

From Vol. 23 of The Data Warehousing Institute's *What Works in Data Integration*
The Business Value of Managing and Integrating Information

The energy industry is one of the most knowledge- and capital-intensive industries in the world today. Over the past decade, energy companies have recognized that the use, management and preservation of information and data¹ are key to achieving business success. Since the industry's inception, information has been gathered and utilized to help identify, exploit and produce oil and gas assets throughout the world, and is used daily to make multi-million dollar decisions daily about every aspect of operations, from exploration to production to distribution.

Because of its information-intensive nature, there are many lessons that can be learned from the experiences of the energy industry – lessons which can be applied to organizations in other industries that are dealing with optimizing the use, management and preservation of their information. Petris has more than ten years of experience working with oil and gas companies; this paper reflects many of the 'lessons learned' in helping these companies better manage and integrate their information.

The Business Hierarchy of Information

Though each organization has some differences, in general there are three general areas of business activity that produce and use information. These are:

- **Administrative** – This includes information that resides in financial and HR systems, and information that supports regulatory compliance, management reporting, risk reporting and general business operations.
- **Production/Operations** – This includes information residing in systems which control operations, ERP and workflow. It feeds consolidated information (results/summaries) to the Administrative systems. In the upstream energy industry, this area focuses on oil & gas production – the activities that take place above ground to measure, produce, transport and refine petroleum products.
- **Technical** – This information resides in specific applications used to design, analyze and operate the various processes that generate revenue for the business. It includes engineering and technical applications, as well as specific built-for-purpose applications used to support the engineering and scientific activities of the organization. In the upstream energy industry, this applies to the applications and data used to support sub-surface activities, including exploration, drilling and well/field management.

Information is best managed by the people who work with it most frequently. More often than not, information exists within 'silos' or 'islands' that are properly managed themselves but have limited connectivity to similar information sources in other part of the organization. This can work fine to some extent as long as information is tended by its owners, but as organizations grow and evolve, valuable information may become isolated.

Integration can provide the ability to see further into the information, providing a more comprehensive view. The preservation of valuable information, coupled with the ability to see a broader view through improved access, are two of the benefits that can be gained from integration.

Structured and Unstructured Data: Business Value Creation

Each area of business activity produces structured data. This is generally managed within applications, including ERP and financial systems, engineering design systems and technical applications, and CRM and support applications. The management of this information is a full-time activity within any business organization, and its status and preservation are treated as critical to business operations.

¹ Within the context of this paper, the terms 'information' and 'data' are used interchangeably.

There is almost always a much larger volume of related unstructured information, which includes data from diverse sources such as spreadsheets, word-processing applications, presentations, email, web sites and a plethora of interactive web-based tools (such as chat and bulletin boards). It also includes the notes and documents used to generate structured information, such as notes and observations produced by an interpreter analyzing data from a well in the energy industry.

This unstructured information is a rich resource of useful knowledge. However, because it lacks structure and has minimal quality checks, it requires different processes and methods to handle and manage. The ‘knowledge management’ revolution of the past decade heightened the sense of urgency to preserve and reuse the non-structured, informal information that an organization produces, while stimulating the development of tools to better manage it.

Integrating unstructured with structured information is not a simple task; there must be business value to be gained in order for organizations to invest in the hard work of making the integration. Based upon Petris’ experience in the oil and gas industry, there are a number of key business advantages to be obtained through integration:

- 1) **Unstructured information provides insight into the activities and the thought processes leading to the production of structured information.** When this is well-documented, it allows others to follow the steps involved, enabling transfer of both knowledge and insight about how a process is conducted.
- 2) **The linkage of unstructured to structured information provides the ‘complete story’,** giving users the most comprehensive view of the information, what led to its creation and how it may have changed over time. When activities span years and multiple waves of workers, having the information live on
- 3) **Having a comprehensive view of all information helps optimize decision making processes, reducing risk.** Every manager’s nightmare is to discover, after the decision has been made, there was additional information that would have changed the conclusion. Having the best information available allows higher confidence in decisions, while understanding the range of possibilities allows mitigation of many of the attendant risks.

The most successful organizations are those that can maximize the value of their business assets through the effective combination of structured and unstructured information. However, even with these benefits, many companies do not give the level of attention to unstructured information that it deserves, largely because the value is much harder to recognize, and because establishing the linkages between structured and unstructured information is largely ad-hoc due to the discrete nature of the systems.

