Data Integration is Changing Roles for Petrochemical Professionals

By James W. Pritchett, CEO, Petris, Inc.
Presented at Petroleum Network Education Conference (PNEC) 2012
1. Introduction

Improving data management and integration is a common goal for exploration and production (E&P) companies and has become increasingly important with the growth in data and its derived interpreted artifacts. Better data integration is vital in helping to break down traditional operational/domain silos and to improve collaboration so petrotechnical professionals can better understand complex reservoirs and new, fast-paced unconventional plays, ultimately improving recovery and reserves.

This paper examines how data integration is changing our roles, including oil company data managers and information technology professionals, domain experts, and consultants. For those who have ventured there, implementing integrated solutions and workflows has delivered many lessons learned and is delivering real value to E&P organizations. Review of some recent projects that have been documented in previous papers and presentations as well, as an analysis of our own operations and services, provide a snapshot of the changing roles, some key challenges in the process, and recommendations for continued progress towards achieving the promise of the digital oilfield.

1.1 Industry Challenges: Why Data Integration is Important for E&P

In general, information technology (IT) professionals have been working on data integration since the development of databases (Ziegler and Dittrich). An integrated view of data can make it easier to access and reuse data, and the combination of integrated data sources can provide powerful new insights not available from the individual sources (Ziegler and Dittrich).

Petrotechnical and IT professionals attending PNEC know deeply and personally the E&P industry challenges driving the need for data integration, including: Huge volumes of complex, scientific data, real-time operations data as well as unstructured data (emails, spreadsheets, documents, etc.) that must be securely and reliably managed.

- Multiple, disparate systems with conflicting data and information cause confusion and waste time, as petrotechnical users work to find the right data for their current projects. Also, individual systems support the status quo of domain silos, which hinder team collaboration, limit new insights, and block potential for workflow transformations.
- There is a growing need to improve workflows that use data and information to make better and faster decisions to improve recovery and reserves.
- Emergences of complex, unconventional resources, such as shale gas (characterized by drilling more wells faster), means even more data to manage and greater urgency to be able to use that data quickly to improve operations, safety, and recovery.

1.2 Benefits of Integration

Benefits of data integration include:

- **Optimize and transform workflows.** Data integration can help reduce data duplication, boost confidence in data quality, and enable automation, which combine to streamline workflows (making them shorter and faster) and reduce uncertainty, resulting in improved decision making.
• **Better business intelligence.** Integrated, accessible, and reliable data can be further analyzed for patterns and trends, with the potential to deliver comprehensive knowledge about field and business operations.

• **Master data management.** Data integration and its cousin, data quality initiatives, are key components of master data management, establishing “gold” data records that serve as the single source of truth. Master data management reduces or eliminates uncertainty for domain professionals and the time wasted trying to determine if they have the correct or latest version of data.

1.3 Challenges to Integration

The challenge for data integration is NOT technology; the technology is there. Though there are different philosophies and approaches, as IT professionals, we all know how to connect systems and map data. Trying to integrate data between different companies may pose some additional security and related technical infrastructure challenges, but these issues are also manageable, especially if understood at the beginning of an integration project and the IT support team is part of and responsive to project needs along the way.

The most difficult challenges to E&P data integration include:

• Many integration projects fail because of business reasons, not technical reasons. By definition, integration projects go across management boundaries; some short distances and others that go across silos that have become very large boundaries. The broader the scope of an integration project, the larger value it may have, but the higher the executive sponsorship needs to be. Integration budgets can become non-existent if that leadership gets tired from long drawn out difficult integration issues that were not considered at the beginning of the project.

• Lack of an industry standard taxonomy or naming system (Ziegler and Dittrich, Archer and Pritchett). Within the industry, different vendors and systems use different names for the same things. For example, one vendor’s system may use Well ID, another Well Number, and still another Well Name, to refer to the same unique well identifier. When looking at long-lived assets one must consider the entire life cycle of the well and even more changes in taxonomy, sometimes lasting over 50 years. An integration solution must be able to address and resolve these conflicting names.

• Volume and complexity of the data. Many commercial integration tools from cross-industry software vendors are typically designed to handle business or accounting data, which is characterized by many transaction and well known data types with straightforward relationships (e.g., name, account number, invoice numbers, etc.) Scientific data, such as E&P data, is characterized by huge volumes, complex relationships, and fewer transactions. Viable integration solutions require knowledge of the domain, the data, and the number and types of systems used along the E&P workflow.

• Emergence and resolution of data quality issues. Once data has been exposed to a broad audience through integration, data quality gaps will become visible and must be addressed, which often is another entire project in the midst of an integration project.

