



# BIG OIL, BIG DATA, BIG VALUE

By Sylvain Santamarta, Peter Forbes, Rash Gandhi, and Michael Bechauf

**I**T'S NO SECRET THAT these are tough times for the oil and gas industry. In the four years following oil's steep drop in 2014, price volatility grew 12%. Alternative forms of energy are becoming increasingly popular, and the price of a barrel of oil remains low. In the face of these challenges, oil and gas companies are implementing digital technologies to drive efficiency throughout the value chain and become more resilient. But success has been limited. One key reason that is often overlooked: an inability to fully leverage data.

The IT systems at most oil and gas companies include a multitude of legacy software applications purchased from various vendors in different formats across many different functions. The inflexible architecture of these systems often makes the underlying data inaccessible. In addition, the industry has failed to maintain a strong focus on data quality. Companies have preferred to give each function the flexibility to store data in whatever way it finds most useful and to use the data formats provided by

vendors. This has made it difficult to build repeatable use cases or to combine data in cross-functional use cases. Compounding these challenges, the IT organization is often set up to provide solutions tailored for individual functions rather than for use across the value chain.

To take full advantage of their data and develop the digital capabilities critical for competitive advantage, companies need to build a central platform that includes a data warehouse, a data lake, or both. This approach will make data easier to find and analyze and will aid in the development of digital solutions that deliver value. It will also centralize data governance, allowing for a cross-functional perspective and making it easier to provide the guidance necessary for ensuring data availability, quality, and security for the entire organization. A few oil and gas companies are already seeing the benefits of data platforms. One international oil company, for example, credited its platform investments with annual cost savings of \$7 billion over a recent three-year period.

## Choosing the Best Way to Store Your Data

Oil and gas companies can build a centralized data platform in one of three ways. They can move their information to a data warehouse, a repository of structured data. They can put it in a data lake, a vast repository of structured and unstructured data. Or they can use a warehouse and a lake in tandem. (See the exhibit.)

**Data Warehouse.** Data warehouses are most valuable in use cases with structured data tied to a set of specific and well-known data sources. They are useful for activities such as regulatory and compliance reporting, where the accuracy and consistency of insights are important and data is structured and unchanging. For example, a global oil company that wanted to develop a common reporting solution for a local business unit recently built a data warehouse that would provide a single source of truth for management dashboards.

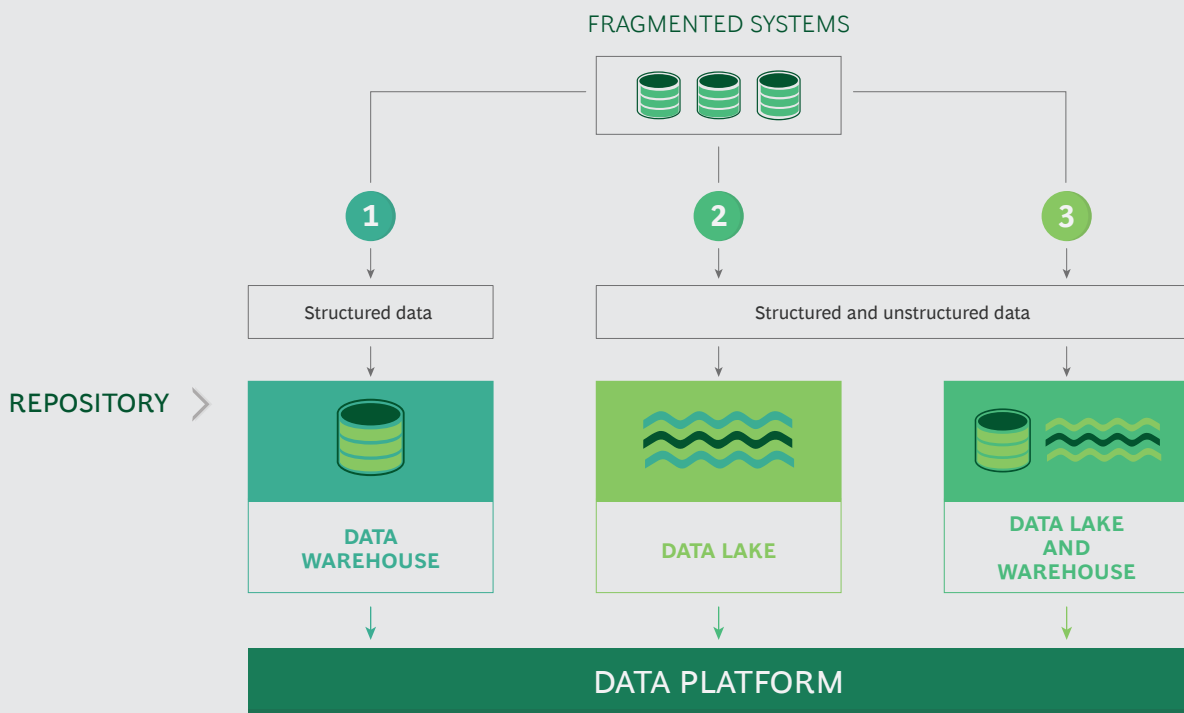
Data warehouses are proven technologies, with many solutions, vendors, and experts

