

# The Value of Implementing DataOps on the Oil & Gas Industry



## INTRODUCTION

In the age of big data, the ability to extract insights from data is a critical capability for businesses in every industry. The oil and gas industry is no exception, with companies relying on data to optimize operations, improve safety, and reduce costs. However, analyzing data in the oil and gas industry can be a complex and challenging task due to the large volumes of data generated by different systems and equipment, as well as the presence of data silos. Data silos occur when data is stored in different systems that are not easily integrated, leading to data fragmentation, and making it difficult to analyze data across different sources. This is where DataOps comes in – a methodology that aims to improve collaboration and communication among different teams involved in the data pipeline, including data engineers, data analysts, and data scientists. In this essay, we will explore best practices for using DataOps to analyze data from different data silos in the oil and gas industry.

## BEST PRACTICES FOR ANALYZING DATA FROM DIFFERENT DATA SILOS USING DATAOPS

### Identify and prioritize data sources

The first step in using DataOps to analyze data from different data silos is to identify and prioritize the data sources that are most important for your analysis. This can involve collaborating with stakeholders from different teams and understanding their data needs and requirements. Once you have identified the key data sources, it is important to prioritize them based on their importance and impact on the business.

### Integrate data from different sources

Integrating data from different sources is essential to enable cross-functional teams to access and analyze data from different data silos. This can involve using tools like Apache Kafka or Apache Nifi to ingest data from different sources and integrating them into a unified data platform. It is important to ensure that the data is integrated in a way that maintains data integrity and consistency, as well as adhering to any data governance policies and regulations.

### Standardize data

Standardizing data is important to ensure that data from different sources is consistent and can be analyzed using the same tools and techniques. This can involve creating data dictionaries and data models to standardize data across different sources. Standardizing data can also help to improve data quality and reduce the time required for data preparation.

## Implement data quality checks

Implementing data quality checks is essential to ensure that data from different sources is accurate and reliable. This can involve using automated data quality checks to flag any inconsistencies or errors in the data. Data quality checks can include data profiling, data cleansing, and data validation. By implementing data quality checks, you can improve the accuracy and reliability of your data analysis and reduce the risk of making decisions based on inaccurate data.

## Enable self-service analytics

Enabling self-service analytics is essential to empower cross-functional teams to access and analyze data from different data silos. This can involve providing access to a centralized data platform and using tools like Tableau or Power BI to enable self-service analytics. Enabling self-service analytics can help to reduce the workload of data analysts and enable other teams to derive insights from the data, leading to improved decision-making and increased efficiency.

## Implement version control

Implementing version control is essential to ensure that changes to the data pipeline are tracked and managed effectively. This can involve using tools like Git to manage changes to code and configuration files. By implementing version control, you can ensure that changes are made in a controlled and consistent manner, reducing the risk of errors, and enabling you to easily roll back changes if necessary.

## Automate data pipelines

Automating data pipelines is essential to enable fast and reliable data analysis. This can involve using tools like Apache Airflow to automate data workflows and schedule data pipelines. Automating data pipelines can help to reduce the time required for data preparation and enable analysts to focus on data analysis and insights.

## Monitor and optimize data pipelines

Monitoring and optimizing data pipelines is important to ensure that data is flowing smoothly through the data pipeline and to identify and resolve any issues or bottlenecks. This can involve using monitoring tools like Prometheus or Grafana to monitor data pipeline performance and identify any issues or bottlenecks. Once issues are identified, they can be resolved by optimizing the data pipeline, such as by adding more resources or changing the data pipeline configuration.

## Collaborate and communicate effectively.

Collaboration and communication are key components of the DataOps methodology. In order to effectively analyze data from different data silos, it is important to establish clear communication channels and collaborative processes between different teams involved in the data pipeline, including data engineers, data analysts, and data scientists. This can involve using collaboration tools like Slack or Microsoft Teams to enable real-time communication and collaboration.

## Continuously improve and iterate

DataOps is an iterative process that involves continuous improvement and iteration. This means that you should regularly review and refine your data pipeline and data analysis processes to ensure that they are meeting the needs of the business. This can involve using feedback from stakeholders and users to identify areas for improvement and implementing changes to improve the data pipeline and analysis processes.

