The Digital Pipeline
"From the Dirt to the Desktop"

Robert C. Maggio, Ph.D.
Vice President, GIS Operations and Development
Petris Technology, Inc.
1900 St. James Place
Houston, TX 77057
www.petris.com
713-582-8892

Abstract

In 2002-3, geo-spatial tools were developed to perform tasks driven by pipeline centerline mapping and regulatory compliance, while in 2004 pipeline databases were the task of the day. In 2005 the trend will be the development of the digital pipeline through the integration of the GIS with other systems such as asset management, accounting, maintenance, document management, and work-order processing. The workflow process may involve field data acquisition of pipeline construction and maintenance information using handheld systems equipped with GPS and GIS data sets. This workflow will facilitate the automated maintenance of an enterprise GIS by allowing field personnel to "check-out" a portion of the GIS, proceed to the field, conduct maintenance patrols, update the GIS, and generate work orders automatically. The tools, the technology, and the vision will be reviewed, and examples of the components will be presented that will bring the digital pipeline to the computer desktop.

Introduction

The pipeline industry in the United States has passed through some notable stages in its development with regard to information technology. The late 19th century saw pipeline construction and maintenance as a simple task - build it and patch it when it leaks. Due to the paucity of pipelines and people in this country, this approach worked. However, as both increased in numbers, the concern for safety and production became of paramount concern. The concern for increased production to serve customers witnessed increases in volumes and operating pressures; and the increase in population and demand for petroleum products created an increased need for safe operations, especially around the populated areas where product was to be delivered. As landuse in this country changed from rural agrarian to interspersed urban, the mode of operations of pipeline systems also changed.

During the 1980s, we witnessed the use of paper copies of "As Built" drawings and route maps of pipeline systems; pipeline operations depended heavily on the personal knowledge of field personnel. The 1990s saw an increased acceptance of automated systems and digital data, especially on newly constructed pipelines. With the advent of this new era came the dilemma of managing both digital and paper information systems. Typically, pipeline operations maintained the old records in paper form and the new records in digital form - this was just the way it was. Also in the 1990s, regulatory compliance requirements became more stringent, and the demand for the safer operation of pipelines became of greater concern to the public. This increased demand for safety, heightened the need for more complete pipeline records and spurred a need for the conversion and integration of the paper records of old, with the digital records of the present.

As regulatory compliance became a driving force in the operations of pipeline systems, operators discovered the need to have a single source for all pipeline operations data - thus was born the pipeline database. These information systems were designed to contain all the information that might be extracted from the paper records of old and the newly generated digital data of the present. Standard pipeline data models were developed that provided structure for the pipeline databases. These models were tried and tested until they became all-inclusive, and now serve as the pipeline data models that guide the storage of pipeline data from "operations" to "regulator compliance".

As the late 1990s faded into the 2000s, pipeline operations began to adopt more and more automated procedures for monitoring pipeline systems. SCADA, gas control operations, field data recorders, automated One Call ticket processing, and distributed computing technology have driven automation into all facets of pipeline operations. However, a new challenge has arisen, the implementation of "The Digital Pipeline - From the Dirt to the Desktop."

Vision

As pipeline operations moves forward, system integration will become the order of the day. However, since pipeline systems currently exist, procedures currently exist, and personnel currently exist, the implementation must be a slow, thoughtful process. If new pipeline companies were to build new pipeline systems, complete integration could be the norm. However, since this situation does not exist, the staged approach must be executed. Nonetheless, the vision of the operation of a digital pipeline system from the field to the corporate office is attainable. Any vice president of operations would like to enter his or her office each morning and say:

"Computer, what happened on my pipeline yesterday?"

Approach

In order for pipeline companies to approach The Digital Pipeline, certain facts of life must be reviewed - the digital era is here, regulatory compliance will only get worse, systems integration is mandatory, budgets will shrink, and corporate knowledge is mobile. In the human body, the heart is at the center of man's health and well-being; in the digital pipeline, the data repository is at the heart of the system. For the modern day pipeline company, this repository is the pipeline database. Considerable time and effort has been dedicated to the development of the pipeline data models that exist today. One such model is the Pipeline Open Data Standard (PODS) and another is the ArcGIS Pipeline Data Model (APDM). These are by no means the only models being used by pipeline companies.

