



WHITEPAPER

# You Don't Need More Data Engineers. You Need DataDevOps.

The missing link between your DevOps culture and your data stack.

BY LAKSHMI VENKATASWAMY, FOUNDER & CTO, EASL  
AND JOHN P. DERHAM, FOUNDER & CEO, EASL

The term "DataDevOps" is a new concept that is quickly emerging in forward-looking technology circles. It combines three main disciplines in enterprise technology:



If the term sounds vaguely familiar to you, it's probably because you've heard the term "DevOps" mentioned in startup tech companies, cloud computing communities, or even Agile teams. DevOps is a set of technology practices that combines software development (Dev) and IT operations (Ops) to improve collaboration, automate processes, and enable faster and more reliable software delivery and deployment.

---

## DevOps Solved One Set of Problems, but Data Has Its Own

DevOps was born out of frustration in the early 2000s, when traditional programming processes followed a strict and linear pattern. The developers would code their parts and then toss them over the wall to the next team, who would audit and implement them in IT operations. This siloed process consistently led to slow implementations, production failures, and the inevitable finger-pointing. In

2008-09, a Belgian software engineer named Patrick Debois sought a solution to this problem by organizing a conference called "DevOpsDays." The goal of the conference was to create new processes of collaboration, automation, and accountability between the development (Dev) and operations (Ops) teams in technology organizations, thereby coining the term "DevOps."

The early benefits of this approach were evident, and the principles were quickly adopted by the most innovative and fastest-moving technology organizations over the next decade.

Infrastructure-as-Code, continuous integration, and automated testing became standard concepts that gave rise to the emergence of microservices and container environments, shifting infrastructure innovations away from legacy computing cores and into containerization and orchestration.

## The Cultural Shift That Made DevOps Work

One of the most significant outcomes of DevOps was the cultural shift that happened as the concept moved from grassroots to standard operating procedures in the fastest-moving organizations. High error rates and the blame game were replaced by collaboration and feedback loops. This enabled organizations to redefine success metrics in terms of new and previously unattainable measures, such as deployment frequency, lead time reduction, mean time to repair (MTTR), and time to value (TTV). The organizations that adopted a DevOps framework quickly had efficient and happy technology teams.

The success of DevOps practices only makes it more obvious that large parts of data and software development still have a long way to go to become lean and mean technology machines.

So, when we think about the term “DataDevOps,” we can associate it with the application of DevOps principles specifically to data pipelines, data

platforms, and AI/ML workflows. It aims to break down silos between the data engineers, analysts, data scientists, and operations teams—just like traditional DevOps did for software developers and IT. Essentially, DataDevOps helps us unify data-centric workflows and remove the barriers between data engineers and software teams.

Sources like Gartner and McKinsey have reported that nearly 40% of “technical debt” in enterprise environments stems from poor data quality, resulting from failing and inadequate data integration systems that remain in place long after their deficiencies are known. The reality is that the current roles of data engineers, referred to as “DataOps,” are not in line with the rest of the DevOps teams. In reality, the disconnect between the two teams is eerily similar to the siloed challenges that made DevOps necessary over 15 years ago.

“Data” is no longer an adjacency to core processing—it is the core of your technology and the core of your business. DevOps processes can create efficient and effective environments for processing, but poorly implemented data pipelines can undermine all of those benefits. The current levels of frustration in this ecosystem are incredibly high and are likely to increase as the adoption of AI drives business needs at an accelerated pace. Now is the time to think differently about this problem and plan for the future. Legacy

tools and approaches will not solve the issues you are facing—or are likely to face—without an innovative approach to DataDevOps.

## Why DataDevOps Is More Than Just Merging Teams

DataDevOps is not merely about integrating your DevOps teams with your DataOps teams. Your DataOps team members likely lack system access and authority to engage in operational platforms and business rules. In contrast, the DevOps team members may not have access to production databases. The DataOps teams are often constrained by legacy tools and “data connectors” so outdated and limited that starting from scratch feels like the only option for each new data source.

Whether you decide to build this capability completely in-house (which we strongly advise against) or integrate with an enterprise third-party platform like an iPaaS or ETL, you should focus on the following capabilities to drive your strategic goals and build a true DataDevOps foundation:

### Think “Workflows,” not “Connectors”

Legacy systems that have historically helped enterprise organizations connect data sources—often referred to as ETLs (extract/transform/load)—were developed 15+ years ago to

perform standard data movement tasks. The world has changed, while these platforms have not. Generic data connectors don’t work with the complexities of today’s data movement needs because these needs are never generic. CI/CD workflow (continuous integration/continuous delivery) is a standard DevOps concept that should be applied to data workflows. Workflow-based infrastructure enables a DevOps mindset, empowering the automation of tasks involved in building, testing, and deploying code changes for fast-changing data pipelines.

