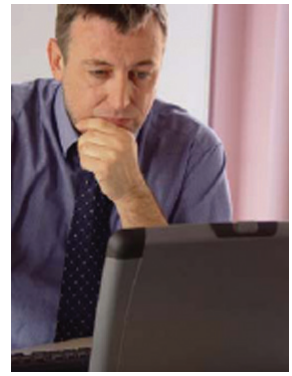
Creating a Dynamic, "Standards Neutral" Environment



By Dr. David Archer, VP,
and James Pritchett, CEO,
Petris Technology

Few efforts in the E&P industry are as challenging as those to develop common data standards.

Along with rapid advances in PC technology in the 1990s came literally hundreds of specialized E&P software applications, most of which were not designed to interoperate with other applications. The results are that valuable E&P data are scattered throughout data storehouses that few can access fully, and many applications simply cannot exchange information with other applications. These data access issues apply equally to so-called "structured" technical data, as well as to supporting "unstructured" information, such as memos, spreadsheets, presentations and e-mails.

Consistent data is difficult, in part, because different software vendors refer to the same concepts with different names. For example, one vendor identifies a well by a field called "Well ID" and another vendor calls it "Well Name;" or one may refer to the depth of the well as "Total Depth" and another might refer to it as "TD," etc. To share data between two such applications, you have to somehow make the appropriate associations from the field names in each application. This is a data semantics problem.

A further challenge is that data models – or the descriptions of how the data is organized and stored – are generally different in different systems. Further complicating the situation is a range of issues caused by different technology platforms – including those due to differing operating systems, application servers or database engines. Moreover, new data types are being invented every day, as people use raw data to create new data through interpretation and analytical tools.

Clearly, we need a means to manage all of today's data that can also handle the data that will soon be invented or discovered. This must be done in a way that neither restricts innovation nor limits the ability to use applications that best suit given needs.

Industry Standards

What role do across-the-board industry data standards play in addressing these issues?

Perhaps it is instructive to consider the role of industry standards in the electrical industry.

Electrical industry standards have enabled the use of a wide variety of appliances and equipment, made by thousands of different manufacturers, so that the devices can be interchanged (connected) conveniently and safely to the electricity grid throughout the US. The standards, developed by several key groups, are continually evolving to improve the safety and uniformity of electric and electronic appliances, equipment and components. Such standards have undoubtedly been an enabler of innovation for over a hundred years.

Why has this approach been so successful in the case of electricity and less so with E&P data standards? Why can't sharing data across a long list of different software applications be as easy as sharing electricity?

Standards Approaches

Think of the electric network as the framework that powers and allows for the connectivity of appliances and the sharing of electricity. This roughly compares with a data integration/connectivity network and its ability to connect software applications (the appliances) to share data (electricity).

In the electricity example, the standards specify not only physical connectivity (plugs), but also the type of current (AC/DC), the magnitude of the current and the frequency.

Shouldn't there be standards that tell the application manufacturers how to connect? That is, is there not a need for standards that govern the interfaces between applications?

The answer to these questions is: "Yes."

However, E&P industry has taken a different approach. Instead, the E&P industry has set out – in many different groups – to define a common data model or model subsets (business objects) along with sets of business rules intended to allow everything to talk to everything else. The main standards organizations within the oil and gas industry include Energistics, formerly the Petrotechnical Open Standards Consortium (POSC), the Public Petroleum Data Model Association (PPDM), the Pipeline Open Database Standards Association (PODS) and the Petroleum Industry Data Exchange (PIDX.)

While there has been progress in many areas and several data standards have been developed, data standards adoption has been sporadic and has lagged the industry's need for a systematic way to integrate all types of data, dependably and consistently.

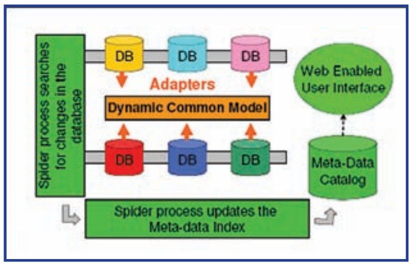


Figure 1

In a 2006 report, Cambridge Energy Research Associates (CERA) noted: "Lack of standardization of information technology systems has long been recognized as a cost carried by the industry. One estimate places the total value of such non-standardization at $40 billion, with the bulk of that tied to improved recovery of reserves. Given the opportunity to capture value, over the years the industry has launched a number of standardization initiatives. However, it is generally recognized that many of these initiatives failed to reach the levels of performance, adoption or benefit promised."

In contrast to the electrical standards situation, with E&P software, multitudes of applications (the appliances) were developed long before anyone came up with a way to interconnect them. Connectivity required both practical data standards and standardized application interfaces or integration platforms. To retrofit these applications according to existing or emerging data standards could require, at least in the short-term, non-economic changes to many applications.

Only Part of the Problem

There are a number of reasons why E&P data standards are only a part of the data integration problem. One, standardization can be a lengthy process. It often takes years for groups to agree on a standard. Then, as we have seen with E&P data standards, it takes even longer for a significant number of vendors to adopt the standard. Also, if they do adopt it, there is no guarantee that all vendors will implement the same standard the same way.