The Pipeline Data Model (PDM) is structured to address major areas of interest to the pipeline industry - Centerline, Crossings, Facilities, and Regulatory Compliance. These major portions of the PDM each contain numerous tables with numerous descriptors - attributes. For example, the Facilities section contains a table for Valves, which contains attributes for valve - Name, Nominal_Inlet_Diameter, Type_CL, Model_Number, Date_Installed, etc. This represents very structured data storage - every piece of data has a predefined home. Therefore, regardless of when and where it is captured, it has a storage destination. The data flow in The Digital Pipeline could go so far as to say that the Nominal Inlet Diameter's home is in the PDM *Section* - **Facilities**; PDM *Table* - **Valv**e; PDM *Attribute* - **Nominal Inlet Diameter**. Regardless of whether this data is captured from a historical construction document, from an alignment sheet, from a CADD drawing, or from a field data recorder currently being used by a "line locator" on a pipeline patrol, it has a pre-defined home - from capture to storage!

Throughout the lifetime of a pipeline system, data is generated during construction, operations, and maintenance. Most of the data generated has a pre-defined home in the PDM. Therefore any system (digital or otherwise) that generates data should be designed with the intent that critical data are destined to be stored in (for example):

Section	**- Facilities;**
Table	**- Valve;**
Attribute	**- Nominal Inlet Diameter.**

Accomplishing this task requires that when new systems are designed, they are evaluated to ascertain which data elements will be captured and where in the PDM will they be stored.

Integration "from The Dirt to the Desktop" requires not only the adoption of a PDM and the realization that all data that is generated must be stored in its pre-defined home, but it also requires that independent systems be implemented with this approach in mind.

Systems

Most digital, data acquisition systems being implemented in the pipeline industry today record data in a file structure that is either proprietary or to one that is completely open, such as .xls, .dbf, or .txt. Typically the justification for using a proprietary format is to protect market share in the industry and to compel users to depend on selected vendors for integration. This approach has typically prevented the smooth integration of digital systems and perpetuated the fragmentation of information while presenting roadblocks to the basic desire to store all data in a central repository - **the pipeline database**. Another factor that has prevented smooth integration of independent system is the fear of change and the loss of control over existing systems.

If there is a lesson to be learned from the present conditions in the pipeline industry with regard to systems integration, it is to be aware of the data formats that the various information systems are using that are currently in place or that are about to be implemented. The critical question to be asked is: ***How will the data generated from this "tool" be passed to, and stored in, the pipeline database?***

Since the pipeline systems that we strive to manage cover large geographic extents, the data we strive to capture and store also has a geographic component. The critical question during the construction, operations, and maintenance of a pipeline system always contains a "where" component. All actions occur at some point on the pipeline system. Therefore, most of the data recorded during construction, operation, or maintenance of the pipeline system must contain a location parameter. In the pipeline database, this location parameter is either latitude and longitude or engineering station. It only makes sense that all "tools" that are used to collect construction, operations, and maintenance data on the pipeline system contain a location parameter or a link to one. The link need not be latitude and longitude coordinates if a unique identifier is collected, and if this unique identifier is currently stored in the pipeline database with a latitude and longitude coordinate.

Tools

Tools as used here denote software applications that generate digital data regarding the construction, operations, and maintenance of the pipeline system. These tools may be used for field data collection to integrity management modeling. However, the common thread among all tools is that they generate valuable data that must be captured and stored in the pipeline database - not all data that is captured must be stored in the pipeline database.

Centerline Mapping and Pipeline Database Development

Most pipeline operators have developed a digital centerline of their system, which contains the basic attributes of pipeline name/ID, diameter, system name, operator name, etc. This type of centerline satisfies the requirements for the National Pipeline Mapping System (NPMS) submissions and other state submissions such as the Texas Railroad Commission (T4). Additionally, it is suitable for registration of pipeline systems to state One Call centers.

The development of a pipeline database requires a considerable commitment of time and money. However, it does not have to be completed in a single project. Most pipeline companies spend several years developing the pipeline database and a lifetime updating it.

Pipeline Drawings

Most pipeline operators use "As Built" construction drawings and P&ID drawings as a source of data for the digital centerline. The "As Builts" were developed during construction, and each is the product of the engineering firm that designed and/or constructed the pipeline. They typically exist in paper form because the digital files that created them have long since been lost. For pipeline companies that operate 3-5,000 miles of pipeline, there may be as many as 6,000 sheets of "As Builts", P&ID drawings, or alignment sheets. Regardless, these drawings are worth their weight in gold, and are coveted by field operations personnel as though they are gold.

Since most pipeline drawings are not geo-referenced to a world coordinate system and do not exist digitally, it is difficult to piece together this collection of drawings into a seamless pipeline centerline without extensive field work. The approach taken by many operators combines both field work and digital, aerial photography. Having the pipeline system flown with high quality, geo-referenced aerial imagery tends to be the preferred method to begin the process of building the intelligent pipeline centerline.

