META-MODEL FOR MASTER DATA

Baba Piprani
Suneil Dham
Canada
Your Presenter: Baba Piprani
MetaGlobal Systems, Canada

- Over 30 years experience teaching/implementing Object Role Modelling and implementing 100% rules in SQL DBMSs
- Canadian Delegate in ISO SQL and Metadata Standardization in ISO/IEC JTC1 SC32 since 1986, and in the ISO WG on Conceptual Schemas(TR9007)
- Canadian Delegate in ISO SOA (Service Oriented Architecture) Working group in ISO/IEC JTC1 SC38
- Past Chair - ISO Conceptual Schema Modelling Facilities WG
- Past Chair - Govt. of Canada Treasury Board IT Standards (TBITS SQL)
- Author of several publications
- Highly successful DW and metadata based web implementations using Model Driven Architecture using SOA principles
Abstract

The term ‘Master Data’ brings up different interpretations and connotations, especially with vendors espousing the “single version of the truth”. Is there a single version of the truth? Practical realities suggest that we need to continue to live with existing versions – at least until the “single version” is reached. So how do we handle the co-existence of “multiple versions” of the truth?

This paper examines a metamodel that defines what master data is, the types of master data, and criteria for determining master data. The metamodel visits master data administration, privileges and master data services including change management, metadata migration and data migration for master data is also reviewed.

The impact of changes to an enterprise going the master data route, along with the emphasis to influence future management direction is also examined. Several administration aspects make use of established ISO standards in metadata IS:11179-3 and IS:19763.
Agenda

1. Role of Metadata
2. Master Data
3. Master Data Process Model
4. Walkthrough of Semantics and definitions involved in selected Processes
5. Overview of semantics in remaining Processes
6. Conclusion
Role of Metadata...

• Data Processing, Web Services and Electronic Data Interchange rely heavily on accurate, reliable, controllable and verifiable data recorded in databases or in some persistent state

• prerequisite for correct and proper use and interpretation of data is that both users and owners of data have a common understanding of the meaning and representation of the data (Helsinki Principle)
'Helsinki Principle' from ISO TR9007

Note: These utterances are to be interpreted (recursively) as international English utterances:

“Any meaningful exchange of utterances depends upon the prior existence of an agreed set of semantic and syntactic rules. The recipients of the utterances must use only these rules to interpret the received utterances, if it is to mean the same as that which was meant by the utterer.”
Metadata and Standards...

• A number of characteristics, or attributes, of the data have to be agreed upon and defined common understanding towards implementation

• i.e. “metadata”, that is, “data that describes data”

• ISO/IEC 11179-3 Metadata Registry standard provides for the attributes of data elements and associated metadata to be specified and registered as metadata items in a Metadata Registry
Types of Metadata

• Two types of metadata
  – technical metadata and
  – business metadata

• Technical metadata
  – used by administrators and software tools
  – provides the technical descriptions of data and operations
  – includes information about source data, target data, and the rules that are used to extract, filter enhance, cleanse, and transform source data to target data
Types of Metadata

• Technical metadata could be created by a relational database management system, by data warehouse and transformation tools, or by the data warehouse operations personnel

• Examples of these include database statistics, descriptions of transformations, scheduling
Types of Metadata

• Business metadata
  – used by business analysts and end users
  – provides a business description of informational objects
  – assists end users in locating, understanding, and accessing information in the data mart, data warehouse, or other informational sources
  – Examples include calculation used to create a particular value, the data and time a report was created, or a description of the approval status of the projected forecast.
So Let’s get to Data, and Master Data...

- Business is concerned with the ‘truth’ value of the data it operates with
- Data is assumed to be sacrosanct since our friendly and reliable computer system is giving it to us, right?
- Wrong! The truth of the matter is that in more than 50% of the cases, data quality and data integrity of the business data is questionable, see Gartner
- The search goes on...for the ‘single version’ of the truth.
Enter Master Data Management...

• refers to data relevant to the conduct of the business on which transactions or analyses are based
• master data is non-transactional data about people, things, places, and concepts
• Creating and maintaining good quality master data has become a necessity, and actually is critical
Master data ...

• basically two kinds of master data
  – common reference data - like types and categories pertaining to the properties or characteristics of data (e.g. Country code, Address Type)
  – ‘master file’ type data like, customers, clients, vendors, products, etc...essentially data that the organization uses for tracking through transactions
Master Data Model....