## Ensure data security and compliance

Ensuring data security and compliance is essential in the oil and gas industry, which is subject to strict regulations and data privacy requirements. This can involve implementing data security measures like encryption and access controls to ensure that data is protected from unauthorized access. It is also important to ensure that the data pipeline and analysis processes comply with relevant regulations and policies, such as the General Data Protection Regulation (GDPR) and the Payment Card Industry Data Security Standard (PCI DSS).

## Build a data-driven culture

Building a data-driven culture is essential to enable effective data analysis in the oil and gas industry. This means fostering a culture that values data and encourages data-driven decision-making. This can involve providing training and resources to enable employees to become proficient in data analysis tools and techniques, as well as establishing metrics and KPIs to measure the effectiveness of data analysis and use of data in decision-making.

## CONCLUSION FOR THE OIL & GAS INDUSTRY

Analyzing data from different data silos in the oil and gas industry can be a complex and challenging task. However, by implementing best practices for using DataOps, businesses can effectively analyze data from different sources and derive insights that can help to optimize operations, improve safety, and reduce costs. Key best practices include identifying and prioritizing data sources, integrating data from different sources, standardizing data, implementing data quality checks, enabling self-service analytics, implementing version control, automating data pipelines, monitoring, and optimizing data pipelines, collaborating, and communicating effectively, continuously improving and iterating, ensuring data security and compliance, and building a data-driven culture. By following these best practices, businesses in the oil and gas industry can effectively analyze data from different data silos and gain a competitive advantage through data-driven decision-making.

Plant managers in the oil and gas industry face a variety of challenges when it comes to managing operations effectively. They are responsible for ensuring that their facilities are operating efficiently and safely, while also managing costs and adhering to strict regulatory requirements. One key area where plant managers can benefit from using DataOps is in analyzing data from different data silos to gain insights into operations and make data-driven decisions. In this article, we will explore the benefits of using DataOps to analyze data from different data silos and provide practical guidance for plant managers looking to implement DataOps best practices.

# The importance of DataOps in Plant Operations



In the oil and gas industry, data is a critical component of plant operations. From monitoring equipment performance to tracking production levels and safety incidents, data provides valuable insights into the health of plant operations. However, the challenge for plant managers is that data is often scattered across different data silos, including operational databases, process historians, and data warehouses. This makes it difficult to access, integrate, and analyze data in a timely and efficient manner.

Moreover, data silos often contain data that is not consistent, complete, or accurate, which can lead to errors and incorrect conclusions. For example, data may be missing or duplicated, or it may be stored in different formats that are not compatible with each other. These data quality issues can make it difficult to trust the data and make informed decisions based on it.

This is where DataOps comes in. By implementing best practices for integrating and analyzing data from different data silos, plant managers can gain a complete and more accurate picture of plant operations. They can identify areas for improvement, optimize processes, and reduce costs. They can also improve safety by identifying potential hazards and taking corrective action before incidents occur.

## Benefits of using DataOps in Plant Operations

There are many benefits to using DataOps in plant operations.

Some of the key benefits include:

- 1** Improved operational efficiency: By analyzing data from different data silos, plant managers can identify bottlenecks and inefficiencies in plant operations. They can optimize processes and reduce downtime, resulting in improved operational efficiency.

- 2** Increased safety: By analyzing data from different data silos, plant managers can identify potential safety hazards and take corrective action before incidents occur. This can help to improve safety and reduce the risk of accidents.

- 3 **Reduced costs:** By analyzing data from different data silos, plant managers can identify areas where costs can be reduced. This can include optimizing energy usage, reducing waste, and streamlining processes.
- 4 **Improved product quality:** By analyzing data from different data silos, plant managers can identify areas where product quality can be improved. This can help to increase customer satisfaction and reduce the risk of product recalls.
- 5 **Better decision-making:** By analyzing data from different data silos, plant managers can make more informed decisions based on data-driven insights. This can help to improve overall business performance and competitiveness.

## Best Practices for using DataOps in Plant Operations

To effectively use DataOps in plant operations, plant managers should follow best practices for integrating and analyzing data from different data silos.