readily available. And because they use preconfigured combinations, they can produce analytics results fast. But they have their disadvantages. Time to value is high because data must go through a lengthy structuring process before it can be stored in a warehouse. This means that the user must know upfront which data will be stored. In addition, the need for structuring makes incorporating new data sources difficult; for data from unstructured or semi-structured sources, it's nearly impossible.

**Data Lake.** Building a data lake requires simply loading the information. This greatly accelerates the time to value because the data does not need to be structured first; that can be done as needed for specific use cases.

In addition, because data lakes get their storage capabilities from distributed commoditized hardware and open-source software rather than from legacy systems, they can scale to enormous capacity. And structured and unstructured information can be stored in infinite combinations. Data lakes can, for example, include large data stores

Three Ways to Build a Data Platform



Source: BCG analysis.

that are essential to the high-frequency time series data common in both upstream and downstream production. Such flexibility and scalability allow companies to adopt the latest digital technologies.

But there's a downside: because the information hasn't been structured, data lakes require more rigorous governance and management than warehouses. They also require people with data lake architecture and data engineering skills, who are far scarcer than data warehouse experts.

When a national oil company embarked on its digital transformation journey recently, it put all its data in a data lake. Given that there was no data warehouse in place, this provided the shortest time to value as well as the greatest flexibility.

**Data Warehouse in Tandem with a Data Lake.** Companies that already have a data warehouse can also deploy a data lake for new use cases that have additional requirements. They can migrate use cases that combine structured and unstructured data into a data lake, while limiting the use of a data warehouse to use cases that require high-performance, repeatable, and auditable results. Some modern cloud-based data warehouses such as Snowflake or Amazon Redshift also allow combined queries across structured, semistructured, and unstructured content. But maintaining multiple repositories adds complexity and cost.

An international oil company recently chose to add a data lake, because it had an established production-reporting data warehouse but needed a data lake for new digital use cases that required both structured and unstructured data from several different sources. Only by adding a data lake could the company meet these new requirements and have a short time to value.

## Building and Deploying the Data Platform

Once a company has determined which type of data repository will work best, the next step is to build out the platform with

use cases and the data they require, and scale it to full realization. The following practices are key.

**Decide on your optimal solution.** Companies can build their own platform using a generic cloud service such as AWS, Google, or Azure, or they can purchase a data suite from an industry vendor such as Schlumberger DELFI or Palantir. In either case, the choice must offer enough flexibility to support the digital transformation. A generic cloud service will offer high flexibility and many components that can be used across different functional and cross-functional use cases. But it may take more time to ingest data; it also may require custom components.

Vendor expertise and ready-made components for data ingestion, storage, and modeling help data suites reduce time to value. But they can restrict solution design to the options that the suites employ.

**Define what belongs in the optimal data platform.** The company's strategic goals should dictate which kinds of use cases and data to store in a platform. Specifically, this means selecting use cases that will leverage the platform's functionality to provide business value.

**Decide how to integrate your data.** There are two options. Companies can build either a single data platform shared across the company or multiple data platforms with a common mechanism for integrating data sources (a virtualization layer).

A single platform provides seamless access to data across the value chain but requires more effort. Moreover, the wide variety of data involved may create technical challenges.

Data residence regulations can sometimes force oil and gas companies to have multiple platforms. These require less effort to build but can lead to performance issues when working across data sets. In addition, it's less clear where to store cross-functional use case results. Establishing some guidelines—such as sticking to a single technology solution

and colocating data to minimize its movement across platforms—will help avoid performance and cost issues, reduce application latency, and simplify data governance.

#### Build out and scale up incrementally.

Often companies want to show value by adding many types of data to the platform all at once. But if they don't take the time to integrate the data, the use cases won't be scalable and will provide little additional value. Instead, we recommend first developing and integrating individual use case minimum viable products (MVPs) in a simplified architecture stack. This requires building automated pipelines that feed data from source systems to the data platform, which is then able to generate analytics and visualizations to deliver insights to business users.