• Challenges for change management. The only thing that is certain in E&P integration is change. New applications, new versions, new data types, and sensors are emerging on a regular basis.
Changes to corporate workflows and new business partners; mergers and acquisitions all bring new challenges and opportunities to integration. Being able to accommodate these changes in a timely and cost effective manner is a huge challenge and a fundamental/high value need that must be addressed by any integration strategy.

2. Case Studies: How Integration Works in the Real World

This section provides several real-world examples based on just a few of the projects that Petris has done with our clients (some of which were first discussed in previous PNEC papers) and provides an overview of the approach and the technology used.

The projects discussed in this paper include the following:

- Full lifecycle management of wellbore data at Statoil (Nordtveit. et al.). Work to combine systems (following a merger) and automate previously manual process for receiving, loading, quality checking, and approving wellbore data including raw and processed log data. Result: A system that allowed acquisition companies to upload their raw data, then automated processes to: load and QC data; notify the acquisition company of rejected data or notify petrophysicist of data requiring approval; and publish data upon approval. The solution includes PetrisWINDS Enterprise and Recall.
A project at Baker Hughes to integrate product line data silos and improve cross-product opportunities and service (Lui and Dumothier). Baker Hughes criteria for an integration solution included: text and map/GIS-based search capabilities through the Web; minimal disruption to existing systems (and those investments); avoidance of data migration and creation of new data stores; and the use of commercially available off-the-shelf software to leverage industry knowledge. PetrisWINDS Enterprise is the foundation for this solution.

The ups and downs of automated workflows in Eni’s Well@Desk initiative (Lunghi et al.). Eni has developed new automated workflows to improve the loading and verification of borehole data into Eni’s Recall Corporate borehole database, which includes a wizard-driven automatic data load process and automated data verification process using PetrisWINDS Recall’s Raven module.
2.1 Approach
Integrating multiple complex applications and data sources is a complex process with many moving pieces and inter-related parts (The devil is in the details!). It requires a coordinated effort from a team of people. Projects typically have phased implementations with near-term and long-term objectives and goals.
For Petris projects, the process begins with planning workshops that involve representative stakeholders, including: client petrotechnical, IT, and data management (DM) professionals and technicians, and Petris consultants. These workshops can accomplish at least two key elements:
- To make sure that all stakeholders are clear on the scope, objectives, and goals of the project.
- To get all issues, obstacles, and potential pitfalls out on the table so that they can be properly addressed by the solution and the project team.

Depending on the initial scope and complexity, additional workshops are typically held.

Following the planning workshops, the team begins designing the solution and launching a proof-of-concept (POC) or pilot project to demonstrate the solutions viability. With validation from a pilot or POC project, production development begins.

The full workflow from service company acquisition to the delivery to applications for analysis is identified. Connectivity mechanisms for each component need to be examined. For commercial applications, it’s best
to use the vendor API for connectivity to preserve the application’s integrity. For Web-based or other delivery from the Service companies, each mechanism must be understood with respect to infrastructure capacity and security. Business rules related to completeness and quality must be defined. Keep in mind, all business rules, mappings, and data discrepancies must first be resolved by the project team, and then can be implemented in the solution. Petris shares “best practices” on all of these considerations.

To implement automated workflows, each workflow is mapped out and the decision points clarified with the appropriate countermeasures and responses. Domain-specific analytics are combined with data governance standards to establish the final handling of the information coming into the organization. For the preservation of interpretation results, naming practices or other ways to designate final results must be defined so that they can be extracted and returned to the archive data store (PW Recall) for long term preservation.

2.2 Solution Technology Overview
Many different integration approaches exist based on the level at which you want to integrate using a layered architecture model (e.g., user interface, application, data, etc.), and many different technological approaches for implementation have emerged (e.g., portals, data warehouses, Web services, etc.) (Ziggler and Dittrich).

Businesses working in E&P need integration solutions that address the integration challenges, help them address the industry challenges, and realize the benefits of integration (all described in Section 1).

This section briefly describes the technology used in the integration solutions of the E&P organizations in the case studies.

2.2.1 Comprehensive Data Integration and Search Engine
PetrisWINDS Enterprise (PWE) is a flexible framework that integrates data from various disparate E&P applications (Figure 4), though data remains in the native format and data store (no data warehousing required).
Data access is provided through PWE adapters, which connect each of the various systems (and require no modification to the systems being connected). The connections are mapped, maintained, and synchronized in the PWE lightweight Metadata catalog, which contains the synchronized information about the identity and location of all connected structured and unstructured collections, and includes viewing permissions, data management events, and workflow conditions.

