The Role of Master Data Management (MDM) for Data Centric Automation Strategies

Tracy J. McLane, Engineering Virtual Project Delivery (VPD) Lead
Bechtel Infrastructure
@tjmclane
Tracy McLane has been a Bechtel Distinguished Scientist since 2012, and served as the GIS Technical Discipline Lead for the company from March 2007 to February 2018. In this role, she was responsible for the development of the GIS Technical Center of Excellence, the Enterprise GIS database architecture and a centralized GIS Knowledge Bank for the Bechtel.

Ms. McLane served as the Engineering Virtual Project Delivery (VPD) Lead for the 4D, 5D Program for the past year and a half, with some of her personal focus being on 3D GIS and BIM integration and the development of Master Data.

Ms. McLane has worked with Oracle Spatial technology for over 16 years, and currently serves as the Chair of the Independent Oracle Users Group (IOUG) Oracle Spatial and Graph Special Interest Group (SIG) Board. Ms. McLane has worked in the GIS industry for more than 24 years, with past GIS experience at the Tennessee Valley Authority and the U. S. Department of Energy Savannah River Site Ms. McLane now works in Bechtels Headquarters in Reston, Virginia.

Ms. McLane holds a Master of Science degree in Geography from the University of Tennessee in Knoxville, Tennessee and a Bachelor of Arts degree in International Business from Eckerd College in Saint Petersburg, Florida.
Agenda – Oracle Spatial and Graph

1:30 pm  “The Role of Master Data Management for Data Centric Automation Strategies” – Tracy McLane

1:55 pm  “Measuring Public Citation using Graph Modeling and Analysis” – Keisuke Honda & Mio Takei

2:30 pm  “Spatial Analytics with Spark and Big Data”
          – LJ Qian & Siva Ravada

3:40 pm  “Using GeoJSON in the Oracle Database”
          -Albert Godfrind
Objectives - A Spatial Person’s Journey to Graph

- Overview of Master Data Selection and Cataloging
- Classification of Master Data Types
- Development of MDM Data Dictionary
- Automation for MDM Structure Creation and Content Migration
- Comparison of RDF View vs. Native Semantic Model vs. Virtual Semantic Model
- Automation for Semantic Model Architecture Creation and Enrichment
- User Stories for Spatial: Comparison of Traditional RDBMS MDM Uses and Automation for Identification of Future Use Cases for Graph:
  - Data Normalization
  - BIM Classification and Other MDM Hierarchies for Spatial Data Classification
  - MDM Data Enrichment and Interoperability of Spatial Data
What Is Master Data Management?

- Master Data Management (MDM): comprehensive methodology that raises the accuracy, quality and consistency of our information (data and systems).
- Essential for organizations transitioning to data centric execution and enterprise systems
- Effective enterprise MDM provides:
  - Data Governance
    - Policy and Rules
  - Data Quality
    - Constraints and Standards
  - Integration
    - Cross Functional Relationships

Oracle MDM RDBMS Decode Example:

```
DECODE
(<spatial table name>..<column name>,
<domain table>..<domain value column>,
<domain table>..<domain description column>,
<default value>)
<column alias>
```
Bechtel VPD
Master Data Management: Types and Organization

RDF Views
(RDF Content in Views with only metadata stored as RDF)

Constraint MDM
(Value List and Range Domains)
D_MDM_TYP: MDM_COLMN

Rule-Based MDM
(Classification and Data Exchange Criteria)
D_MDM_TYP: MDM_CLS

Software MDM
(Data Dependencies for API and Software Development)
D_MDM_TYP: MDM_SFT

Topology MDM
(Spatial Rules)
D_MDM_TYP: MDM_TOPOL

Relationship MDM
(Graph and RDB Types)
D_MDM_TYP: MDM_RELAT

Structured Data

Domains Types
D_DOMAIN_TYP
List Domains
D_DOMAIN_LIST
Range Domains
D_DOMAIN_RANG

Rules Types
D_RULE_TYP
Classification Rules
D_RULE_CLASSIF
Specification Rules
D_RULE_SPEC
Rules of Credit
D_RULE_CREDIT
BOM Quantity Rules
D_RULE_BOM

Software Domains
D_SW_DOM
Rule Groups by Software
D_RULE_GRP
Rule Group Membership by Software
D_RULE_GRP

Topologies Types
D_TOPOLOGY_TYP
Topology Subtypes
D_TOPOLOGY_SUB

RDF Views
D_RDF_VIEWS

RDF Standard Properties
D_RDF_PROP

Ontology Types
D_ONTOLGY_TYP
RDF Naming Conventions
D_RDF_NAM
RDF Rules Conventions
D_RDF_RULE

Source: Enterprise Database Design Developed by T. McLane in combination with original SDSRE v. 2.6 design concepts, plus additional Bechtel-specific content, metadata, MDM architectural breakdowns/hierarchies, relational/RDF view concepts and functional requirements for Engineering VPD Jobs-to-be-Done (originated: August 11, 2017)
Spatial and MDM Data Governance