These best practices include:

- 1 **Identify and prioritize data sources:** Plant managers should identify the data sources that are most important for plant operations. This may include operational databases, process historians, and data warehouses. They should prioritize data sources based on their relevance to plant operations and their potential to provide valuable insights.
- 2 **Integrate data from different sources:** Once data sources have been identified, plant managers should integrate data from different sources. This may involve using tools like data integration platforms and data pipelines to extract, transform, and load data from different sources into a central data repository.
- 3 **Standardize data:** To ensure that data is consistent and accurate, plant managers should standardize data from different sources.
- 4 **Ensure data quality:** To ensure that data is accurate and complete, plant managers should implement data quality controls. This may include data profiling to identify data quality issues, data cleansing to remove errors and inconsistencies, and data validation to ensure that data meets quality standards.
- 5 **Implement data governance:** To ensure that data is used in a consistent and responsible manner, plant managers should implement data governance processes. This may involve creating policies and procedures for data management, data security, and data privacy. Data governance processes can help to ensure that data is used ethically and in compliance with regulatory requirements.
- 6 **Use data visualization tools:** To make data more accessible and understandable, plant managers should use data visualization tools. These tools can help to create visual representations of data that are easy to interpret and analyze. Data visualization tools can help plant managers to identify trends and patterns in data that may not be apparent in tabular or numerical form.

**7** Leverage machine learning and AI: To automate data analysis and gain insights more quickly, plant managers should leverage machine learning and AI technologies. These technologies can help to identify patterns and anomalies in data that may not be immediately apparent to human analysts. They can also help to automate routine data analysis tasks, freeing up time for plant managers to focus on more strategic activities.

**8** Collaborate across teams: To ensure that data is used effectively and to its full potential, plant managers should encourage collaboration across different teams. This may include data scientists, data engineers, and subject matter experts. Collaboration can help to identify new insights and opportunities for optimization that may not be apparent to individual team members.

**9** Continuously monitor and improve data processes: To ensure that data processes are effective and efficient, plant managers should continuously monitor and improve data processes. This may involve using metrics to track performance and identify areas for improvement, and implementing process improvements to optimize data workflows.

**10** Invest in training and education: To ensure that staff are equipped with the skills and knowledge to effectively use data, plant managers should invest in training and education. This may include training on data analysis tools and techniques, as well as on data governance and security best practices. By investing in training and education, plant managers can ensure that staff are able to use data effectively and responsibly.

## CASE STUDY - 1

### Shell

Shell is a multinational oil and gas company that has implemented a DataOps strategy to improve its operations. The company has a vast amount of data from its operations, including data from oil wells, refineries, and distribution channels. However, the data was often fragmented and siloed, making it difficult to analyze and gain insights.

To overcome this challenge, Shell implemented a DataOps strategy that involved integrating data from different sources, standardizing data, and implementing data quality controls. The company also implemented data visualization tools and machine learning algorithms to automate data analysis and gain insights more quickly.

As a result of these efforts, Shell has been able to optimize its operations and reduce costs. For example, the company was able to reduce drilling times by up to 30% by analyzing data from drilling operations and identifying areas for improvement. Shell has also been able to optimize its distribution channels, reducing transportation costs by up to 20% by analyzing data from shipping operations.



# CASE STUDY - 2

## *ExxonMobil*

---



ExxonMobil is another multinational oil and gas company that has implemented a DataOps strategy to improve its operations. The company has a large amount of data from its operations, including data from oil wells, refineries, and distribution channels. However, the data was often fragmented and siloed, making it difficult to analyze and gain insights.

To overcome this challenge, ExxonMobil implemented a DataOps strategy that involved integrating data from different sources, standardizing data, and implementing data quality controls. The company also implemented data visualization tools and machine learning algorithms to automate data analysis and gain insights more quickly.

As a result of these efforts, ExxonMobil has been able to optimize its operations and reduce costs. For example, the company was able to optimize its drilling operations by analyzing data from drilling rigs and identifying areas for improvement. The company was also able to reduce transportation costs by optimizing shipping routes based on real-time data.

In addition to optimizing operations, ExxonMobil has also used DataOps to improve safety and environmental performance. The company has implemented predictive maintenance models to identify potential equipment failures before they occur, reducing the risk of accidents and downtime. The company has also used machine learning algorithms to monitor emissions and reduce environmental impact.

Overall, the implementation of DataOps has helped ExxonMobil to stay competitive in a rapidly evolving industry. By using data to make informed decisions and optimize operations, the company has been able to reduce costs, improve safety, and reduce environmental impact.

