REALISE THE VALUE OF PIPELINE DATA MANAGEMENT ACROSS THE ENTERPRISE BY EXPLOITING LEGACY DATABASES

By Stuart Clouston, Jay Smith & D Book
PII Pipeline Solutions, Cramlington, UK

Abstract

Within most pipeline organisations, maintenance and facility departments use a range of separate data sources and applications to manage the integrity, maintenance and safety of their pipelines. These databases represent a significant investment over many years and are an integral part of day-to-day operations.

It is evident that integration of data into a single, coherent data management system can provide significant benefits. However, the cost of implementing entirely new systems – with intensive data capture programs – is difficult to justify given the earlier investments. As a result dedicated risk management software using static and separately maintained data is often used as a quick, low cost alternative to meet regulatory compliance commitments.

Experience has shown that, with the right technology and an understanding of the specific needs of an organisation, a phased approach to integrated data management can be achieved at minimum initial cost by exploiting legacy data. This provides a low cost yet scalable solution that can grow with the changing needs of the business.

In addition to the benefits of legacy data integration, we will also look into the benefits of technologies for distributed data access to provide simple, process-focussed reporting tools.

Introduction

When you think about oil or gas transmission pipeline assets, the likelihood is that if you were to compare it to a telephone cable network serving an entire city, you would think managing the phone network data would be considerably more challenging and the pipeline system relatively simple. However, at the risk of sounding obvious, the fact is they are distinctly different. In the case of a phone network, cable routing, connections and line failures are a major concern. Where in the case of pipelines, the issues relate to understanding and managing integrity threats such as corrosion and 3rd party damage to ensure safe operation throughout the life of the pipeline. A pipeline, although simple in nature, has a complex range of continually changing, interacting events and operating conditions that must be carefully understood and managed.

The Challenge Of Managing Pipeline Data

So, what are the challenges of managing pipeline data? Ask almost any pipeline operator and they will tell you a similar story, integrity and facility critical data exists in various forms throughout their organisations and at best in "islands of automation" where small databases, records or software are being used to carry out specific tasks. In addition, other departments often duplicate work and little or no data is shared.

Typically the following key barriers exist to efficient data flow and management:

- Multi-format data (paper records, digital (spreadsheets), databases)
- Data distributed and managed in remote and often independent locations
- No means to align and compare related data
- No clear procedures for maintaining and controlling data
- No methods for communicating when data is updated

- Poor tools for sending data to and from the field
- Information held as knowledge with individuals
- No means for sharing data across the organisation

Given these common issues, why has the pipeline industry been slow to react to the ever-increasing need for better management of data? Of course a wide range of reasons exist but there are some common trends that have influenced progress. The major pipeline operators were some of the early adopters of Geographic Information Systems (GIS) and AM/FM CAD packages. These systems were used primarily to capture the pipeline facilities and geographically reference them using aerial photography (or vector landbase) to allow representation of centreline position as it relates to the real world. Typical applications were for landowner management for onshore systems and more generally for generation of alignment sheets, as-built drawings and response planning. The investment in this type of system at the time was high, and though implemented well, was unable to demonstrate sufficient payback to warrant further investment. As a consequence these type GIS systems maintained by the pipeline owners have been rarely used in day-to-day operations. The reason such systems were not truly successful is that they did not really consider the way a pipeline operator or engineer needs to use his data and were little more than just another source of data for interpretation.

Why Integrate Data?
The Need For Managed Data

To a pipeline maintenance engineer, gathering data is just the first step in an entire process of maintaining the condition and integrity of the asset. The next step is to take various sources of data, review and combine them, interpret the data and taking action as appropriate. As an example, gathering corrosion data is critical and easy to identify from an in-line inspection report where repairs are immediately necessary to ensure integrity. However, where the immediate integrity is not a threat, but corrosion is identified and must be understood in terms of causes, future inspections, future assessments and integrity, other related data is necessary to draw conclusions. This data may come in formats such as: performance of corrosion protection, soil data, product content etc. at the given location. Data availability, accuracy and quality are just the first piece of a much larger puzzle. The challenge is how to build a system that takes data and turns it into information in the context of managing a pipeline asset. To do that, an inherent understanding of how data needs to be used, compared and its relationships to other data needs to flow through every aspect of a data management system's design.

