LINKING DATA TO GET THE FULL PICTURE

The 2010 Pacific Gas & Electric (PG&E) San Bruno tragic accident, in which a 30 in. intrastate natural gas pipeline ruptured in a residential neighbourhood of San Bruno, California, has resulted in increased scrutiny on pipeline operators and the systems they use to manage information about their pipelines. A 5th March, 2012 report, issued by the California Public Utilities Commission, illustrates the importance of data integration.

Janet Sinclair, Executive Director, the PODS Association, USA, explains why data integration is both a legal responsibility and a benefit to smooth pipeline operations.
PG&E had no definitive view, or a single master source of information, relating to each of their pipeline-related assets. This information is distributed across the organization in localized silos of information, that are often unconnected, poorly catalogued with little referential integrity between the various collections.

The Pipeline Open Data Standard (PODS) data model is an industry standard, used by pipeline operators to provide a “single master source of information,” and to eliminate “localized silos of information that are often unconnected.” As the US and other parts of the world increase their focus on pipeline integrity management (PIM) for natural gas pipelines is covered in ASME B31.8S ‘Managing System Integrity of Gas Pipelines’, last updated in 2004.

The ASME B31.8 piping codes and standards govern the design, fabrication, installation, inspection, testing, and integrity management of pipeline facilities used for the transportation of natural gas. Note that these piping codes and standards are internationally recognised industry best practices, but differ from government regulations. In some parts of the world (i.e., the US) strict adherence to these piping codes and standards is required. In others locations, pipeline operators may be left to determine the extent to which they choose to follow these codes and standards.

A single master source of information

PIM depends on good data, and easy access to that data, to facilitate conducting the prescribed analysis. The data required for PIM is listed in Table 1 of ASME B31.8S, and is categorised as follows:

- Attribute data.
- Construction data.
- Operational data.
- Inspection data.

Attribute data includes pipe diameter, wall thickness, seam type, material of construction, and manufacturer. Construction data includes date installed, bending method, joining method, depth of cover, pressure tests, soil type, backfill procedures, inspection reports, casings, and coatings.

The best source for this attribute and construction data is the ‘as-built’ data package. This information is typically turned over by the EPC contractor at the completion of pipeline construction, and usually takes the form of dozens of boxes of paper files, and stacks of USB flash drives loaded with electronic files.

The job of sifting through and deciphering the contents of this turnover package, and uploading the pertinent data into the corporate data management system, may take several months to complete, if it is performed at all. Missing, incorrect, and conflicting data is often found during this process, long after the pipeline is buried. This makes it particularly difficult to investigate and reconcile any problems found with the data.

This discussion of the ‘as-built’ turnover package assumes that the pipeline is recently constructed, but what about pipelines built 40 or 50 years ago? The level of effort involved to perform this same process for pipelines constructed many years ago may be significantly higher. This will be wholly dependent on the company’s record keeping practices at the time the pipeline was built, and the document retention policies and procedures since.

It is also important to note that the term ‘as-built’ also applies to subsequent modifications to the pipeline, such as re-route, replacement, repair, or addition. Records associated with modifications also need to flow into the corporate data management system. The process of deciding which records stay in tact, and which records need to be updated due to the modification, needs to be carefully performed and can be tedious.

Operational data is data gathered after the pipeline has been placed into service and includes gas quality and flowrate, normal, maximum and minimum operating pressures, leak detection results, coating assessment test results, performance information regarding the cathodic protection (CP) systems, ROW encroachments, vandalism, and repairs and maintenance.

Inspection data includes physical inspection reports, inline inspection (ILI) reports, hydrostatic pressure test results, geometry tool inspections, CP inspections, close interval survey (CIS) inspections, and audits. A single ILI tool run can generate millions of new records. These large inspection data sets must also be integrated into the corporate data management system.

Figure 1. An ILI tool (photo: ROSEN Inspection).
It is clear from the above description that the data required to support PIM comes from multiple sources. It is also clear that the amount of data is voluminous, and it grows both with cumulative length of pipelines supported, and with time.

**Role of a data model**
The PODS Association is a not-for-profit industry standards organisation whose mission is to develop and support open data storage and interchange standards to meet the specific data management needs of pipeline operators.

The PODS data model provides pipeline operators with an enterprise database architecture that is comprehensive, open, vendor-neutral, highly-scalable, and proven. It is an industry best practice integration platform for pipeline records and location information.

The data model is widely used in the US, and its use around the globe is growing as the focus on PIM is increasing. The data model was specifically designed for managing the data required to support PIM.

