.
    – For example, ETL links
  – Dependency (mapping)
    • The fact that changing one element has an impact on others.
    • Used in impact analysis
    • For example
      – Entities \(\rightarrow\) Tables
      – Tables \(\rightarrow\) Software
      – Versions
IMM Traceability Metamodel
IMM: Lineage Metamodel
• Each open rental must be **guaranteed by a credit card that is held by the renter who is responsible for the rental**
• the **actual pick-up date/time of an open rental must occur before the scheduled return date/time of the rental**
• It is prohibited that a barred driver is **a driver of a rental**
• It is obligatory that each driver of a rental has a **valid driver license at pick-up date-time of the rental**
• It is prohibited that a rental car is driven by a **driver who is intoxicated**

Vocabulary written using Business Vocabulary Modeling Language (SBVR)
Work Remaining

• Streamline wording
  – Make more spec-like
• Complete LDAP metamodel and profile
• Extend worked example to ODM integration