Unit 2 – A training on the PODS 7 data model and component databases
Intended Audience

- GIS/IT professionals
- New to pipeline industry
- Little or no exposure to PODS

PODS Training – both PODS Basics and PODS Advanced – create a better understanding of PODS Standards and PODS implementations through geospatial and relational database applications.
2 Unwrapping PODS 7

An introduction to the PODS 7 Data Model and component databases
Webinar Series Overview

- **Unit 1** – Introduction to PODS 7
- **Unit 2** – Unwrapping PODS 7, Part 1
- **Unit 3** - Unwrapping PODS 7, Part 2
- **Unit 4** – Modifying & Extending PODS
- **Unit 5** – Introduction to Utility Network
- **Unit 6** - Information and Analysis – Risk and Reports
Introduction

Data Model

Database

Tables

Pipeline Open Data Standard
www.pods.org
2 Unwrapping PODS 7

- The PODS 7 Schema
- Contrasting to Legacy
  - Review of PODS data model terms and key concepts
  - What is different about PODS 7
- What does PODS 7 contain? Rolling out parts of it – demonstrating
- Databases – essential building blocks. PODS Data Standards and using the data dictionary to meet the standard. Meta data, business rules, units of measure
Our Goals for this Unit

1. Become familiar with the PODS 7 data model.
2. Learn general data management terms and concepts relating to PODS 7.
3. Differentiate between PODS Conceptual, Logical, and Physical models.
4. Exploring what sets PODS 7 apart from previous versions.
How would you convert this into an information system?
How do you design a container for all this mission-critical data?

- Geographic boundaries
- Compression
- Physical pipeline facilities
- Regulatory compliance
- Operating measures
- Geographic feature crossings
- Inline inspection
- Risk assessment
- Work history
- Site facilities
- External documents, reports
- Close interval surveys
- Offshore lines
- Cathodic protection
- Leak survey
It’s not just about the container...

It’s about the DATA itself.

• Information is only as good as the DATA it’s built on.
• Data must be organized
• Data yields information.
• Data determines the TRUTH
• Data through analytics derives MEANING from data
How do you design a container for all this mission-critical data?

- Geographic boundaries
- Compression
- Physical pipeline facilities
- Regulatory compliance
- Operating measures
- Geographic feature crossings
- Inline inspection
- Risk assessment
- Work history
- Site facilities
- External documents, reports
- Close interval surveys
- Offshore lines
- Cathodic protection
- Leak survey

What data?

Authoritative?

Organization?

With logical groups of tables!

Data Quality?

Governance?

Derivatives?

You need a data model.
Relational Databases

Because of relational databases entire pipeline systems and ancillary data are managed within the PODS 7 data model.
A Database is a collection of related data stored in **Tables**

- Tables are organized into Rows (records) and columns (fields)
- Databases also store prescribed rules, roles, and relationships for data within it
- Controlled by a DBMS (Database Management System)

**Records**

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Operational Status</th>
<th>Type</th>
<th>System Type</th>
<th>Interstate</th>
<th>Product ID</th>
<th>OBJECTID</th>
</tr>
</thead>
<tbody>
<tr>
<td>711T-A-T-T-T</td>
<td>71st St Lat 36°</td>
<td>ACTIVE/IN-SERVICE</td>
<td>TRANSMISSION</td>
<td>TRANSMISSION</td>
<td>YES</td>
<td>(06455000-19A4-4219-62625)</td>
<td></td>
</tr>
<tr>
<td>LEID-A-T-T-T</td>
<td>Leidy Line “A”</td>
<td>ACTIVE/IN-SERVICE</td>
<td>TRANSMISSION</td>
<td>TRANSMISSION</td>
<td>YES</td>
<td>(06455000-19A4-4219-62625)</td>
<td></td>
</tr>
<tr>
<td>LEID-B-T-T-T</td>
<td>Leidy Line “B”</td>
<td>ACTIVE/IN-SERVICE</td>
<td>TRANSMISSION</td>
<td>TRANSMISSION</td>
<td>YES</td>
<td>(06455000-19A4-4219-62626)</td>
<td></td>
</tr>
<tr>
<td>LEID-C-T-T-T</td>
<td>Leidy Line “C”</td>
<td>ACTIVE/IN-SERVICE</td>
<td>TRANSMISSION</td>
<td>TRANSMISSION</td>
<td>YES</td>
<td>(06455000-19A4-4219-62627)</td>
<td></td>
</tr>
<tr>
<td>LIEGA-T-T-T</td>
<td>Long Island Ext Lat 26°</td>
<td>ACTIVE/IN-SERVICE</td>
<td>TRANSMISSION</td>
<td>TRANSMISSION</td>
<td>YES</td>
<td>(06455000-19A4-4219-62629)</td>
<td></td>
</tr>
<tr>
<td>MAIN-A-T-T-T</td>
<td>Mainline “A”</td>
<td>ACTIVE/IN-SERVICE</td>
<td>TRANSMISSION</td>
<td>TRANSMISSION</td>
<td>YES</td>
<td>(06455000-19A4-4219-62631)</td>
<td></td>
</tr>
<tr>
<td>MAIN-B-T-T-T</td>
<td>Mainline “B”</td>
<td>ACTIVE/IN-SERVICE</td>
<td>TRANSMISSION</td>
<td>TRANSMISSION</td>
<td>YES</td>
<td>(06455000-19A4-4219-62632)</td>
<td></td>
</tr>
</tbody>
</table>