For those operators with aging pipeline networks many are faced with the issue that they hold much of their data in many different formats, have missing data and prior to the age of digital based storage have limited understanding of the quality and consistency of their facility, maintenance and inspection data.

Recognising the value of managing and controlling pipeline data in the overall safe operation, regulators in the US and Europe are placing significant focus on initiatives and regulations that will guide pipeline operators to demonstrate effective and auditable management of their pipeline assets. While not yet prescribed, it is inevitable that integrated data management of pipeline systems is the future for integrity management and will allow risk-based maintenance approaches to move to the next level. As an example, if we look at the US Liquid and Gas Rules for (Integrity Management) and (associated requirements for inspection, assessment and remediation within specified time frames), it is hard to imagine that integrated data management systems are not a pre-requisite for success.

Meeting The Needs

There are many customers of data in any company, each with different needs. This slide shows a simple overview of how integrated data management systems can meet "customer goals"

from a departmental level in the Engineering and Maintenance divisions of an organisation to the high-level corporate goals.

This next slide represents a generalised view in how data flows throughout the integrity management cycle of a pipeline. A constant interchange of related data passes through the various stages of data collection, verification, maintenance and monitoring. Clearly many users and software/database applications accessing a common database of well-maintained and controlled data will significantly enhance process efficiency in almost all parts of the cycle.

What Should You Consider When Defining Systems For Pipeline Data Management

Data Models – a foundation for success

At the heart of any successful database is the data model. The data model defines the types of data that can be stored (e.g. pipe segment, dent, tap, tee, corrosion data, valve, etc.) along with the specific attributes that must also be managed. (e.g. material, wall thickness, length, width, depth, serial no., manufacturer etc.) Additionally, a location reference for all data is required for comparison and identification. Design of the data model requires understanding of the processes and applications the data will be used for. Let's look at a simple example to accurately locate a corrosion feature. While a GPS (Global Positioning System) co-ordinate is great to identify where to excavate, a linear distance to a girth weld and clock-position are necessary to be confident that the correct feature is identified. The same applies for all other processes generating temporal data whether it be capturing pipeline burial depth surveying, land owner management or storing operational data.

Over the last few years, acknowledging that pipelines in general have similar requirements to each other no matter where they are, and that they are somewhat unique (meaning…linear over large distances with large volumes of continuously changing data), the industry has steered towards the establishment of standard data models. Established and sponsored by operators, vendors and research bodies, a number of models have gained recognition. (PODS, ISAT and APDM) and are accepted as a good basis of any pipeline transmission system.

Establishing a common linear reference

The standard method of referencing the location of any item on the ground is by use of an X-Y based location using either the geographic coordinates (latitude-longitude) or some locally established Cartesian coordinate system (easting-northing or X-Y). Corridor applications (such as transmission pipelines, fibre optic lines, railroads, highways, etc.) are unique in that a linear-based, one-dimensional positioning system is commonly used to reference the location of an event (both on and off the centreline). This linear (one dimensional) system locates an event by specifying the distance from a known beginning point to the event in feet, meters, miles or kilometres (to name a few). This linear location system is referred to as *station location* or *engineering stationing*.

In this system, every location along the centreline is displayed in the following format: *aa+bb* where *aa* = distance from the starting reference point in hundreds of feet and *bb* = feet from the previous 100m location. As an example, a point with a station value of 143+28 is 14,328m from the starting reference point of the line.

In an ideal system where no modifications to the original centreline have been made, the starting reference point is usually the beginning of the line or line segment–usually a metering station or pump / compressor site. Modifications to the centreline route will generally cause new reference points to be introduced to the measuring system. The use of a "station equation" at the point of correction, relates the "behind" stationing to the "ahead" stationing. Although station equations complicate the linear location system, the basic concept still remains with each point being located by referencing its distance from some known starting point along the centreline.