Bechtel MDM Data Governance Authority

- Design Authoring: Data Originator
- Design Agency: Data Custodians
- Design Authority: Data Stewards (SMEs) and IS&T
- Data Trustees: MDM Council (Membership of Organizational Trustees)

Spatial Data Governance:
- Process and governing roles were developed at Savannah River Site (SRS) for GIS Data with the involvement of many disciplines, stakeholders and data owners (Bechtel and Westinghouse)
- Bechtel Corporate now in the process of a data governance structures for Master Data

Source: Role Names developed at Savannah River Site, with above definitions originally developed by T. McLane for publication in the 3DG-K04-00030: SPATIAL DATA MANAGEMENT IN BECHTEL'S ENTERPRISE GIS DATABASE, Rev. 000, 2010 December 9
Mastering Engineering Data: Content Authoring Process

Figure 1: BIM Classification Hierarchy
Uniformat Level 4 Master Data

A1010 - 10
A = Substructure
A10 = Foundations
A1010 = Standard Foundations
A1010.10 = Wall Foundations

@tjmclane
MDM Semantic Modeling with RDF Views: Oracle Spatial and Graph Process
Oracle RDF Semantic Object Metadata

- All Semantic Models: $mdsys.sem_model$
- Virtual Semantic Models: $mdsys.sem_vmodel_info$
- Unique Virtual Semantic Model View: $mdsys.semu_$
- Duplicate Virtual Semantic Model View: $mdsys.semv_$
- Rulebase View Columns Prefix: $mdsys.semr_$
- Rulebase Information: $mdsys.sem_rulebase_info$
- Rulebase View Prefix: $mdsys.rdfr_$
- Semantic Entailments (Indices): $mdsys.sem_rules_index_datasets$
- Semantic Entailment View Prefix: $mdsys.rdfi_$
- Semantic RDF Triples: $mdsys.rdf_value$_
Oracle Graph RDF Views: Mapping Triples

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>PREDICATE</th>
<th>OBJECT</th>
<th>GRAPH</th>
</tr>
</thead>
<tbody>
<tr>
<td>r:2511a1833C0CA143A4167n36</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n37</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a9786887AB2D0B05n20</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a9786887AB2D0B08n421</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n76</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n77</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n112</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a333C0CA143A4167n113</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n60</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n61</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a9786887AB2D0B05n423</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n43</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n71</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n74</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n101</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n102</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a9786887AB2D0B06n39</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n49</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n59</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n50</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a9786887AB2D0B08n411</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n45</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n135</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n137</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n38</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n39</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n99</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n98</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n130</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n131</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a9786887AB2D0B08n415</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n135</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n451</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n455</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n67</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n67</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
<tr>
<td>r:2511a1833C0CA143A4167n138</td>
<td><a href="http://www.w3.org/ns/r2ml#ObjectMap">http://www.w3.org/ns/r2ml#ObjectMap</a></td>
<td>r:2511a1833C0CA143A4167n138</td>
<td><a href="http://xmlns.oracle.com/oc4rdf/r2graph">http://xmlns.oracle.com/oc4rdf/r2graph</a></td>
</tr>
</tbody>
</table>
Oracle Graph RDF Views: Schema Triples
sem_api.create_rdfview_model (  
    model_name => <\textit{model name}>,  
    tables => SYS.ODCIVarchar2List (<\textit{table 1}>,<\textit{table 2}>),  
    prefix => <\textit{URL prefix}>,  
    options => <\textit{priority option: Conformance or Key-based}>)  

*Note: Input tables are RDBMS content, not triples

@tjmclane
Triples Tables: Staging Tables

SEM_APIS.BULK_LOAD_FROM_STAGING

```sql
semApis.bulk_load_from_staging_table (  
    model_name => <model name>,  
    table_owner => <table owner>,  
    table_name => <staging table name>,  
    flags => <optional flags>  
)
```
Triples Tables: Staging Tables for RDF Views

SEM_APIs.EXPORT_RDFVIEW_MODEL

sem_apis.export_rdfview_model (<model name>,
<project schema>,
<output staging table name>,
NULL)

*Note: The sem_apis.export_rdfview_model function supports both formats of staging tables

@tjmclane
Oracle Triples Tables: Native Semantic Models

```
sem_apis.create_sem_model(
  model_name => <model name>,
  table_name => <table name>,
  column_name => <name of Triple column>,
  model.tablespace => <model tablespace>)
```
Combining Content into a Virtual Semantic Model

**SEM_APIS.CREATEVIRTUAL_MODEL**

**Data Type Input:**
- Models
- Rulebases
- Entailments (Indices)

**Advantages:**
- Single point for SEM_MATCH queries
- Can combine content of native semantic models and RDF view exported content
- Supports integration of RDBMS and services based content

*Note: RDF View content is not supported directly, but can be integrated from view content exported to staging tables and rewritten as SDO_RDF_TRIPLE_S triple content*
MDM-Based Pattern Definition Exposure