A companion procedure to the use of aerial photography is the use of GPS for the acquisition of key assets on the pipeline system as well as major "points of inflection" (PI). The combination of the GPS control points superimposed upon the aerial imagery provides the critical link between the drawings and the "real world." Once the imagery and the GPS points have been merged in the GIS, the task of extracting the data from the drawings must begin.

The first task is to create an "acceptable," geo-referenced centerline. Pipeline centerline accuracy is quite variable. Many state and federal regulations set requirements for acceptable positional accuracy. However, the requirements for most pipeline companies are more stringent. More often than not, pipeline companies adopt a "crude" accuracy for centerline mapping initially and improve it as time and money permit. Mapping the centerline from drawings as an overlay to aerial imagery typically allows the positional accuracy of the imagery to dictate the positional accuracy of the centerline. Modestly priced aerial imagery has an accuracy of 10-15 meters. Some people refer to this as "quad map accuracy" due to the fact that the U.S. Geological Survey digital ortho quarter quadrangles have a map accuracy of 10 meters.

Once the centerline of the pipeline system has been established, other pipeline assets must be placed on the system. Many features that exist on the drawings are not visible on the pipeline in the field, e.g. points of inflection, belowground valves, pipe welds, etc. These assets are referenced on the drawings by engineering stations that were computed or measured during construction and must be assigned positions on the pipeline. In the GIS environment, features on the pipeline are assigned positions by heads-up digitizing the features from aerial photography, capturing GPS coordinates in the field, or by using stationing. If features are not visible in the field, the use of aerial imagery and GPS must be ruled out.

The GIS provides a wide range of geo-processing tools that will facilitate the placement of entities on the pipeline using the engineering stations extracted from the pipeline drawings. Identifying the point of beginning of the pipeline, whether it be at engineering station 0 or some other station, and building a data table of engineering stations and associated feature names begins the process. Attention must also be paid to the station series within which a station exists. On some pipelines, distances along the pipeline must re-start when pipeline re-routes are encountered. This re-setting of stationing must be taken into consideration when attempting to place features on the pipeline using this method. Regardless, once the stationing and feature names have been compiled, a geo-processing procedure is used to trace along the pipeline centerline until the distances represented by the engineering stations have been accumulated. At this point, a marker is placed in the GIS database and the feature name and stationing is entered. This process is followed to place all entities on the pipeline that were extracted from the drawings.

The database that accompanies the GIS software will provide a very simple storage medium during data compilation. However, a pipeline data model must be selected that will house the extensive asset attributes that must be stored. Most pipeline companies use the GIS database to compile the asset information and migrate it to a pipeline database at a later date.

Field Data Capture Applications

The pipeline industry has employed field data capture equipment to collect maintenance data, read flow meters, and capture GPS data for re-routes, etc., for a number of years. Each of these systems collects data about changes or events on the pipeline system. However, most of these systems were purchased and implemented prior to the adoption of pipeline databases. Additionally, the design and implementation of a pipeline database has routinely been the task of Information Technology or GIS departments and not the Operations department. Therefore, there was no need to integrate them. However, with the advent of stricter regulatory compliance issues, the need has arisen.

Regardless of the hardware and software used to collect field data, the data to be captured and its storage location must be identified in advance of deploying these field systems. Often times it is imperative to step back and take a holistic look at the dataflow within a pipeline company - from **"the dirt to the desktop."** An analysis of the maintenance and regulatory compliance data that must be collected will reveal data elements that should be stored in the pipeline database. Theoretically, the data elements collected by these systems should be assigned the Section, Table, and Attribute from the Pipeline Data Model where they should be stored. This process will eliminate the necessity for specialized software applications to extract data from proprietary systems. A standard transfer data format should be adopted.

Operationally, the procedures to streamline this task are:

1. Field personnel should use the software applications to capture the data,

2. Return to the office, dock the equipment with a desktop (or use wireless communications from the field or office if available) system,
3. Upload the data from the handheld, stage it in a temporary database,
4. Perform QA/QC checks, and
5. Push the data directly to the pipeline database.

This process would not be a "push-button" execution, but would prevent data from disappearing into isolated databases.

Pipeline Patrol

Routine pipeline patrols require that valves, pipe, and other assets be inspected every three (3) months for some pipeline systems. Selected portions of this inspection information may be stored in the pipeline database. The software used to capture the data could be configured to collect data regarding the inspection and condition of valves. Field data collection forms and equipment are gaining in acceptance for this task. Some of these systems have integrated GPS and camera equipment to augment the data captured by field personnel. The GPS not only captures latitude and longitude coordinates, it date-stamps the field visit, thereby verifying that the facility was inspected at the date and time indicated by the GPS.