Tools that dynamically segment Data

Segmentation is a method used with linearly referenced data to collect areas with common characteristics together for analysis. This is accomplished dynamically when the data is needed so that the individual linear features can be stored most efficiently. This slide represents a range of typical pipeline characteristics on any one section of pipe. Pipelines can be made up of many event ranges along the pipe. (…an example would be a section of pipe with common coating) A segment would be where a group of defined multiple event ranges do not change. We use these segments to apply specific analysis that would be influenced should one of the parameters change in value or location along the pipe. In this case, in the upper part of slide we have defined six segments for analysis. However data is continuously changing on pipelines and even facility data changes with time. The lower part of the figure demonstrates the impact of a pipe replacement where a section of pipe with three wall thickness changes occur, is replaced by a single section of common wall. In this case when segmented we have now only four segments for analysis. Dynamic segmentation is where the software automatically segments the pipe prior to analysis so that any changes to integrity critical data is considered prior to analysis. This way any analysis always verifies that the most recent data is used.

Scalability

Significant investments may have already been made to manage the integrity, maintenance and safety of pipelines. To protect these investments, the solution must incorporate a phased approach that leverages legacy data. The solution must allow individual applications to be implemented against the existing data sources to accomplish immediate assessment needs. The system, however, must be also able to achieve long-term objectives to go beyond departmental solutions to share data and applications across the entire enterprise.

Data Alignment

Effective decision making to a pipeline integrity engineer is directly related to how well multi-year inspection data, environmental, geographical, facility and operational data can be aligned to identify potential threats, validate historical trends and establish mitigation measures.

Accurate alignment of data along the common linear reference is the only true way to align data for assessment. Inherent inaccuracies from various measurement methods render matching absolute position (e.g. GPS co-ordinates) impractical and of little value when it comes to aligning multiple data sets. The challenge, therefore, for any data management system is to take multiple data types (e.g. point data, linear data, raster data and spatial data) and relate it to one another to a high degree of accuracy. Furthermore, techniques to calibrate data are essential to match facility location data from other more accurate data sources such as field surveys. This ability to continuously improve data location accuracy will have an ongoing benefit throughout the life of the asset.

An accurately aligned data set is the first and most critical step to improving processes and analysis throughout the integrity management cycle.

Managing and Maintaining Data

Mapping and Storing is not Managing

Advances and the availability of GIS software and technology has made it very simple to take almost any data, overlay it in a geospatial reference frame and relate them to other features (e.g. corrosion features) …and display them on a map. However, at this point a very clear distinction must be made. Visualising datasets relative to one another is only a part of a successful data management system. Many people have implemented a simple GIS to visualise

data on maps, however, without purpose built tools to maintain that data given the large volume of data that continuously changes along a pipeline, they soon become unmanageable and most likely redundant. Alternatively their use is so restricted, that the real value of integrating data will never be realised.

Therefore prior to moving down the path of investing in a pipeline GIS or data management system the following key questions should be asked:

1. Is there a good method for loading and aligning large volume data e.g. inspection data?
2. Is there linear reference system common to all data?
3. If I make a pipeline modification or re-route how will I update all related data?
4. Do I have methods to calibrate out inaccuracies in measurement methods for locating data along the pipeline?
5. If I run analysis or 3rd party applications from the system, how will they be affected by changes in pipeline facility, geographic, operational, inspection or maintenance data?

Let me demonstrate the significance of these questions with a couple of examples. Assume we are performing a risk assessment on a section of pipe using a 3rd party risk software application to prioritise maintenance activities, and that data is being provided via a common company database. In the event that risk critical data changes (e.g. coating survey or class location data), then we need a means to ensure that there is a foolproof method (or the software is intelligent enough) to ensure that the risk calculations are always correct. So given the discussion on dynamic segmentation earlier… to prevent significant errors, all data required in any risk equations must always be dynamically segmented prior to calculation to capture changes unknown by the engineer performing the risk analysis.

As a second example, consider we have made a re-route of a pipeline. Given that all data is linearly referenced along the pipeline as described earlier, if we make a change to that reference system (e.g. an insert of an additional 50m of pipeline), then all other data downstream of the re-route will be affected. The data management system needs to be able to automatically update all data affected by such a change seamlessly.

In order to do this the software needs to understand and manage specific data relationships. While this example may vary depending on the size of the pipeline system, a more likely event having similar impact would be if you were to make a modification to a pipeline that introduced and additional weld, all data sets that reference weld numbers for their location (e.g. corrosion features), would need updated data attributes in the database.

Only by establishing data maintenance capabilities such as these can you establish a full proof system for managing integrity data.