Petris’ patented Dynamic Common Mode™ (DCM) enables integration across the broad and complex scientific and technical data of E&P by expanding dynamically to include all data types whether or not they fit into a standard data model. New data types can be mapped to existing ones or added as new instances.

Third parties can access all of PWE’s services and even connect vendor source data files through PWE’s web services module.

PWE delivers enterprise search through the definition-based Metadata catalog, which is optimized for search speed. Both map and text search functions access the entire catalog, which can automatically be updated with a user-programmable “spider” process that periodically combs through all connected applications and data sources synchronizing them with the catalog.

**2.2.2 Borehole Data Management - an Example of Data Integration/Automation**

PetrisWINDS Recall Data Management is a solution for loading, managing, storing, processing and distributing borehole data. Recall delivers all of the capabilities to manage both data used daily by project teams and corporate data stores and the archiving of well data and its fundamental design, including raw and edited tape image files of borehole data, films, plots, photos and reports. It also handles the loading and moving of data between corporate and daily stores.

The Recall database is seamlessly integrated with PetrisWINDS Enterprise described above.

The PetrisWINDS Recall Raven data quality and workflow suite delivers a robust, automated validation engine for Recall through extensive investigations into data loaded into Recall databases and rule-based evaluations of data quality and metrics. Raven significantly reduces petrotechnical staff’s time in verifying data and boosts their confidence in borehole data, resulting in better collaboration, streamlined workflows, and quicker decision-making.

**3. Results**

**3.1.1 Now Possible for Organizations to Systematically Resolve Data Quality Issues**

There is no avoiding it: Integration brings into the full light of day data quality problems. Ideally, the right integration solution should include the tools and the ability to automate the process of cleaning up data and implementing an automated QC and correction process to keep the data quality high.

All three project teams had to deal with data quality issues upon integration of systems.
In the case of Statoil and a merger with another company, the team was able to step back and set clear goals for the integration and rationalization processes of systems and databases, which included (Nordtveit. et al.):

- Improved data management with more efficient work processes and reduced cycle times
- Improved availability of all data, especially raw data and interpreted high-business value data
- Improved reporting to authorities and partners
- Improved decisions for exploration, field development, and business development
- Improved ability to perform evaluations according to best practices

For Eni, as part of the integration project the team developed borehole data standards, for example, naming conventions for log curves, well attributes, and for identifying raw, edited, and interpreted data sets. These standards made it possible for users to more easily locate, identify, and retrieve the correct data they required.

The application of these standards to existing data flagged incorrect and duplicate names. Some of these could be resolved automatically, but many required human review and intervention to correct. However, the solution provided a systematic method for the identification, review, and correction process. Implementation of standards into the solution and accompanying changes in work processes should significantly reduce—and potentially eliminate—these types of errors in the future.

Baker Hughes also reports that their integrated solution is providing a systematic way to identify and resolve similar types of data discrepancies.

3.1.2 Automated Workflows
Automated workflows will be a crucial part of transforming E&P operations to reduce cycle times and realize new value. Some of the details described in these papers make it absolutely clear that reliable quality data is paramount for automation.

At Statoil, the team developed a process that let acquisition companies upload data that then kicked off an automated data verification and QC process. If the data failed the test, automatic notification was sent to the supplying vendor. If the data passed the test, automatic notification and a summary report were sent to the responsible petrophysicist who could then review and approve (or reject) the data. Approval by the petrophysicist launched the procedure to publish the data to the corporate data store.

At Eni, the team implemented a solution to automate data loading and verification to the company’s Recall corporate database. From a desktop wizard, data loaders can prepare and submit request and then a series of QC checks is automatically performed. If the data passes all tests, it is loaded into the Recall Database.

Each of these automated processes eliminated manual, time-consuming, and error-prone processes.
3.1.3 - Plans to Include More Data Types, More Automation, and Focus on Domain Processes

Upon the completion of the current phase of these projects, the next steps for each project team were nearly unanimous—MORE!

Eni plans to look at expanding the scope to integrate additional systems and data sources to include mud logs and deviations surveys. The project team also plans to look next at improving its pre-drilling processes and procedures.

Statoil reports several ideas to explore in the future which include: investigating the ability to pre-scan data sets submitted by vendors to pre-populate the metadata associate with each file in the data sets that the vendors upload; automation of data distribution from Recall Corporate databases to application/project data stores; and extending the data tracking system for other Statoil initiatives to track all documents delivered to the company.

With several key systems now integrated, teams at Baker Hughes have been focusing on resolving discrepancies and cleaning up data. They are beginning to explore how better availability and access to quality data can optimize their workflows and allow them to better exploit cross-product-line opportunities.