Two, a common industry model generally represents the lowest common denominator (the intersection of all common data types from each application) of the disparate application data models. This means you might be able to exchange only a certain subset of all the data – and data elements not included standard can be lost. This is akin to a "filter" through which all data must pass and through which only data types defined by the standard can pass.

Three, E&P companies must access large volumes of historical data. Implementing new data standards as a forward-reaching process does not do much to enable access to historical data. Companies need be able to use historical information on the same basis as current information in order to make decisions and to optimize operations.

Finally, the single-minded focus on data model standards – especially throughout the early 1990s – provides only part of the solution which requires, in addition to standardized data models, standard protocols and interfaces for communicating among applications.

This is not to say the industry cannot use standards or that it has not benefited form the various standardization efforts. To the extent that data standards have been widely implemented, they are providing value. In fact, standards that have focused on both information content and implementations have been successful. WITSML (Wellsite Information Transfer Markup Language) and PRODML (Production XML) are examples of very successful standardization efforts; yet they represent only a small but growing subset of the overall E&P data world.

Data standards alone, therefore, cannot address the full range of access and interoperability problems. What is critical is a way to access and manage all the data that exists now, and which anticipates all the new data that will be discovered in the future and which works even in the absence of data standards that cover the complete universe of E&P information.

Rather than trying to impose a data standard upon application vendors (as the electrical industry was able to do), the more effective approach is to develop an integration framework that can accommodate all data models, naming conventions and connectivity formats. In effect, each application suite can have its own data standard, with all of these recognized by the integration framework. Such technology indeed exists.

A Standards-Neutral Approach

To address information management problems arising from disparate data sources and non-standard data representation, Petris developed the patented Dynamic Common Model (DCM.) DCM is the only data integration model that assumes a constantly changing, or dynamic, data environment. The DCM is designed to enable the rapid, easy addition of new data types by application users or subject matter experts such as engineers, geologists and geophysicists. Use of the DCM eliminates the costs and time that would be needed for IT staff to write and maintain complex systems to connect required data types throughout an enterprise. The DCM is the core of the PetrisWINDS Enterprise (PWE) integration framework.

As stated previously, a central problem with most integration systems is that they rely on standard/canonical data integration models. In these models, when new type of data needs to be included, it must either conform to an existing definition, or be added to every data store and application with the potential to use the new data. Furthermore, when new applications are introduced that have several types of data, any data not pre-defined in the model are simply lost.

All data types – not only the ones on a pre-defined list – are represented in the DCM. In fact, a major strength of the DCM is that it stores the relationships between the types of data, not the data itself.

The DCM is designed for change; it permits the data model to grow dynamically, independent of the data itself. It allows data to be accessed and moved to other applications or data stores – without the need to add another set of data definitions to an industry standard/canonical model. Only the DCM can guarantee no data loss as data moves between applications or data stores. All of the important features are preserved.

When a new application or data store is added to the DCM, all of its data types are mapped to the DCM, not just the ones that are already represented in the standard/canonical model in use. The user defines the relationship between the application data types and the ones that already exist in the DCM. If particular data types do not exist in the DCM, they are added to it. Applications and data stores connected in the future then utilize this expanded list of data types and relationships in the DCM. Adding data stores or changing data types within a data store has no effect on the relationships between existing applications/data stores and the DCM. As in the case of industry standard data models, the DCM is not a complete integration solution on its own. An equally flexible and dynamic integration framework also is required.

Such a framework employs a Web-enabled metadata catalogue. Functioning with simplicity and visual clarity similar to popular search engines such as Google and Yahoo, the metadata catalog, like an index, contains summaries of all the available data types located in data stores connected to the framework. In E&P, these include data types such as seismic, well header, production, log curves, lease ownership, images and more. A user-configurable spider process searches for changes in a company's databases and updates the metadata catalogue on demand or automatically. A Web-based interface provides the additional advantage of providing access to data anytime, from any Internet connection.

Another key building block in this integration approach is the adapter, a special server side-program that transparently exposes the data types in applications or data sources connected to the framework and that are represented in the DCM. XSLTs (Extensible Stylesheet Language Transformations) provide bi-directional mappings between data types in connected data stores and the DCM.

PWE also includes an application called the Semantic Manager, which allows a subject matter expert to create and maintain mappings between data types from different applications themselves, using a visual, non-programming interface that recognizes the context and true meaning of each term. The subject matter expert uses the Semantic Manager to connect "like with like" across the two data sets by simply pointing and clicking.

Competitive Advantage

Oil and gas companies lose competitiveness when faced with overly long cycle times in locating, retrieving and transforming data or having to re-create missing data. While data standards development shows promise in some areas, especially around Web-based exchange standards, most companies choose not to get involved in the standardization process.

The DCM serves as the driver of the only technology available that combines full integration of all structured and unstructured data types with a framework that provides rapid text and geospatial, map-based search capabilities. As a result, data that otherwise can take days to locate, load and use can be displayed in a Web-based view of all available enterprise data presented in a form ready to be managed and transferred between applications – in minutes flat.

Success has always hinged upon the ability to adapt to change. Today, the ability to adapt smoothly to new information, new information sources and delivery mechanisms as well as to technological change is critical. This need to be agile and to adapt to change is the driver for the Petris data management and integration approach. This approach is vendor-neutral, extensible and flexible – in a word "dynamic."

www.petris.com