OPERATIONALIZATION AND ORCHESTRATION: THE KEYS TO DATA PROJECT SUCCESS

To operationalize data projects, organizations need automated, orchestrated, end-to-end visibility across every stage of the data pipeline. Without a structured operationalization framework, even the most sophisticated data initiatives stall before reaching production. A unified workflow orchestration platform that abstracts the complexity of disparate data tools is the critical bridge between a working prototype and enterprise-scale deployment.

Data is abundant—and growing exponentially. But simply having the data isn't enough. Businesses consistently struggle to move data projects from development into production, and the cost of that gap is significant. In 2018, Gartner® predicted in their report entitled "Predicts 2019: Artificial Intelligence Core Technologies" that only 15 percent of cutting-edge data projects would make it into production by 2022—meaning 85 percent would fail to produce results. Separately, in its Top Trends in Data and Analytics, 2022 report, Gartner warned that organizations without a sustainable data and analytics operationalization framework risk seeing their initiatives set back by up to two years.

Why do data projects struggle to reach production?

The core challenge is scale. A project can work well in prototype in one location, but if it can't be scaled nationally or globally, it has essentially failed.

As companies recognize the need to build operationalization into their plans, the industry has refocused on IT operations (ITOps)—generating a range of discipline-specific Ops frameworks: DataOps for data, MLOps for machine learning, AIOps for artificial intelligence,

and ModelOps for analytics modeling. This proliferation has even produced the catch-all term XOps—a placeholder, as some in the industry put it, for "we don't know what's coming next but it will involve Ops somehow." The problem isn't ambition. It's complexity.

What are the four stages of a data pipeline?

Every data project shares the same four foundational stages, which form the building blocks of data pipelines:

1.Data ingestion

Orchestrating data from traditional sources such as enterprise resource planning (ERP) and customer relationship management (CRM) systems, financial platforms, and other systems of record—combined with data from devices, sensors, social media, weblogs, and IoT sensors and devices.

2. Data storage

Where and how data is stored depends significantly on persistence, the relative value of data sets, the rate of refresh for analytics models, and the speed at which data can move to processing.

3. Data processing

Processing requirements vary widely: How much compute is needed? Is it constant or variable? Is workload scheduled, event-driven, or ad hoc? How are costs minimized?

4. Insight delivery

The last mile—moving data output to analytics systems. The insights layer is complex and shifts constantly as markets adopt new technologies. A new data analytics service that isn't in production at scale delivers no actionable insights and no business value, whether measured in revenue generation or operational efficiency.

Why is end-to-end orchestration essential for data pipelines?

The operational goal is to run data pipelines in a highly automated fashion with minimal human intervention and full visibility across every component. But almost every technology in the data pipeline comes with its own built-in automation utilities and tools—tools that are often not designed to work with each other. Stitching these together for end-to-end automation and orchestration is where teams hit the wall.

This challenge has driven the rise of application and data workflow orchestration platforms: solutions that operate with speed and scale in production and abstract the underlying automation utilities so teams can focus on outcomes rather than infrastructure.

How does Control-M help operationalize data projects?