3.2 - Changing Roles

Integrated data and applications promote collaboration among petrotechnical professionals and members of asset teams. But to develop integrated solutions, you must first have collaboration among members of the project teams (as described in Section 2.1 above) including the client organizations petrotechnical, IT, and DM professionals and technicians, and consultants.

For Petris, our solutions have been designed to be flexible and easily integrate with all major applications. However, because of the uniqueness of each client’s IT environment, workflows, and processes, the project teams must collaborate closely to ensure best implementation of the solution, which is typically an iterative process to work through the issues and fine-tune the performance.

As a result, people’s roles will typically change some during the development process and with the implementation of an integrated solution.

3.2.1 - Petrotechnical Professionals

Petrotechnical professionals are the ultimate end-users—they will define requirements and verify prototypes and pilot projects. They must be involved all along the way.

If tasks previously done by data management technicians move to new “self-serve” features, that will change the role of petrotechnical professional. Processes within the solution should be well-designed to make them as simple as possible so that petrotechnical professionals can continue to focus on their domain work.
For example, in the Statoil solution where after the automated verification process a Petrophysicist must approve all data before it is published to the corporate database, a well-designed approval request email is sent, clearly showing the results of the test, providing links for the approver to drill down on data, and with an approval link that launches the wizard to kick off next steps in the process.

3.2.2 - IT and Data Management Professionals and Technicians
Client IT and DM professionals must help consultants understand the necessary details of the client environment and systems so clients can accurately and effectively make recommendations and collaborate with IT professionals to implement the integration solution.

While data integration solutions should ultimately reduce the workloads of IT and DM workers, working the bugs out of newly integrated solutions and automated processes can mean an initial increase in workload and responsibility.

At Eni, the new automated workflows were more complex and created some intricate interdependencies in the system. In the beginning, discrepancies or errors in the process would flag problems that would have to be addressed by the Recall database administrator and systems specialists. Overtime, if the system has been properly designed, the issues will reduce, which was the expectation of Eni.

There is also the potential for some technician jobs to change dramatically (or perhaps be eliminated). At Eni, technicians who used to manually load the data now perform data loading electronically using the submission wizard.

3.2.3 - Consultants
At Petris, we’ve known for a long time that you don’t just “sell” sophisticated and robust software solutions, plug them in and walk out the door; consulting services typically help clients to more fully realize all the potential of such solutions faster.

With integration solutions, the above projects show how the best results come when consultants work as key members of the collaborative team, listening to the needs and challenges of the client, contributing based on their expertise and experience, and leading in areas around the solution technology and implementation.

Additionally, we’ve found that implementation of these solutions is not the end of the engagement. Still, in relatively early stages, once clients begin to integrate their data and see the results, they see many more possibilities and seek assistance with the next phases and new projects.

3.2.4 - Service Companies
Service companies working with operators must be made aware of and educated on new processes and data formats required. In the case of Statoil, service companies need to understand the required data formats and the requirements and processes for getting a user ID and password and submitting data to the system.
4. - Conclusions and Recommendations

4.1 - Conclusions

- E&P businesses are realizing the benefits of data integration—now. The process of moving to an integrated solution may be initially more complex and labor intensive, but the value is there. The payoff of data integration is more reliable, better quality data that can be used for workflow optimization and automation, better decisions, and better business intelligence. Organizations typically integrate with specific goals in mind. However, after integration, more opportunities and benefits of integration quickly become apparent.
- When you integrate and can see all your data, you will definitely discover data quality problems. The good news is with the right integration solution you can also have the tools to address those data quality problems and resolve them with automated solutions to ensure data quality is managed at the point of entry and maintained.
- For integration and automation to work successfully, a great deal of collaboration from all parties is required. It will be an iterative process to get the solutions where you want them to be.
- Roles will change for many people within the organization and for external partners, such as consultants, service companies, and vendors who interact with the organization around these transformed processes. New roles are also being established such as the Data Management specialist, a person, or even a department, well versed in IT, Information Management, and E&P (for example the PPDM Petroleum Data Manager (PDM) credentialed program, currently being developed by the Petroleum Education Task Force).

4.2 - Recommendations

- Plan appropriately for the time and resources required to effectively implement integrated solutions. Understand your data and environment; map them carefully during the planning and development process; and implement, troubleshoot, and fine-tune processes and technology.
- Involve representatives from all key stakeholder groups, from the initial discovery and development phase of the project, throughout the full project life cycle, and into subsequent phases as necessary.
- Do your research and choose the right tools and technology: flexible, powerful solutions that allow you to integrate best-of-breed applications from any vendor and solutions that are field-proven.
- Work with experienced consultants who understand the data and challenges of E&P as well as the information technology and solutions and can help you realize the full potential of data integration.
References