• Establish a model for Master data:
  – Define scope via a Process Model Decomposition
  – Define ORM schemas as applicable for managing Master Data

________________________________________

Process Model Decomposition
  – highlights typical business processes involved with going the Master Data route
Process Model Decomposition

Baba Piprani, Suneil Dham: Meta-model for Master Data
ORM Workshop, Crete, Greece, 2010
Real world usage....

• Process model is borrowed from current user experience involved in establishing and stabilizing:
  – master data hub and
  – associated master data hub services to address Service Oriented Architecture (SOA) requirements
P1 - Manage Master Data Definition...
Master Data Definition using a subset of ISO 11179-3

Metadata Registry....
Paper Coverage Contents and Acknowledgement

• this paper discusses in detail the Business Processes P1, P2 and P3 in Figure 1, for which ORM/NIAM metamodels are provided
• notation used in the ORM/NIAM schema for Fact Types is using Fact Type Identifiers and omits role names for the purpose of making the diagrams less crowded due to space limitations in this publication.
• Thanks to Ooriane Corp. for the use of Ooriane Designer and Ooriane Semantic Analyzer for ORM/NIAM graphics
Baba Piprani, Suneil Dham:
Meta-model for Master Data
ORM Workshop, Crete, Greece, 2010
<table>
<thead>
<tr>
<th>#</th>
<th>Object</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Name</td>
<td>Business term name in its unabbreviated form.</td>
</tr>
<tr>
<td>1.2</td>
<td>Identifier</td>
<td>Any user defined data attribute identifier used for reference purposes</td>
</tr>
<tr>
<td>1.3</td>
<td>Version</td>
<td>Version number of metadata definition.</td>
</tr>
<tr>
<td>1.4</td>
<td>Version Date</td>
<td>Version date of the version number</td>
</tr>
<tr>
<td>1.5</td>
<td>Abbreviation</td>
<td>Recommended short form of the attribute name using standard abbreviations, if any.</td>
</tr>
<tr>
<td>1.6</td>
<td>Acronym</td>
<td>Recommended acronyms of the attribute name</td>
</tr>
<tr>
<td>1.7</td>
<td>Associated Concepts</td>
<td>Perspective, abstract idea or another attribute with which this data attribute is associated with, or is a property of, or can be mapped to.</td>
</tr>
<tr>
<td>1.8</td>
<td>Classification Scheme</td>
<td>Descriptive information for a generic level arrangement of objects into groups based on characteristics that the objects have in common.</td>
</tr>
</tbody>
</table>
### Object Definition

2.1 Life Cycle Status

Definition: Condition of the metadata in its lifecycle (draft, proposed, approved, retired).

2.2 Replaces

Definition: Name of attribute (old attribute), which is being replaced by this attribute.

2.3 Replaces on date

Definition: Date of replacement for the old attribute.

2.4 Replaced by

Definition: Name of attribute (new attribute), which replaces this attribute.