# CASE STUDY - 3

## *Chevron*

---



Chevron is another multinational oil and gas company that has implemented a DataOps strategy to improve its operations. The company has a large amount of data from its operations, including data from oil wells, refineries, and distribution channels. However, the data was often fragmented and siloed, making it difficult to analyze and gain insights.

To overcome this challenge, Chevron implemented a DataOps strategy that involved integrating data from different sources, standardizing data, and implementing data quality controls. The company also implemented data visualization tools and machine learning algorithms to automate data analysis and gain insights more quickly.

As a result of these efforts, Chevron has been able to optimize its operations and reduce costs. For example, the company was able to reduce maintenance costs by up to 20% by analyzing data from equipment sensors and identifying potential failures before they occur. The company was also able to optimize its supply chain by analyzing data from shipping operations and identifying opportunities for consolidation.

In addition to optimizing operations, Chevron has also used DataOps to improve safety and environmental performance. The company has implemented predictive maintenance models to identify potential equipment failures before they occur, reducing the risk of accidents and downtime. The company has also used machine learning algorithms to monitor emissions and reduce environmental impact.

Overall, the implementation of DataOps has helped Chevron to stay competitive in a rapidly evolving industry. By using data to make informed decisions and optimize operations, the company has been able to reduce costs, improve safety, and reduce environmental impact.



## CONCLUSION

DataOps is an approach to data management that emphasizes collaboration, automation, and continuous improvement. By implementing DataOps best practices, plant managers in the oil and gas industry can effectively analyze data from different data silos, identify trends and patterns, and make informed decisions to optimize operations.

To implement DataOps best practices, plant managers should focus on integrating data from different sources, standardizing data, ensuring data quality, implementing data governance, using data visualization tools, leveraging machine learning and AI, collaborating across teams, continuously monitoring and improving data processes, and investing in training and education.

The implementation of DataOps has helped companies like Shell, ExxonMobil, and Chevron to optimize operations, reduce costs, improve safety, and reduce environmental impact. By following these best practices, plant managers in the oil and gas industry can also benefit from the advantages of DataOps and stay competitive in a rapidly evolving industry.



## REFERENCES

- "DataOps: An Overview." Agile Data Warehousing Project Management. Accessed 15 October 2021. <https://www.agiledatawarehouseproject.com/dataops-an-overview/>
- "DataOps: A New Approach to Data Management." Deloitte. Accessed 15 October 2021. <https://www2.deloitte.com/us/en/insights/focus/tech-trends/2019/data-ops-data-management.html>
- "DataOps: The Agile Way to Manage Data." CA Technologies. Accessed 15 October 2021. <https://www.ca.com/content/-dam/ca/us/files/ebook/dataops-the-agile-way-to-manage-data-ebook.pdf>
- "DataOps: Improving Data Quality in the Era of Big Data." Informatica. Accessed 15 October 2021. <https://www.informatica.com/content/dam/informatica-com/global/amer/us/pdfs/solution-briefs/br-dataops-improving-data-quality-in-the-era-of-big-data.pdf>
- "Shell Uses DataOps to Streamline Analytics." CIO. Accessed 15 October 2021. <https://www.cio.com/article/3201445/shell-uses-dataops-to-streamline-analytics.html>
- "DataOps: Best Practices for Data Management." DataKitchen. Accessed 15 October 2021. <https://www.datakitchen.io/blog/dataops-best-practices-for-data-management/>
- "DataOps: Data Management for the Agile Enterprise." Gartner. Accessed 15 October 2021. <https://www.gartner.com/smarterwithgartner/dataops-data-management-for-the-agile-enterprise/>
- "ExxonMobil Invests in DataOps to Fuel Digital Transformation." TechHQ. Accessed 15 October 2021. <https://techhq.com/2021/04/exxonmobil-invests-in-dataops-to-fuel-digital-transformation/>
- "Chevron Uses Machine Learning to Optimize Operations." Datanami. Accessed 15 October 2021. <https://www.datanami.com/2020/10/26/chevron-uses-machine-learning-to-optimize-operations/>
- "DataOps: Managing Data for Analytics in the Modern Enterprise." IBM. Accessed 15 October 2021. <https://www.ibm.com/downloads/cas/8YJYVKRZ>