

# OPERATIONALIZATION AND ORCHESTRATION: THE KEYS TO DATA PROJECT SUCCESS



Data is vital to the success of every company. The amount of data available is staggering (and growing exponentially). But simply having the data isn't enough; companies must also utilize it correctly. Unfortunately, businesses struggle to get into production the data projects that turn all this data into insights. In fact, in 2018, Gartner® predicted in their report entitled "[Predicts 2019: Artificial Intelligence Core Technologies](#)" that through 2022 only 15 percent of cutting-edge data projects would make it into production. Looking at this from the other side—**85 percent of data projects will fail to produce results**. Pretty staggering, right? In its [Top Trends in Data and Analytics, 2022 report](#), Gartner points out that by 2024, organizations that lack a sustainable data and analytics operationalization framework will have their initiatives set back by up to two years.

As companies start to recognize that they need to build operationalization into their plans, the industry has begun to put a renewed focus on IT operations (ITOps). That has resulted in a plethora of variations around data (DataOps), machine learning (MLOps), artificial intelligence (AIOps), and analytics modeling (ModelOps). This boom has even spawned the term XOps, which some people in the industry are interpreting on a lighter note as, "we don't know what's coming next but it will involve Ops somehow, so we'll fill in the blank later." Ultimately, businesses know that they can have a project that works well in prototype in one location, but if it can't be scaled nationally or globally, the project has essentially failed.

Another reason data projects are so difficult to move to production is the sheer number of moving parts involved. Every data project has the same four basic stages, which are the building blocks of data pipelines: data ingestion from multiple sources, data storage, data processing, and insight delivery. Each of these stages involves a significant amount of technology and moving parts.

## Four Stages—Building Blocks of Data Pipelines

1. Data ingestion
2. Data storage
3. Data processing
4. Insight delivery

Looking at each stage, it quickly becomes apparent there are a lot of components across many application, data, and infrastructure technologies. Ingestion involves orchestrating data from traditional sources like enterprise resource planning (ERP) and customer relationship management (CRM) solutions, financial systems, and many other systems of record. This data is often combined with additional data from devices, sensors, social media, weblogs, and Internet of Things (IoT) sensors and devices, etc.

Storage and processing are also extremely complex. Where and how you store data depends significantly on persistence, the relative value of the data sets, the rate of refresh for your analytics models, and the speed at which you can move the data to processing. Processing has many of the same challenges: How much pure processing is needed? Is it constant or variable? Is it scheduled, event-driven or ad hoc? How do you minimize costs?

The last mile of the journey involves moving the data output to systems that provide the analytics. The insights layer is also complex and continues to shift. When the market adopts a new technology or capability, companies regularly adopt that shiny new thing. This constant innovation of new data technologies creates pressure and churn that can bring even the best operations team to its knees.

It's important to be nimble. You must be able to easily adopt new technologies. Remember—if a new data analytics service is not in production at scale, you are not getting any actionable insights, and as a consequence, the organization is not getting any value from it, whether it is generating revenue or driving efficiencies and optimization.

An obvious goal at the operational level is to run data pipelines in a highly automated fashion with little to no human intervention, and most importantly, have visibility into all aspects of the pipeline. However, almost every technology in the data pipeline comes with its own built-in automation, utilities, and tools that are often not designed to work with each other, which makes them difficult to stitch together for end-to-end automation and orchestration. This has led to a rise in application and data workflow orchestration platforms that can operate with speed and scale in production and abstract underlying automation utilities.

# Vendor Landscape for DataOps

Representative vendors only (in some cases, product names mentioned)

Your decision flow should be:

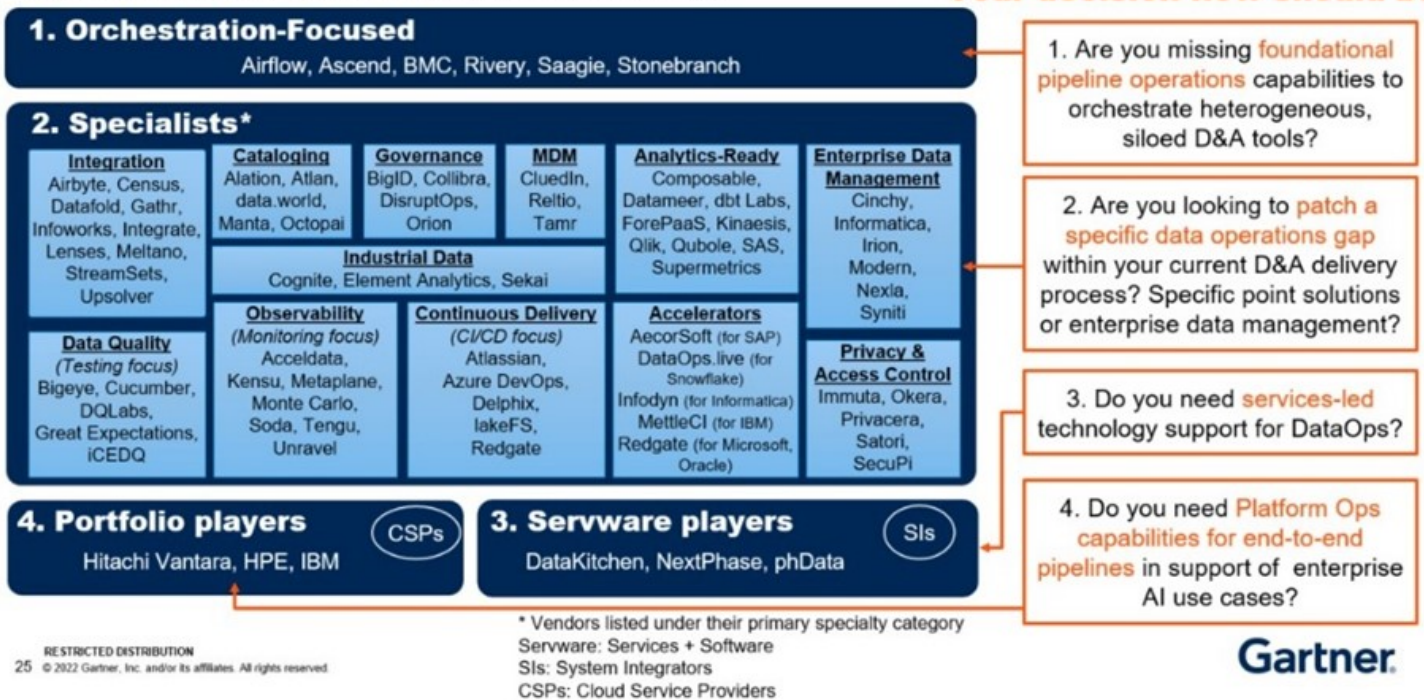


Figure 1. Gartner Data and Analytics Essentials: DataOps by Robert Thanaraj

Control-M from BMC is an application and data workflow orchestration and automation platform that serves as the abstraction layer to simplify the complex data pipeline. It enables end-to-end visibility and predictive service level agreements (SLAs) across any data technology or infrastructure.

Control-M delivers data-driven insights in production at scale, and easily integrates new technology innovations in the most complex data pipelines with ease.

The Control-M platform has a range of capabilities to help you automate and orchestrate your application and data workflows, such as:

- The [Control-M Automation API](#), which promotes collaboration between Dev and Ops by allowing developers to embed production-ready workflow automation while applications are being developed.
- [Out-of-the-box support for cloud resources](#) including Amazon Web Services (AWS) Lambda and Azure Logic Apps, Functions, and Batch to help you leverage the flexibility and scalability of your cloud ecosystems.
- [Integrated file transfers](#) with all your applications that allow you to move internal and external file transfers to a central interface to improve visibility and control.
- [Self-Service](#) features that allows employees across the business to access the jobs data relevant to them.
- [Application Integrator](#), which supports the creation of custom job types and deploys them in your Control-M environment quickly and easily.
- [Conversion tools](#) that simplify conversion from third-party schedulers.

Data projects will continue to grow in importance. Finding the best way to successfully operationalize data workflows as a key part of your overall project plan and execution is vital to the success of your business. An [application and data workflow orchestration](#) platform should be a foundational step in your DataOps journey.

To learn more about how Control-M can help you find DataOps success, visit our [website](#).

*GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.*