**Fields**

Relationships between tables reflect the association of objects in reality.
# Data Dictionaries

Using Supplied Data Dictionaries help you define your data and meet the PODS 7 standard.

## Data Dictionaries

<table>
<thead>
<tr>
<th>Class Attributes</th>
<th>Type</th>
<th>Length</th>
<th>Precision</th>
<th>Scale</th>
<th>Allow Nulls</th>
<th>Default Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PL_COMPANY_NAME</td>
<td>esriFieldTypeString</td>
<td>75</td>
<td>0</td>
<td>0</td>
<td>false</td>
<td></td>
<td>The full name of the company for a given record.</td>
</tr>
<tr>
<td>PL_COMPANY_ACRONYM</td>
<td>esriFieldTypeString</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>true</td>
<td></td>
<td>The acronym of the company name for a given record (e.g., stock ticker, internally designated acronym, etc.).</td>
</tr>
<tr>
<td>NATIONAL_IDENTIFIER</td>
<td>esriFieldTypeString</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>true</td>
<td></td>
<td>A governmentally assigned operator identifier at the national/federal level (for example, NPMS for US regulatory).</td>
</tr>
<tr>
<td>STATE_OR_PROV_IDENTIFIER</td>
<td>esriFieldTypeString</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>true</td>
<td></td>
<td>A governmentally assigned operator identifier at the state/provincial (regional/territory) level.</td>
</tr>
<tr>
<td>IS_OPERATOR</td>
<td>YES_NO_NA_CL</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>true</td>
<td>Unknown</td>
<td>Logical flag indicating that a company may have (is allowed to have) a foreign key value in the PIPELINE table (OPERATOR_COMPANY_ID) which means that the company operates the pipeline. This model assumes that a pipeline is operated by a single company.</td>
</tr>
<tr>
<td>IS_OWNER</td>
<td>YES_NO_NA_CL</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>true</td>
<td>Unknown</td>
<td>Logical flag indicating that a company may have (is allowed to have) a record in the OWNERSHIP table.</td>
</tr>
<tr>
<td>UNIQUE_ID</td>
<td>esriFieldTypeGUID</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>false</td>
<td></td>
<td>This is the primary key field for any table/feature class implemented in a PODS 7.0 data model.</td>
</tr>
<tr>
<td>CREATED_DATE</td>
<td>esriFieldTypeDate</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>false</td>
<td></td>
<td>The date and time the database row was created in a database table.</td>
</tr>
<tr>
<td>LAST_EDITED_DATE</td>
<td>esriFieldTypeDate</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>true</td>
<td></td>
<td>The date and time a database row as last modified or last edited. Modification means populating attributes with values that previously had no values, removing values that existed previously, or modifying the value of an attribute.</td>
</tr>
<tr>
<td>CREATED_USER</td>
<td>esriFieldTypeString</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>false</td>
<td></td>
<td>The name of the user or the software system who created the database record. Typically user IDs are derived from the operating system [OS] ID that the user logged in as. Alternatively, the CREATED_USER attribute could be populated by a relational database.</td>
</tr>
</tbody>
</table>
### Relational Databases

- Normalization of disparate data into a Relational Database
- A database design method that organizes information into multiple related tables to minimize data redundancy
- Information goes into the correct table, free of duplicates
- Tables are related to each other
- PODS database structure is normalized

<table>
<thead>
<tr>
<th>Operator</th>
<th>Pipeline</th>
<th>Meter</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Name</td>
<td>• Name</td>
<td>• Name</td>
</tr>
<tr>
<td>• UID</td>
<td>• UID</td>
<td>• UID</td>
</tr>
<tr>
<td>• Contact</td>
<td>• OP-ID</td>
<td>• PIPE-ID</td>
</tr>
</tbody>
</table>
Relational Database Key Fields

- **Primary Key**
  - Unique
  - Not null
  - Never changes

- **Foreign Key**
  - Not unique
  - Multiple
  - Points to some other Primary Key

<table>
<thead>
<tr>
<th>Operator</th>
<th>Pipeline</th>
<th>Meter</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Name</td>
<td>• Name</td>
<td>• Name</td>
</tr>
<tr>
<td>• UID</td>
<td>• UID</td>
<td>• UID</td>
</tr>
<tr>
<td>• Contact</td>
<td>• OP-ID</td>
<td>• PIPE-ID</td>
</tr>
</tbody>
</table>