Use of the field data capture forms provides an automated process whereby the data is captured and stored. The automation of this process facilitates error trapping at the source of the data by pre-loading acceptable data values (from the pipeline database) into the field data capture equipment. This pre-loading tends to reduce errors, increase efficiency, and streamline the dataflow from the field to the pipeline database for those data parameters that are stored in the pipeline database. Elements of this data should be identified by Section, Table, and Attribute and stored in the pipeline database. A pipeline database loader should be employed to populate data fields in the pipeline database.

Cathodic Protection Survey

Each of the cathodic protection test leads should be captured and stored in the pipeline database. This data may be used to identify which sections of the pipeline system are protected and which are not. The results of routine surveys of the Test Leads should be recorded in the pipeline database and used to update status of protected pipe. Output formats from these systems should include a format that will facilitate the uploading of the data to a staging database and then an automatic loading of the data into the pipeline database - **a pipeline database loader**.

Meter Reading

In many pipeline companies that are not automated, meters are read manually and the data is captured and stored in paper form (or on charts). The location and unique identifier of each meter has a Section, Table, and Attribute in the pipeline database. During field data collection, the meter ID is captured along with the meter reading. Post processing will compile the data in the "Measurements Database" for billing purposes. However, other pertinent information regarding the meter may be stored in the pipeline database. The appropriate Section, Table, and Attribute should be identified and used by the field data collection apparatus to facilitate this transfer and storage.

Pipeline Re-routes and Modifications

Pipeline systems are not static; they are always in a state of flux. Individual pipelines are added to accommodate connections to new wells and new pipelines. As population density continues to increase, re-routes are always a consideration. Since the pipeline database is a "live system," it must be fed.

As re-routes and pipeline renovations occur, this data must be captured and stored in the pipeline database. Prior to or during such events, the appropriate sections of the pipeline centerline may be downloaded into the appropriate field data capture hardware. Integrating the field equipment with a GPS will allow field personnel to digitize the location of the re-route in the field, and to capture the assets placed in service. The format of the data to be captured during this process should be dictated by the location (Section, Table, and Attribute) in the pipeline database where the data will reside. Once the data is captured, **a data loader** could be used to store the new assets in the pipeline database.

Regulatory Compliance

Housing Density

The U.S. DOT requires that some pipeline operators conduct Housing Density (Class Location) studies along the right-of-way of their pipelines. The rules for class location are defined by DOT. Pipeline operators perform class location studies in a number of ways; the most common methods are through field studies or aerial observation. Regardless of the method used, there should be a quantification of the results that may be stored in the Regulatory Compliance Section of the pipeline database. The aerial observation method tends to provide a more quantifiable result.

Using digital, aerial imagery, photointerpreters superimpose the pipeline centerline (extracted from the pipeline database) onto the imagery, buffer the centerline by the appropriate 660' and 300' buffers, examine the imagery, and heads-up digitize each structure that falls within the buffers. Each structure is assigned a unique identifier and a building structure code. An automated process is applied that creates pipeline segments according to housing density. Each segment is assigned a beginning and ending latitude and longitude coordinate, a DOT Class (1, 2, 3, or 4), and a segment length. The format of the results of this process should be such that a pipeline data loader may be used to upload the segments into the pipeline database. If attention is paid to the formats produced by the Class Location tool, this process may be automated.

High Consequence Areas

U.S. DOT requires pipeline operators to identify high consequence areas (HCA) with the intent of recognizing areas where people might reside that have limited mobility. These areas include hospitals, schools, day care facilities, correctional institutions, and outdoor areas such as playgrounds, parks, theaters, and public gathering areas.

Mapping the HCAs is most easily done by extracting the pipeline centerline from the pipeline database, superimposing it onto digital, aerial imagery, and mapping the HCA structures. As with the Class Location process described earlier, the HCA analysis uses a set of rules to process the structures with the intent of producing pipeline segments, each of which is assigned a beginning and ending latitude and longitude coordinate, an HCA designator (Y, N), and a segment length. The format of the results of this process should be such that a pipeline data loader may be used to upload the segments into the pipeline database. If attention is paid to the formats produced by the Class Location tool, this process may be automated.