The Overall Benefits of Integration

In every organization, there are generally numerous disparate information sources that need to be brought together. Integrating these discrete information sources is not a simple task, requiring careful planning and clear definition of the business goals to be achieved. However, it is clear from the experiences of those organizations that have integrated their information that the benefits clearly make it worth the investment of time and effort to achieve the integration.

- 1) **Increased efficiency and productivity.** Better management of information assets yields better understanding of operations, supporting improvements in cycle time and increases in overall operational efficiency. In the energy business, it also results in improved recovery rates from producing assets.
- 2) **Increased usefulness of the information to end users.** Integrating information provides a ‘meta view’ of multiple information sources which, taken together, provide additional insight into the specific workflow or operation.
- 3) **Improved quality of information.** Since all information can be seen together, problems with the information become apparent and can be remedied.
- 4) **Supply chain integration.** Improved forecasting and supply- demand balance by providing production information to all stakeholders, particularly those in supply chain functions.
- 5) **Enhanced employee productivity.** Allowing information to be integrated and shared across discrete applications allows workers to continue using familiar tools, yet have access to all available information

regardless of its format. This is especially valuable during acquisitions and when bringing new people into an organization, as some technical applications may take years to learn.

Obstacles to Integration

A recent survey by Information Builders² provided an in-depth look at the state of integration in the upstream energy business. One of the points discussed was the obstacles to integration. With few exceptions, the respondents mentioned in the report agreed on a common set of challenges facing integration solutions, including:

- Lack of standards and naming conventions;
- “Silos” of information;
- Quality and gaps in data;
- Large quantities of data;
- Lack of experienced people to interpret (analyze) the data; and
- Lack of a common integration framework.

These obstacles are not unique to the energy business. A robust approach to integration addresses and overcomes the majority of these obstacles.

Standards and Integration

To gain all rewards and efficiencies companies expect once integrated, systems must have the flexibility to exchange data and knowledge. However, without commonly understood and widely available interfaces, there is no effective way to connect all of the equipment, applications and people who form the basis of digital energy systems. Standards help define how these interfaces function.

Standards also greatly improve the overall data quality. Why? Because once the data is represented in a common format, it is finally exposed in a way that allows a direct ‘apples to apples’ comparison of quality, which isn’t really possible when the data resides in discrete applications and formats.

The energy industry spearheads numerous efforts aimed at standardizing data, largely focused on the exchange of data between companies. Even though the potential payoff is huge, the industry has been slow to make significant investments in this area.

The standards themselves are not necessarily the end of the story. One important aspect of this discussion is semantics, the derivation of meaning from the nature and structure of a term. Both the word itself and how it is used contribute to its definition – for instance, a ‘log’ in oilfield usage is generally quite different from a ‘log’ in the lumber industry.

Having standards helps greatly in setting the structure for terminology. However, there is not always agreement on the exact terminology employed, both between discrete applications and even from one area of a company to another. Even if a company broadly adopts a standardized approach, there are still many issues associated with conversion of legacy data, as well as the difficulties of enabling data sharing with companies who do not use the same standard.

One lesson to be learned from the ongoing discussions about standards is that nothing is static, and *the approach/technology is selected for integration must enable changes to be made in the future*, including the implementation of new standards, revisions to existing ones, and changes in the way in which terminology is defined.

Dealing with Silos and Islands

Information is best managed by the groups which use it most frequently. In the case of most companies, this means that information is often housed in local repositories for support of local operations – the so-called ‘islands’ and

² “A Snapshot of Integration in Upstream Oil & Gas Operations - A Research Report By Advertas and Upstream CIO Newsletter” – August 2006

‘silos’ of information. Often this information is specific to certain applications, and the more critical/valuable information is kept under secure access. This process works well as long as access to information is only needed locally or by specific business areas.

The internet provides the means to connect these sources, but maintaining local ownership of the data helps ensure that someone tends to it. The integration of these silos and islands of data is largely a people-related issue; however *technology that allows data to remain in its native sources while being indexed for the broader view gives companies the ‘best of both worlds’.*

Data Quality and Integration

To achieve the benefits of integration, data must be of a known quality. Although it’s accepted that unstructured data may be of varying quality, there are tools and processes that help improve and even ‘rate’ the quality and reliability of this data.

When it comes to structured data, quality is also an issue. For instance, when a full collection of data is made visible for the first time, even companies that have maintained fairly rigorous quality standards note that the data is of uneven quality, with many ‘holes’, duplicate data and questionable data sets.