**Operator**
- Name
- UID
- Contact

**Pipeline**
- Name
- UID
- OP-ID

**Meter**
- Name
- UID
- PIPE-ID
Example – Database Linkages

Stationed Centerline

Series related to Measure

Line related to Route
**What’s the difference between a Data Model and a Database?**

<table>
<thead>
<tr>
<th>Data Model</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Defines the connections between data – how it's defined, stored, processed</td>
<td>• Physical implementation of the data model</td>
</tr>
<tr>
<td>• Not the actual data</td>
<td>• A central data repository, i.e. “System of Record”</td>
</tr>
</tbody>
</table>
What is a Data Model?

A data model defines how data is connected, stored, and processed.

It provides the organizational and conceptual model of data relationships.

PODS has 3 types of data models:

1. **Conceptual** – Overview, low detail
2. **Logical** – detailed model including rules, standards, etc.
3. **Physical** – platform-specific, detailed, finished data model with defined data elements
PODS Data Models Over Time – PODS 4.02
PODS 7 Data Model

PODS 7.0 Conceptual Model
PODS 7 – Location Model

Location Model (Linear Referencing System - LRS) [Optional]

- Centerline
- CenterlineSequence
- LRSNetwork
- ContinuousMeasureNetwork
- EngineeringStationNetwork
- CalibrationPoint

Ordering and Measurement:
- Centerline to CenterlineSequence
- LRSNetwork to ContinuousMeasureNetwork
- EngineeringStationNetwork

Assign Measure to CalibrationPoint
What’s Different About PODS 7?

1. Completely re-designed, revised and streamlined data model
   Location information stored in business tables supporting linear referencing (event) and feature-based representations of pipeline features.

2. New approach for Modules makes them highly customizable to your business.
   Extends the model for business-specific processes, regulations, reporting, etc.

3. New Data Exchange Specification
   Improved readability, compatibility of data, and customization.

4. Improved, extended documentation
   Facilitates better end-user implementations
What key drivers shape the PODS Data Models?

- Regulatory Compliance
- Traceability of pipeline equipment, materials
- Quality assurance
- Interoperability with other enterprise systems
- Project management, monitoring
- Industry standards and common language
- Data management strategy
- Consistency through pipeline phases from design to operation
<table>
<thead>
<tr>
<th>Why PODS 7?</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Greater Agility to Build &amp; Extend</strong></td>
<td>Simpler base model. Focus on location of pipeline centerline. Modules allow for extensibility.</td>
</tr>
<tr>
<td><strong>Meet changing Business Requirements</strong></td>
<td>System of record for pipeline centerline location. Support the industry standard GIS software.</td>
</tr>
<tr>
<td><strong>Interoperability</strong></td>
<td>One model, multiple physical implementations. Schema is immutable. Flexible set of abstract classes based on geometry and data type</td>
</tr>
<tr>
<td><strong>Addressing Barriers to Entry</strong></td>
<td>Well documented. All three model types. 6 different physical models. Training. Sample Data.</td>
</tr>
<tr>
<td><strong>Data Exchange</strong></td>
<td>Facilitate transfer of data in standard format between suppliers and consumers be that systems or organizations.</td>
</tr>
<tr>
<td><strong>Beyond North America</strong></td>
<td>Non-regional specific terminology. Units of Measure can be applied to any field and subsets of features on different routes.</td>
</tr>
</tbody>
</table>
Why PODS 7?

Responsive to changing business requirements
Ensure that the location of the centerline. Everything else is secondary and derivative. Add the most important modules =ILI etc.

Optimize Performance
Geometry in the business tables. Standard schema for this. All database updates for attributes and geometry can happen with open SQL.

Open Source
PODS Lite model is published and release for free. Support POSTGIS and SQLite physical database patterns.

Data Validation/QA-QC
Business rules construct. Data requirements for ESRI APR and UPDM.

Metadata
Metadata tables are integral part of model. All have the word META_DATA in their title making them easier to locate in the schema.

Code Lookups
Domains or tables. Tables provide mechanism for retirement. Business rules provide ability for conditional domains.

PODS 7.0 is designed to be the system of record for pipeline centerlines and pressurized containment assets for the safe transport of product.
1. Information is only as good as the data it's built on.

2. PODS Data Models, including PODS 7, define the connections between data – how it's defined, stored, processed. Databases are the physical manifestation of the data model.

3. Relational Databases are the essential building blocks of the PODS Data Model providing connectivity throughout the PODS Data Model.

4. PODS 7 is quite different from previous versions of the PODS data models. The redesigned model is based on simplicity, customization, data sharing, and implementation support.
End of Unit 2

Any questions?
Resource for This Unit

PODS Association web site

https://www.pods.org