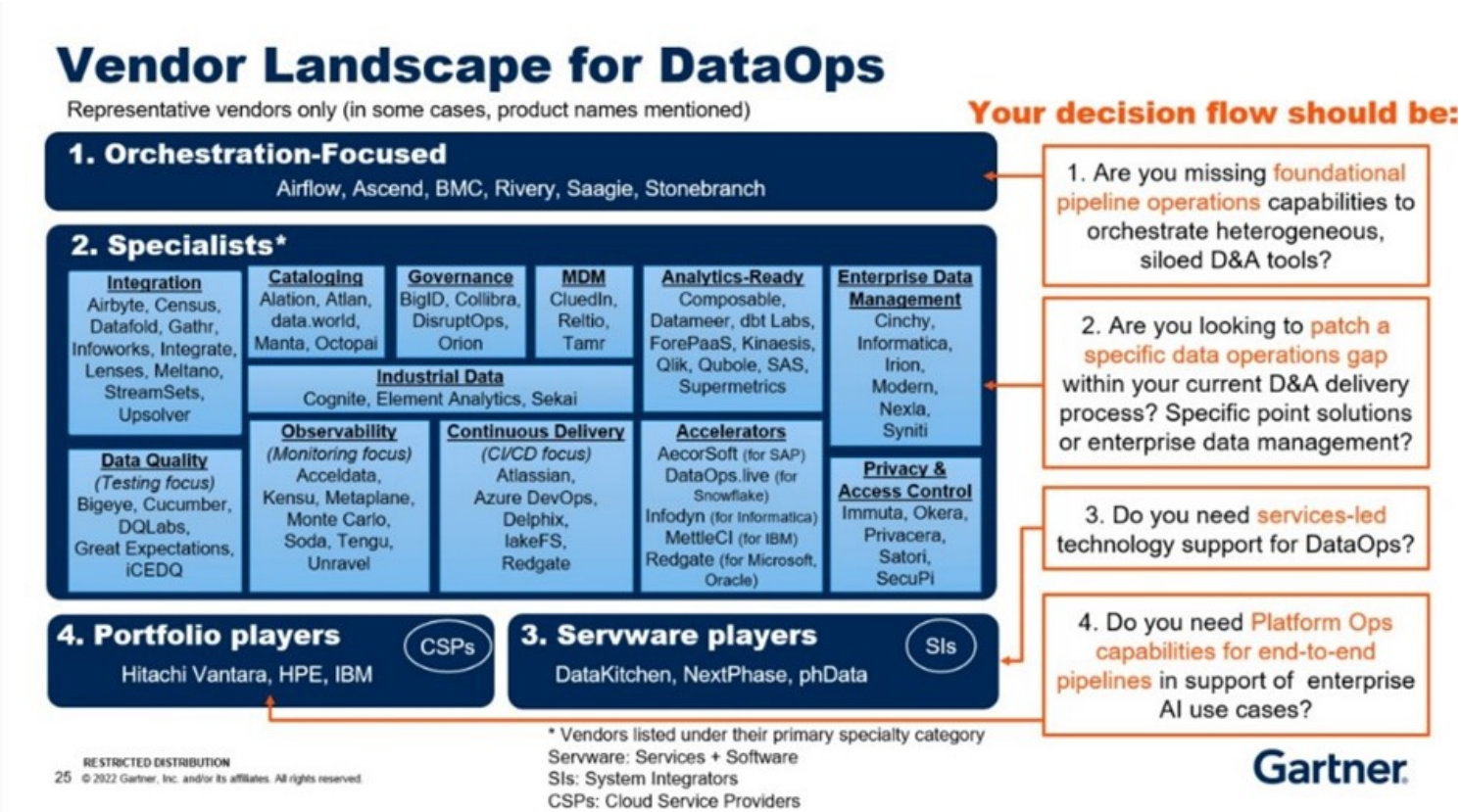

Figure 1. Gartner Data and Analytics Essentials: DataOps by Robert Thanaraj

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. Control-M enables end-to-end visibility and predictive service level agreements (SLAs) across any data technology or infrastructure, delivers data-driven insights in production at scale, and integrates new technology innovations into even the most complex data pipelines.

The Control-M platform offers a range of capabilities to automate and orchestrate application and data workflows:

- 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 strategic importance. Successfully operationalizing data workflows as a core part of project planning and execution is essential to business outcomes. An application and data workflow orchestration platform should be a foundational step in every DataOps journey.

To learn more about how Control-M can help you find DataOps success, visit our website.

Frequently asked questions

What does it mean to operationalize a data project?

To operationalize a data project means to move it from prototype or development into reliable, repeatable production at scale. Operationalization requires automating the core pipeline stages—ingestion, storage, processing, and insight delivery—so that data projects consistently produce actionable results without manual intervention.

Why do most data projects fail to reach production?

Most data projects fail to reach production because of complexity at scale. Each pipeline stage relies on different tools and technologies that don't natively integrate, making end-to-end automation difficult to achieve. Without an orchestration layer to unify these components, teams cannot maintain the visibility and control needed to run pipelines reliably in production.

What is the difference between data automation and data orchestration?

Data automation refers to executing individual tasks or processes without human intervention. Data orchestration coordinates multiple automated tasks across tools, systems, and environments to execute a complete data pipeline as a single managed workflow. Orchestration is what enables automation to work at enterprise scale.

What role does DataOps play in operationalizing data projects?

DataOps applies DevOps principles—continuous integration, automation, and collaboration—to data pipeline development and operations. A DataOps framework helps organizations build operationalization into data projects from the start, reducing the time from prototype to production and improving the reliability of data-driven outcomes.

What should I look for in a workflow orchestration platform?

A workflow orchestration platform for data projects should provide end-to-end pipeline visibility, support for heterogeneous data technologies and cloud environments, event-driven and scheduled workload capabilities, integrated file transfer management, self-service access for business users, and APIs that enable Dev and Ops collaboration during development.

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.

The views and opinions expressed in this post are those of the author and do not necessarily reflect the official position of BMC.