Good data management requires data of known quality. It is essential in the process of integration to recognize and ensure that the proper tools and processes are in place to determine and, when needed, improve data quality.

Security and Integration

In the energy industry, technical data and production information provides invaluable information about potential revenues for a company. Protection of this data from both unauthorized access and possible corruption is the responsibility of data managers within the enterprise – and a potential nightmare for the company. Individual applications and data stores typically have their own security protection through passwords and restrictions managed individual applications where the data resides.

But when data is integrated, this approach to security doesn’t necessarily provide assurances that data won’t be changed or inadvertently made available to someone who is not authorized to see it. Issues around security of the data may arise as the data becomes more broadly accessible through integrated data management techniques and solutions. The integrity of the original data sources must be maintained, and access must be preserved for those who need to use it. And rather than forge a new security schema, *it makes sense in many situations to leave the existing security in place, using an approach that relies on the native security within each application/data source.*

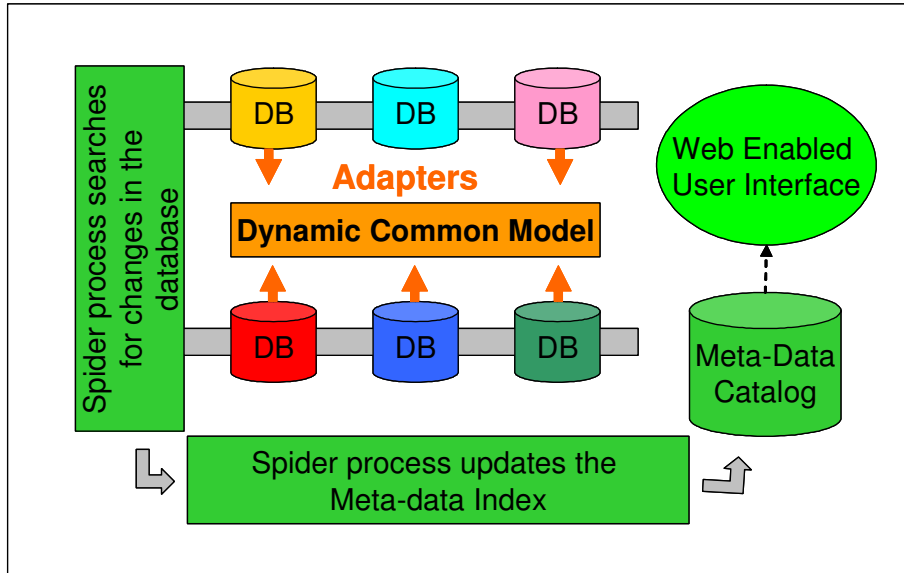
A Flexible Framework for Integration: PetrisWINDS Enterprise

An integration solution which addresses each of the obstacles noted is required to achieve the full business value of the integration.

PetrisWINDS Enterprise provides this solution. PetrisWINDS Enterprise is composed of multiple processing modules assembled through a **service-oriented architecture** (SOA) to provide the data management solution. Using SOA makes the creation and addition of new services and system expansion far simpler, easier, and faster than with older, more conventional IT architectural approaches.

Data in PetrisWINDS Enterprise always remains in its native format and/or data store, without the need for data warehousing. Adapters that do not require any modifications to vendor applications provide access. PetrisWINDS Enterprise’s lightweight **Metadata catalog** contains synchronized information about the identity and location of all connected structured and unstructured collections, and includes viewing permissions, data management events, and workflow conditions.

Integration is accomplished through storing the relationships between each type of data: data mappings. The metadata catalog also stores the data mappings that are at the heart of the patented PetrisWINDS Enterprise transformation technology. The optional Tightly Coupled Catalog provides SQL access to any connected application insuring synchronization of changes made in either the native application or the catalog. It can enable the bulk processing of data clean up operations.



The **Dynamic Common Model™**, or DCM, enables integration across the complex and broad-scoped data models of scientific and technical industries by expanding dynamically to encompass all data types, regardless of whether or not they ‘fit’ within a standard data model. As new data types are added to the system, they are mapped to existing ones or added as new instances.

PetrisWINDS Enterprise achieves its ability to connect to any data source

through the use of its adapters, server-side programs that facilitate the transfer of data to and from the native data source applications. An **Adapter** transparently exposes the data types of its specific data source or application to the DCM. XSLTs within the Data Exchange Model support the transformation of the data from the data source to the DCM, and from the DCM to the data source. By adding the PetrisWINDS Enterprise **Web Services module**, you can have access to any data source through its adapter.