Masterformat Hierarchy

Cytoscape Edge Definitions

Primavera Unifier Parent-Child

@tjmclane
Use Case 2: BIM Classification System Hierarchy
Master Data for 3D Model Object Data

Uniformat BIM Classification Hierarchies for HVAC Objects

Masterformat BIM Classification Hierarchy

Omniformat Table 22 Hierarchies for Electrical Objects

Uniclass Hierarchies for Bridge Objects

Rail Station Model: 113,601 objects, 1,376,276 parameters

Rail Station Model: 8,612 objects, 141,593 parameters

A = Substructure
A10 = Foundations
A1010 = Standard Foundations
A1010.10 = Wall Foundations

A1010.10 - 10

Uniformat BIM Classification Hierarchy
Master Data for Semantic Rulebase Creation

SEM_APIS.CREATE_RULEBASE

Rulebase Types

Rulebase Subtypes

Rulebase Syntax

@tjmclane
Use Case 1: Master Data (Relational and Graph) for 3D Spatial Object Data

- Revit Cable Tray Database Objects with decoded MDM Hierarchy
- Descriptions in the Spatial Views (only codes on the object)

Milestone

- Process
- Rule

Revit Code = -2008130

Material = ALUMINUM

Installation = ABOVE_GROUND

MDM-based Data Rules:
- Parameter (Schema) Normalization Rule
- Constraint (Domain) Rules

MDM RDF-based Rules:
- MDM Classification Hierarchy Rule
- Rulebase Constructs

<table>
<thead>
<tr>
<th>LDOMAIN</th>
<th>DOMAINNAME</th>
<th>DOMAINVALU</th>
<th>DOMAINDESC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:allValuesFrom</td>
<td>OWLXML Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:complementOf</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:differentFrom</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>rdfs:domain</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:equivalentClass</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:equivalentProperty</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:functionalProperty</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:hasValue</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:inverseFunctionalProperty</td>
<td>RDF++ Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:inverseProperty</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:inverseOf</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>rdfs:range</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
<tr>
<td>BDM02-0005137-2016</td>
<td>Oracle Spatial and Graph Rulebase Constructs</td>
<td>owl:sameAs</td>
<td>OWLPrime Vocabulary Constructs</td>
</tr>
</tbody>
</table>
Master Data for Semantic Rulebase Creation

Ontology Types

 Ontologies Supported in Oracle Spatial and Graph (out-of-the-box)

Rules

 @tjmclane

Analytics and Data Summit 2018
Use Case 3: Spatial Data Interoperability & MDM Enrichment: Autodesk Civil 3D Cube Solids

Breakdown of Cube Solid Geometries and Regenerate in Edges as Enterprise Oracle Spatial Database

ST_GEOMETRY Objects:

1) Interoperable with other Client Software
2) Available for In-Database Analytics (example: Centroid)
3) Automated Geographic Work Breakdown Structure
4) Interoperable with other tools (shown in Esri ArcScene)
5) Exposure of MDM-Enriched Data back in original authoring software

Figure 1: Civil 3D Autodesk Cube Edges as Oracle Spatial 3D Lines and Vertice 3D Points shown in Esri ArcScene (edge linework used as input for mdsys.ST_GEOMETRY objects)

Note: Oracle Spatial 3D Polygon Faces examples in red also displayed, but these objects do not directly interoperate back to Civil 3D (but export to Esri Shapefile from Oracle Spatial does interoperate 3D polygon cube faces)
Use Case 3: Master Data Enrichment and Automated Tag Generation: Civil 3D Solid Cube Edges as ST_GEOMETRY

Current State: Architected Spatial Views Exposing MDM and Data Patterns

Desired Future State: Spatial Views become input for Graph
Recognized Data Patterns and Automatic MDM Association
Key Takeaways: Spatial Master Data

• Primary priority should be on identification of important master datasets needed for spatial content
• Second priority should be on the revision process to update content to keep current with the industry publication and to normalize content between MDM datasets (i.e. – data governance process)
• Final priority should be on deployment strategies (both relational and graph) needed to support end user experience
• Master data architecture should have both data dictionary and metadata components to facilitate automation
• Definitions for data quality should include a MDM basis for data completeness, content standards and data transformation
Key Takeaways: Use of RDF Views

- **RDF Views** are a way to expose semantic triples data from traditional relation database content.
- **RDF Views** create: mapping triples, schema triples, rdf object triples.
- **RDF View** triples content can be ported to staging and triplestore tables for the creation of **Native Semantic Models**.
- Rulebases and entailments can be created on the native semantic models exported from RDF Views.
- Semantic matching (SEM_MATCH) queries work against both **RDF Views** and traditional **Native Semantic Models**, BUT…
- Perhaps a hybrid of both RDF Views and native semantic models content combined in a **Virtual Semantic Model** will produce the best results…my comparisons of the three are now in progress.
Questions & Answers