

What is DataOps? Collaborative, cross-functional analytics

Olavsrud, Thor . Olavsrud, Thor.

[ProQuest document link](#)

ABSTRACT (ENGLISH)

According to Gartner, DataOps also aims “to deliver value faster by creating predictable delivery and change management of data, data models, and related artifacts.” DataOps in practice To make the most of DataOps, enterprises must evolve their data management strategies to deal with data at scale and in response to real-world events as they happen, according to Dunning and Friedman. Because DataOps builds on DevOps, cross-functional teams that cut across “skill guilds” such as operations, software engineering, architecture and planning, product management, data analysis, data development, and data engineering are essential, and DataOps teams should be managed in ways that ensure increased collaboration and communication among developers, operations professionals, and data experts. According to Michele Goetz, vice president and principal analyst at Forrester, some of the key areas of expertise on DataOps teams include:

FULL TEXT

What is DataOps?

DataOps (data operations) is an agile, process-oriented methodology for developing and delivering analytics. It brings together DevOps teams with data engineers and data scientists to provide the tools, processes, and organizational structures to support the data-focused enterprise. Research firm Gartner further describes the methodology as one focused on “improving the communication, integration, and automation of data flows between data managers and data consumers across an organization.”

DataOps goals

According to Dataversity, the goal of DataOps is to streamline the design, development, and maintenance of applications based on data and data analytics. It seeks to improve the way data are managed and products are created, and to coordinate these improvements with the goals of the business. According to Gartner, DataOps also aims “to deliver value faster by creating predictable delivery and change management of data, data models, and related artifacts.”

DataOps vs. DevOps

DevOps is a software development methodology that brings continuous delivery to the systems development lifecycle by combining development teams and operations teams into a single unit responsible for a product or service. DataOps builds on that concept by adding data specialists —data analysts, data developers, data engineers, and/or data scientists—to focus on the collaborative development of data flows and the continuous use of data across the organization.

DataKitchen, which specializes in DataOps observability and automation software, maintains that DataOps is not simply “DevOps for data.” While both practices aim to accelerate the development of software (software that leverages analytics in the case of DataOps), DataOps has to simultaneously manage data operations.

DataOps principles

Like DevOps, DataOps takes its cues from the agile methodology. The approach values continuous delivery of analytic insights with the primary goal of satisfying the customer.

According to the DataOps Manifesto, DataOps teams value analytics that work, measuring the performance of data

analytics by the insights they deliver. DataOps teams also embrace change and seek to constantly understand evolving customer needs. They self-organize around goals and seek to reduce “heroism” in favor of sustainable and scalable teams and processes.

DataOps teams also seek to orchestrate data, tools, code, and environments from beginning to end, with the aim of providing reproducible results. Such teams tend to view analytic pipelines as analogous to lean manufacturing lines and regularly reflect on feedback provided by customers, team members, and operational statistics.

Where DataOps fits

Enterprises today are increasingly injecting machine learning into a vast array of products and services and DataOps is an approach geared toward supporting the end-to-end needs of machine learning.

“For example, this style makes it more feasible for data scientists to have the support of software engineering to provide what is needed when models are handed over to operations during deployment,” Ted Dunning and Ellen Friedman write in their book, *Machine Learning Logistics*.

“The DataOps approach is not limited to machine learning,” they add. “This style of organization is useful for any data-oriented work, making it easier to take advantage of the benefits offered by building a global data fabric.”

They also note DataOps fits well with microservices architectures.

DataOps in practice

To make the most of DataOps, enterprises must evolve their data management strategies to deal with data at scale and in response to real-world events as they happen, according to Dunning and Friedman.

Because DataOps builds on DevOps, cross-functional teams that cut across “skill guilds” such as operations, software engineering, architecture and planning, product management, data analysis, data development, and data engineering are essential, and DataOps teams should be managed in ways that ensure increased collaboration and communication among developers, operations professionals, and data experts.