Originally initiated in 1998 by the Gas Research Institute (now GTI), PODS brought together pipeline operators and software providers to significantly extend the ISAT data model, and add important functionality focused on the data requirements for PIM. The first release, PODS 2.0, was offered to member companies in 1998, and it has since periodically released revisions as the data model has been expanded and improved. The current version, PODS 6.0, was just released in December 2012. The PODS Association continues to update the model to include enhancements submitted by member companies, and to capture advancements in inspection technologies and database best practices.

**A common reference for data integration**
The PODS data model is based on a linear reference system that incorporates location information with every record stored in the database. This ‘address’ or engineering station describes either a linear section of the pipeline or a point where that particular record applies.

For example, a pipeline may have a casing that begins at 1000 m from the beginning of the pipeline and is 50 m long. In the database, this casing is represented with a begin station of 1000 m and an end station of 1050 m. Similarly for a point feature, such as a valve, the begin station and end station are the same value (i.e., 500 m).

In a similar manner, the same linear referencing system is used to locate, and effectively overlay, inspection reports, events, activities, or features, which occur on or near the pipeline. The PODS data model contains a comprehensive set of tables and columns to store all of the data required for PIM, and other relevant data. PODS also allows for external documents such as reports, procurement records, hydrostatic test charts, or even photographs, to be linked to the physical assets they describe.

**Providing a ‘definitive view’**
With its comprehensive set of pipeline and inspection tables and its linear referencing system, PODS can provide pipeline operators with a “single master source of information, relating to each of their pipeline-related assets.” Further, viewing PODS data inside of a geographic information system (GIS) can provide pipeline operators with a ‘definitive view’ of their pipeline assets. This combination of PODS and GIS provides powerful analysis tools for performing the threat identification and assessments required for PIM.

“It is important to make the distinction between PODS and GIS. The PODS data model is typically implemented on a relational database management system (RDBMS) such as Oracle or SQL Server. Queries and interactions with the PODS database are written in structured query language (SQL), the most common database programming language. The PODS relational model does not depend on GIS, but a GIS is the most common method for displaying PODS data, either on maps or overlaid on satellite images. Software applications are available that allow loading and manipulation of data within a PODS database through a GIS frontend. The data model itself is GIS neutral, and the fact that it is open allows many software companies to develop applications that can interface with it.

The one exception to the above is PODS Esri Spatial, which is an implementation that was specifically designed for the Esri platform. This implementation allows pipeline operators to take full advantage of the tools inherent to the Esri software suite.

**Differences explained**
Some confusion still exists within the industry regarding the PODS, and the ArcGIS Pipeline Data Model (APDM). The main differences between these are explained below.

Esri released APDM in 2003. APDM provides a core set of conceptual objects, which can be used to describe most pipeline features. APDM is a template that an operator could use to build their own pipeline data model. It would need to be extended to support the data requirements of PIM. APDM can only be deployed on an Esri platform.

In contrast, PODS is already an extensive pipeline data model for storing a comprehensive set of critical features and attributes necessary to support PIM. PODS is GIS-platform independent, meaning it can work with Esri, Intergraph, or any other GIS software. It is available as either a relational database (Oracle and SQL Server are currently supported) or as an Esri geodatabase. The PODS data model has been managed and supported by the PODS Association, a not-for-profit industry standards organisation, since 1998.
Important changes in 2012

In its most recent release, PODS 6.0, the Association has made a fundamental shift in the design of the data model. Starting this year, the model has been modularised into functional groups that can be deployed independently.

The core module of PODS 6.0 maintains all of the information necessary to describe the pipeline centreline and its location on the earth, and is the only required module. Additional modules (there are 30) may be added as needed to support specific functional requirements. For example, pipeline companies located outside the US can drop the US Regulatory tables, and optionally add tables they develop to support their own region’s regulatory requirements.

The primary purpose of modularisation was to facilitate an easier implementation for new adopters. Pipeline operators are now more easily able to plan a PODS implementation in phases, perhaps start by populating just pipeline centreline and attributes, and over time add modules to support other functional requirements. A complete list the 31 modules and their descriptions is available on the PODS website.

Some of the advantages of modularising PODS include:

- It facilitates incremental additions and modifications to the PODS data model through minor releases of individual modules, rather than releasing an update to the entire model.
- It provides flexibility to develop targeted modules, such as region-specific regulatory modules, or modules for a specific industry segment (i.e., gathering).
- It provides pipeline operators the ability to implement only those modules applicable to their operations and desired level of complexity.
- It allows pipeline operators to adopt only desired updates, so long as all implemented modules are compatible with any published dependencies.

Note

The PODS data model is copyright protected, and the right to use it is granted only to member companies. Membership dues support further development of the model, promote the adoption of industry standards, and provide valuable benefits to members, including the annual PODS User Conference, and the online member forums.

References

2. Section 4, ASME B31.8S-2004, ’Managing System Integrity of Natural Gas Pipelines’.