Exploiting Existing Data

Most, if not all company's, have already invested large sums of money and resource into database and software systems that are used in various parts of the organisation during normal pipeline operations. Day to day processes now rely on certain systems that have become essential tools for engineering, management and planning processes. This slide shows diagrammatically how common data flows and interacts with many activities and processes within a pipeline operator's organisation.

Traditionally one of the biggest challenges in getting on the path to an integrated approach to data management has been a combination of a reluctance to change and the inability to demonstrate a defendable value proposition for further investment. One of the main reasons for these barriers was the need to basically replace existing IT infrastructure, software databases and applications with new solutions. However with a combination of proven approaches to handling and maintaining pipeline data and improved software and database

technology, solutions that exploit existing company data offer an opportunity for a strong business case. By communicating and leveraging existing enterprise data sources such as GIS, local databases and business software (e.g. SAP), a low start-up cost ensures a more compelling value story.

This next slide illustrates conceptually, how, by connecting to external (live and legacy) data sources it is possible to feed a purpose built environment for managing pipeline assets without the need to migrate data and change establish practices within the organisation. …In other words get better data and results from existing data with minimum cost and disruption.

Besides the obvious cost implications there are a number of very real and tangible benefits of using this approach that arise from integration with a purpose built software environment for handling pipeline data:

- Establishment of a common linear referencing system (where one didn't exist before)
- Maintenance and management tools for that linear reference
- High accuracy data alignment
- Dynamic segmentation for analysis such as risk assessment
- Additional reporting tools of aligned data sets *(this may include automated alignment sheet generation)*
- Establishment of a rich platform of accurate and controlled data for analysis and integration with pipeline focused engineering software and applications
- Single source of data for controlled distribution throughout the organisation
- A means for distributing data throughout the organisation (e.g. the internet)

This approach brings various data sets together to allow pipeline operators a low cost, scalable solution, creating a single data environment for the entire organisation and establishing improved tools for managing data going forward.

Getting Data To Those Who Need It – Distributed Data and Remote Access

If all we do is to pull data into a single location then really the job is only half done. To realise the benefit of an integrated data management system we need to go a step further to ensure that:

- Data is easily accessible to those who need it across the enterprise
- Data is accessible in the field and at remote locations or other field offices
- Data can be easily collected in-field and updated in the main database

Only by delivering these as part of the solution can there be confidence that the system will be used and data kept current.

Without doubt, one of the obvious advantages of a centralised, IT based integrity data management system is the freedom to access correct and controlled data throughout the organisation. Many GIS and data management systems can be readily web enabled with much of today's software in fact being so called "web native", meaning they were designed and structured to ensure rapid access over the Internet/Intranet.

As discussed previously, an active system is likely to be linked to a multitude of live/legacy database applications and other business management and financial tools. Most software platforms have been quick to adopt and apply the advantages that the Internet or Intranet brings to data distribution and have developed environments and tools to enable almost all data to be accessed over the web. With the inclusion of comprehensive security protocols, data has never been so easily shared and available to those who need it throughout organisations.

This next slide illustrates a typical high-level system design. A central database and environment effectively managing multiple data sources and acting as a single point of access to the enterprise's data. Multiple sites such as field offices have not only read access to data but full control and responsibility for maintaining data if required. With a single data source in a common data model, pipeline engineering applications driving efficiency in day-to-day processes can quickly be established in the knowledge that the data feeding analysis is both current and accurate.

The last piece of the puzzle is a mechanism with which to take data to and from the field. Using detached client technology, replicas or "snapshots" of the database can be downloaded to laptop or handheld device for field access as shown in the current slide. This is invaluable for field verification work, giving instant access to data such as location/co-ordinates, inspection data records, facility data and alignment sheets. With comprehensive data maintenance tools, field records can be instantly entered and corrections to co-ordinate references (e.g. a GPS position of a corrosion feature or valve station), can be logged for audit and permanent entry into the central database system.

Conclusions

Establishing an integrated data environment for pipeline asset and integrity management does not need to be a costly replacement of existing IT and software infrastructure. Given good system design focused on meeting the individual needs of the organisation, existing enterprise data sources in addition to best practices for managing and maintaining their data, pipeline organisations can start to experience the benefits of integrated data management with minimum start-up investment.