Mapping data is only part of the solution – the larger challenge is to enable users to search for and retrieve needed data effectively. Instead of searching native data stores, a method dependent on network and database performance, PetrisWINDS Enterprise searches the definition-based metadata catalog optimized for search speed. The **map and text Search** functions run across all data in the catalog, and a periodic user-programmable ‘spider’ process ensures that the contents of the metadata catalog are synchronized with each connected data source and application.

The power of the PetrisWINDS Enterprise SOA approach allows the integration of ‘best of breed’ components from 3rd-party suppliers, including visualization tools, alternative access methods, geo-referenced indices, and portal integration. In the changing world of Data Management, PetrisWINDS Enterprise provides the most cost-effective and responsive method for managing data to support evolving business needs.

PetrisWINDS Enterprise and Standards: A “Semantic GPS”

Adoption of a standardized approach to data naming, structure and exchange requires that current and legacy data be reformatted and converted to be compliant with the new standard. This is true regardless of the standard selected.

The PetrisWINDS Enterprise integration platform provides the means by which a company can adopt and implement any standard, across any application they may use. Through the use of adapters specific to each application, data can be mapped to the new standard and moved as needed. The lightweight catalog within PetrisWINDS Enterprise allows users to search across all data collections, regardless of format, and then move selected data sets as required for their analysis. Data can be moved across and stored in the new standardized format, or left in its existing data stores to be accessed and moved only as needed.

PetrisWINDS Enterprise and Security

The Petris solution of a lightweight searchable metadata catalog allows authorized users to know what data exists and its parameters. The data itself remains in its native data source. A ‘shopping cart’ model permits a copy of the data to be moved from one application to wherever it’s needed. The original data remains untouched. Adapters specific to the source and target applications are used which ‘convert’ the data from one format to the other, without loss or change.

Once the analyses are complete, the results may be retained in whatever format the user desires, and indexed by PetrisWINDS Enterprise. A user searching afterwards would then see both the original data source as well as the new analysis, along with any other analyses that exist. This approach maintains data integrity while preserving access rights and licenses to those users who are authorized to actually work with the data. Companies recognize the benefits of integration without compromising data security – in essence, they get the best of both worlds.

“Rules of Thumb” for Integration

Petris has been working with companies in the energy industry for more than 10 years. Over that time frame, a number of good ‘rules of thumb’ have emerged that provide a good framework for the evaluation and procurement of a flexible and expandable solution for integration and data management. These include the following:

1. **Select only “vendor neutral” data management and integration platforms.** Some providers offer “solutions” for integrating and managing E&P data that in reality lock you in to using only that provider’s software. To compete in today’s market you need the flexibility to choose software that is most suited to your operation’s specialized situations, regardless of the vendor.
2. **Focus on applications that provide user-friendly features,** such as screens with clear icons that almost anybody can learn intuitively – think “Google.” An effective integration/data management solution will be largely a point-and-click proposition – not something that poses a multi-year learning curve.
3. **Make sure that the data management and integration solution selected handles ALL of your information types – and can easily expand to handle others.** These include “structured” technical and financial data, plus reports, interpretations, Excel charts, PowerPoint slides, maps, diagrams, memos and relevant e-mails.
4. **Select a provider with experience in your industry.** Domain-specific knowledge on the part of your provider will speed implementation and help you all speak the same language from the start. And for their part, they should bring the in-depth IT knowledge necessary to make the project work, in addition to their understanding of E&P.

Why Petris?

At Petris, we’re working hard to bridge the gap between industry’s daunting array of software applications (and formats) and the provision of anytime, anywhere access to data in any needed format. As we continue to expand our own portfolio of software integration and analytical tools, our goal is to provide the easiest, most direct access to the data you need, regardless of its original format or source, with the tools to better understand and apply it to your operations.

As challenging as that might seem, we know that high quality integrated data helps asset managers make better decisions. Our solutions apply industry-wide, from projects designed to optimize production in older fields, to the management of pipelines and digital oil fields of the future.

In today’s high-tech businesses, projects often function or fail based on the quality of their data, coupled with the quality of the analysis. Petris has the level of domain knowledge required to help companies tackle their most difficult data and analyses challenges, with the solutions needed to solve them.