To scale up successfully, companies need to follow a fail-fast approach, integrating and scaling each use case only after it has demonstrated value. This method makes it possible to adjust the solution as technology develops. And it will provide a broad range of functional and cross-functional solutions that provide new business insights. Even simple use cases, such as the combination of time series sensor data and maintenance work order data into a single view, can provide engineers with a valuable source of information for troubleshooting. The same data can then enable predictive maintenance solutions for the medium term.

#### Establish governance on three fronts.

To prevent unnecessary use cases and other sources of inefficiency, it's essential to put solid controls in place in the following areas:

- **Use Cases.** To ensure that the platform houses only use cases that lead to value, the governing group will need to prioritize the use cases that will produce the highest return.
- **Technology.** It's also important to decide which technologies to deploy. Since not all requirements will be clear initially, we suggest giving teams the freedom in the early stages to select

technology components for rapid development, then later standardizing and rationalizing these components. This approach will allow flexibility in the short term before requirements and then components are locked in place.

- **Data.** Equally important, the governing group must appoint someone to determine where master data will be held, what that data will look like, and who will be responsible for managing different data elements to ensure accessibility, security, and quality. This will help steadily improve data quality over time, thus enabling rapid analytics.

**Acquire the necessary capabilities.** Deploying data technologies requires skills and experience that some oil and gas companies don't have. In order to create their own comprehensive data platform, companies need to either build internal capabilities or team up with a partner. Third parties can provide a range of architectural options, from data lake distributions, which are suitable for companies that want an onsite data center or cloud-based implementation, to full platform-as-a-service offerings.

Whichever option is preferable, oil and gas companies will require some internal capabilities to build the platform. Data architects, data engineers, and others with data skills will be essential in designing and implementing the solution and engineering the data for use cases and analytics. These employees will also be needed after the platform is built to ensure maintenance and further development.

## Reaping the Benefits

Once fully realized, the data platform provides tremendous transparency not only across the organization but throughout the entire value chain, from exploration to production and beyond. By freeing data from siloes, a flexible data platform enables new cross-functional use cases—such as automated well design, which requires data from a variety of functions, including geology, geophysics, and engineering.

This type of platform provides the same benefits across the entire value chain. Upstream subsurface data stores can provide nonexploration business segments better access to exploration data. Downstream, data from different channels can be combined to develop a better understanding of the customer, allowing the company to better target offerings. Additionally, predictive analytics can be used to optimize ordering, storage, and utilization of materials across the value chain. The potential is huge.

**A**S THE NEED for digitization grows, so will the need for data platforms. It's therefore important for companies to regard them as an investment that will require much in the way of time and resources but will offer long-term benefits. By adding more and more data in a managed way, companies can make a data platform a powerful source of advantage.

### About the Authors

**Sylvain Santamarta** is a managing director and partner in the Amsterdam office of Boston Consulting Group. You may contact him by email at [santamarta.sylvain@bcg.com](mailto:santamarta.sylvain@bcg.com).

**Peter Forbes** is an associate director in BCG's London office with expertise in digital energy. You may contact him by email at [forbes.peter@bcg.com](mailto:forbes.peter@bcg.com).

**Rash Gandhi** is a partner and associate director in the firm's London office with expertise in data and digital platforms. You may contact him by email at [gandhi.rash@BCG.com](mailto:gandhi.rash@BCG.com).

**Michael Bechauf** is a partner and associate director in BCG's Silicon Valley office with expertise in enterprise software. You may contact him by email at [bechauf.michael@BCG.com](mailto:bechauf.michael@BCG.com).

Boston Consulting Group partners with leaders in business and society to tackle their most important challenges and capture their greatest opportunities. BCG was the pioneer in business strategy when it was founded in 1963. Today, we help clients with total transformation—inspiring complex change, enabling organizations to grow, building competitive advantage, and driving bottom-line impact.

To succeed, organizations must blend digital and human capabilities. Our diverse, global teams bring deep industry and functional expertise and a range of perspectives to spark change. BCG delivers solutions through leading-edge management consulting along with technology and design, corporate and digital ventures—and business purpose. We work in a uniquely collaborative model across the firm and throughout all levels of the client organization, generating results that allow our clients to thrive.

© Boston Consulting Group 2019. All rights reserved. 8/19

For information or permission to reprint, please contact BCG at [permissions@bcg.com](mailto:permissions@bcg.com). To find the latest BCG content and register to receive e-alerts on this topic or others, please visit [bcg.com](http://bcg.com). Follow Boston Consulting Group on Facebook and Twitter.