One Call Ticket Processing

A part of the Integrity Management Plan that depends on field operations is the quantification of the threat of third-party invasion. One Call tickets that are processed by field personnel may be used to generate density levels of ticket occurrences. With each One Call ticket mapped by latitude and longitude coordinate and superimposed onto the pipeline system, a quantification of the density of One Call tickets along the pipeline may be produced. This approach will generate pipeline segments that are attributed by the number of tickets per unit of pipeline, beginning and ending latitude and longitude, and segment length that occurred within a pre-defined buffer. The output of this analysis should identify pipeline segments that will be stored in the pipeline database under the Regulatory Compliance section. The format of the results of this process should be such that a pipeline data loader may be used to upload the segments into the pipeline database.

External Systems

Once the pipeline database has been populated (initially), it stores the inventory of the entire system. However, since the pipeline database is merely the inventory, external software applications and systems must be used to add value to the data through interpretation and analysis. Linking the external systems to the pipeline database is facilitated through the standard pipeline data model architecture. Software applications that require pipeline inventory data may access the database by mapping the data structure. As long as pipeline companies maintain the standard data structure of the pipeline data models developed by the Pipeline Open Data Standard organization or the ArcGIS Pipeline Data Model (APDM), applications will be developed that add value to the data stored therein.

Risk Modeling and Integrity Management Plan

Ranking pipeline segments according to risk factors identified by the U.S. DOT consumed considerable time and money in the pipeline industry in 2003 and 2004. The objective was to identify those segments of pipeline that were at the highest risk level for failure. The factors that were identified as part of a risk model tended to characterize the pipeline by both physical pipeline factors and by external factors. The internal physical factors are those data elements that are stored in the pipeline database such as diameter, maximum allowable operating pressure, etc. The external factors are the threat of third-party invasion (extracted from One Call ticket analysis), U.S. DOT housing density study results, high consequence area identification, etc. Pipeline companies use both types of factors to risk rank pipeline segments. Risk-ranking a pipeline system is very difficult if the pipeline database has not been developed. The quantification of the necessary factors requires considerable time.

Asset Management and Work Order Processing

The pipeline database is the heart of the pipeline information system. Eventually, vast quantities of data will be stored in the system. However, other departments of a pipeline company must have access to the data to accomplish the goals of their departments. Scheduling asset maintenance and work order processing are two systems that must interact with the database. All maintenance and repairs occur on the pipeline system; and since the GIS (pipeline database) does not possess the capabilities to perform these tasks, external systems must be integrated to the database to facilitate these tasks.

The Digital Pipeline must be perceived as a virtual system because performing all of the processing necessary within a single system is not practical. However, developing linkages among systems that depend on standard data formats and standard data models is prudent. Each entity stored in the pipeline database is assigned a unique identifier within the architecture of the system. However, this unique ID is typically not viable as a unique link to external systems. Depending on the GIS software and the pipeline data model, the internal unique ID may change as geo-processing is performed on the pipeline data. Therefore, creating a unique ID for each asset (pipe segments, valves, meters, compressors, etc.) stored in the pipeline database is mandatory. Management of this unique ID will be imperative to the successful linkage of external systems to the pipeline database. Typically, neither the GIS nor the asset management software has this capability built in; it must be developed externally. This linkage does require custom developed software, but there seems to be no option at this time.

The linkage of the GIS database to external asset management systems and work order processing systems is typically a two-way street. Frequently new assets are added to the GIS database that has not yet been added to the asset management system, and so the dataflow is from the GIS database to the asset management system. In this scenario, the unique ID is assigned by the GIS. However, should the dataflow move from the asset management system to the GIS, the asset management system must assign and maintain the unique ID. In some rare cases, both systems may share this task if a common database is used to track unique IDs. Regardless of the dataflow, the task must be accomplished successfully.

The need for the maintenance on pipeline system assets is typically initiated by field personnel during routine pipeline patrols. In The Digital Pipeline, the need to repair an asset is captured using a field data recorder and electronic forms. When the asset is inspected in the field, it is identified by its location and unique ID, both of which are stored in the pipeline database. The field capture of this information triggers an event that will generate a work order. However, the work order will not be generated by the GIS, but by an external work order processing system. Once again, as with the asset management system, one system or the other must claim ownership of the work order generation. However, the unique asset ID that is to be repaired will be the link between the two systems.

The Digital Pipeline Summary

Each pipeline company must decide if "The Digital Pipeline" is a "pipe dream" or if it is the future of pipeline operations. The intent here is not to present an approach to an all-inclusive system that will some day crater under its own weight. Rather, it is to provide pipeline operators with a vision of what can be. Additionally, The Digital Pipeline need not become reality tomorrow or the day after. As reviewed here, it is made up of components that may fit together. The complete puzzle need not be assembled today. The Digital Pipeline should be approached with caution. When components of the system are considered for acquisition, ascertain whether these new pieces will fit well with, and communicate with, the existing pieces. "Some assembly required!"