Data scientists may also be included as key members of DataOps teams, according to Dunning. “I think the most important thing to do here is to not stick with the more traditional Ivory Tower organization where data scientists live apart from dev teams,” he says. “The most important step you can take is to actually embed data scientists in a DevOps team. When they live in the same room, eat the same meals, hear the same complaints, they will naturally gain alignment.”

But Dunning also notes that data scientists may not need to be permanently embedded in a DataOps team.

“Typically, there’s a data scientist embedded in the team for a time,” Dunning says. “Their capabilities and sensibilities begin to rub off. Someone on the team then takes on the role of data engineer and kind of a low-budget data scientist. The actual data scientist embedded in the team then moves along. It’s a fluid situation.”

How to build a DataOps team

Most DevOps-based enterprises already have the nucleus of a DataOps team on hand. Once they have identified projects that need data-intensive development, they need only add someone with data training to the team. Often that person is a data engineer rather than a data scientist. DataKitchen suggests organizations seek out DataOps engineers who specialize in creating and implementing the processes that enable teamwork within data organizations. These individuals design the orchestrations that allow work to flow from development to production and ensure that hardware, software, data, and other resources are available on demand.

Many teams are built of individuals with overlapping skillsets, or individuals may take on multiple roles with a DataOps team, depending on expertise.

According to Michele Goetz, vice president and principal analyst at Forrester, some of the key areas of expertise on DataOps teams include:

- Databases
- Integration

- Data to process orchestration
- Data policy deployment
- Data and model integration
- Data security and privacy controls

Regardless of makeup, DataOps teams must share a common goal: the data-driven needs of the services they support.

DataOps roles

According to Goetz, DataOps team members include:

- Data specialists, who support the data landscape and development best practices
- Data engineers, who provide ad hoc and system support to BI, analytics, and business applications
- Principal data engineers, who are developers working on product and customer-facing deliverables

DataOps salaries

Here are some of the most popular job titles related to DataOps and the average salary for each position, according to data from PayScale:

- Analytics manager: \$72K-\$134K
- Associate data scientist: \$61K-\$101K
- Data analyst: \$46K-\$89K
- Data architect: \$81K-\$162K
- Data engineer: \$67K-\$134K
- Data scientist: \$70K-\$137K
- Director of analytics: \$85K-\$177K

DataOps tools

The following are some of the most popular DataOps tools:

- Census: An operational analytics platform specialized for reverse ETL, the process of synching data from a source of truth (like a data warehouse) to frontline systems like CRM, advertising platforms, etc.
- Databricks Lakehouse Platform: a data management platform that unifies data warehousing and AI use cases
- Datafold: A data quality platform for detecting and fixing data quality issues
- DataKitchen: A data observability and automation platform that orchestrates end-to-end multi-tool, multi-environment data pipelines
- Dbt: A data transformation tool for creating data pipelines

- Tengu: A DataOps orchestration platform for data and pipeline management

DETAILS

Subject:	Data management; Data analysis; Machine learning; Change management; Collaboration; Teams; Architecture; Software engineering; Automation; Engineers
Business indexing term:	Subject: Machine learning Architecture Automation; Industry: 54133 : Engineering Services
Classification:	54133: Engineering Services
Publication title:	CIO; Framingham
Publication year:	2022
Publication date:	Dec 22, 2022
Publisher:	Foundry
Place of publication:	Framingham
Country of publication:	United States
Publication subject:	Business And Economics--Computer Applications
ISSN:	08949301
Source type:	Trade Journal
Language of publication:	English
Document type:	News
Publication history :	
Milestone dates:	2022-12-22 (Published)
ProQuest document ID:	2756923897
Document URL:	https://www.proquest.com/trade-journals/what-is-dataops-collaborative-cross-functional/docview/2756923897/se-2?accountid=10382
Copyright:	Copyright Foundry Dec 22, 2022
Last updated:	2024-12-18
Database:	ProQuest One Academic,SciTech Premium Collection

LINKS

[Linking Service](#)

Database copyright © 2025 ProQuest LLC. All rights reserved.

[Terms and Conditions](#) [Contact ProQuest](#)