2.5 Replaced by on date

Definition: Date on which this attribute was replaced.
<table>
<thead>
<tr>
<th>#</th>
<th><strong>Object</strong></th>
<th><strong>Definition</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Data Stewardship Group</td>
<td>The organization(s) responsible for the accuracy of the attribute’s definition.</td>
</tr>
<tr>
<td>3.2</td>
<td>Data Steward</td>
<td>The person(s) responsible for the attribute meta-data definition (name, contacts, definition, business rules).</td>
</tr>
<tr>
<td>3.3</td>
<td>Source of Reference</td>
<td>Reference number and/or title of an adopted data/metadata standard from an external source.</td>
</tr>
<tr>
<td>3.4</td>
<td>System or Reference of</td>
<td>The manual or automated system that serves as the authoritative source for accurate data values.</td>
</tr>
<tr>
<td></td>
<td>Record</td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>Custodian</td>
<td>The person(s) responsible for the maintenance and quality of the actual data in the system of record.</td>
</tr>
<tr>
<td>#</td>
<td>Object</td>
<td>Definition</td>
</tr>
<tr>
<td>----</td>
<td>-----------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>4.1</td>
<td>Definition</td>
<td>The textual description of the attribute as per ISO/IEC 11179-4, see [5]</td>
</tr>
<tr>
<td>4.2</td>
<td>Source of Definition</td>
<td>The publication, directive, standard, system, organization or person(s) responsible for developing the attribute definition.</td>
</tr>
<tr>
<td>4.3</td>
<td>Comments</td>
<td>Additional information to aid users in understanding the purpose and use of attribute.</td>
</tr>
<tr>
<td>#</td>
<td>Object</td>
<td>Definition</td>
</tr>
<tr>
<td>-----</td>
<td>-------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>5.1</td>
<td>Exchange Display Role</td>
<td>The data element representation mode whether for data exchange or data entry/display formatting purposes</td>
</tr>
<tr>
<td>5.2</td>
<td>Size</td>
<td>The maximum &amp; minimum allowable lengths for the raw data.</td>
</tr>
<tr>
<td>5.3</td>
<td>Data Type</td>
<td>The kind of data. Examples are: alphabetic, binary, numeric, alpha-numeric</td>
</tr>
<tr>
<td>5.4</td>
<td>Case Sensitivity</td>
<td>A specification of whether or not the data is to be upper, lower, or mixed case.</td>
</tr>
<tr>
<td>5.5</td>
<td>Format</td>
<td>A specification of the way the raw data should be arranged.</td>
</tr>
<tr>
<td>5.6</td>
<td>Scale and Precision (if numeric)</td>
<td>The total number of digits and positioning of the decimal point if applicable.</td>
</tr>
<tr>
<td>5.7</td>
<td>XML Tags</td>
<td>Specific XML (Extensible Markup Language) tags associated with attribute</td>
</tr>
<tr>
<td>5.7</td>
<td>Example</td>
<td>Provide example(s) of valid raw data and descriptor, if appropriate.</td>
</tr>
<tr>
<td>#</td>
<td><strong>Object</strong></td>
<td><strong>Definition</strong></td>
</tr>
<tr>
<td>-----</td>
<td>------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| 6.1 | Business Rules  | Contains rules that are business subject area specific and are used to identify the business rule data violators to provide data quality and integrity. Integrity violations on the data will be reported on using this category classification. The corresponding error reports will be forwarded to the data stewards.  
  The Business Rules for definition call for:  
  - Cardinalities of relationships (one-to-zero/one/many, many-many, recursive, parent-child update/delete rules)  
  - Value domains  
  - Super/sub-type  
  - Nullability  
  - Derivations (computed values)  
  - Unique constraints for business keys (not surrogate keys)  
  - Other business rules  
  The Business Rules for Execution call for:  
  - Handling exception conditions  
  - Notification  
  - CRUD sequencing (e.g. can it recognize cardinality rules e.g. insert supertype first then subtype) |
Manage Master Data Administration (P2)
Manage Master Data Administration (P2)

• accountability and responsibility for managing master data administration
• Infrastructure administration
• Stewardship and Trustee
  – Roles and Responsibilities
  – Governance Model
  – Master Data Administration (Models and Data)
• Metadata Management (OSFI Form 19, a subset of ISO 11179-3 Metadata Registry Standard and ISO 19763 Metamodel for Interoperability Standard)
Change Management (P3)

Baba Piprani, Suneil Dham: Meta-model for Master Data
ORM Workshop, Crete, Greece, 2010
Change Management (P3)

- Impact assessment (e.g. if the data model for master data is changed, what systems are affected)
- Lineage management (e.g. what systems (applications, databases, reports) are using the data)
- Version Control (e.g. versioning of the master data data model)
- Migration (e.g. migration of a new master data model version through deployment environments like Development/Test/User Acceptance/Production as well as data)
Master Data Services (P4)
Master Data Services (P4)

- Interface types (SQL, Web Services...)
- Create/Retrieve/Update/Delete/Search
- Federated or Consolidated master data hub approaches
- Managing Referential Integrity between master data and any referencing data
- Service description and discovery
- Service presentation
- Services interface definition
- International Language support
Privileges (P5)
Privileges (P5)

- Define Privileges
- Manage Privileges
- Manage Roles
Migration of Metadata and Data (P6)
Migration of Metadata and Data (P6)

- Existing schemas and data to new Master Data registry
- Existing data models
- Existing business rule declarations
- Existing Excel or other spreadsheet ad hocs
- Parallel versions of ‘multiple versions of the truth’ i.e. the master data version as the single version and, the local legacy version accompanied with mappings during transition, etc.
Conclusion

• Migrating towards Master Data needs a lot of consideration and planning
  – data quality issues,
  – establishing semantic equivalence,
  – co-existence/transitioning ....

• Necessitates a SOA type approach for using and managing Master Data hub

• Integrating data for master data involves data harmonization, data cleansing and governance...and development of common services for legacy systems...

• Be patient....Don’t expect magic to happen overnight.....😊