Generic data connectors don’t work for the complexities of today’s data challenges because these challenges are not generic.

### Require Strict Environmental Controls for Data and Pipelines

This may seem like an obvious request to your technology teams, but reality often falls well short of expectations. Simple rules, such as version control for changes to data movement coding, are easy to achieve with a DataDevOps practice. Storing all raw data feeds as immutable records before transforms is another easy-to-achieve, high-value feature. These two simple structures enable the auditability of processes and raw data from multiple sources, allowing quick changes and error



tracking. They're foundational to any solid DevOps platform, and just as essential in data environments.

### **Monitor All Steps in Data Movement**

The basic tenet of a DataDevOps solution is end-to-end observability. The transparency of all data functions, including adherence to expected schemas, data quality, pipeline outages, and data drift, as well as record-level and variable-level specification (e.g., a date field should be a date field), is often lacking in many data engineering solutions. Data problems usually arise when a user identifies them in an operational or analytic endeavor where the data does not conform. The best way to address a problem is to identify it in the data before it becomes a significant issue. Auditing data pipelines and their stages becomes a straightforward exercise when monitoring systems are correctly set up.

### **Enforce Collaboration Across All Data Users and Sources**

The mindset of specific "keepers of data" should be retired. This specialization is often the root cause of poorly documented data dictionaries and unclear communication around bespoke data source use cases. The user communities of data can typically be divided into two groups: those who use the data to "operate the business" and those who "analyze the business." While these two groups may have different day-to-day goals,

they are inextricably linked and need to share. Make all data shareable (with privileges) across functions, use the data to experiment with new policies and procedures, and ensure that all results are shared across teams. You'd be surprised what others don't know, and how vital that gap is to delivering an efficient DataDevOps function.

### **Apply IaC (Infrastructure-as-Code) to Data Deployment**

This mindset enables your team to focus less on building data environments and more on automating and managing them through business logic. This automation eliminates the need for developers to manually provision data handling servers, operating systems, and databases, as well as manage the provisioning of resources through awesome platforms like Databricks, Snowflake, Azure, and AWS. This means the configuration files can contain infrastructure specifications, making it easier to edit and deploy various configurations, which results in greater efficiency, lower capital expenditures, and significantly reduced operational costs.

### **Prioritize Security and Governance**

As AI/ML models expand into all corners of our lives, the standard for reviewing data will increase at a rapid rate. Your data can be easily protected at rest by encrypting it, hashing PII variables, and backing it up as required by prevailing policies and laws. In the ever-changing

world of DataDevOps, you will need to limit access to much of your data based on the roles of users within the community. Your data engineers and ML engineers will need different levels of access than your platform engineers and analytics engineers. Role-based access to every content library—and sometimes even down to the field

level—is an excellent policy that can be managed relatively easily within the context of DataDevOps practices. These types of roles and rules are always considered the cornerstone of good data governance. Remember that these provisions are designed to expedite things, not slow them down.

---

## **DataDevOps Is Inevitable. The Question Is When.**

The world is changing rapidly and will continue to evolve at rates that we cannot yet fathom. Just as DevOps became essential nearly 25 years ago, DataDevOps will emerge from a fringe theory and into a core process for all organizations. Its core value is simple: helping organizations integrate data processes into everyday business functions with the same agility, discipline, and speed that DevOps brought to software.

To learn more about the history of DevOps and its relationship to the emerging DataDevOps revolution, please refer to the [DevOps Handbook](#).



## **Built on Experience, Designed for Tomorrow**

EASL was conceived by a team with 35+ years of experience moving data at massive scale. Our platform integrates this deep expertise with cutting-edge technologies to solve the acute challenges of scaling data implementation and processing capabilities that face any high-growth company. Our SOC2 Type I & II certified platform operates with zero-record-loss according to the highest compliance, audit and security standards.

**Learn how EASL can help you get your data right.**

Visit: [easltech.com](https://easltech.com)

### **ABOUT LAKSHMI VENKATASWAMY**

Lakshmi Venkataswamy has spent three decades solving the kinds of technical problems most people walk away from. As EASL's Founder and CTO, he brings deep engineering discipline to the challenge of scaling data infrastructure in real-world environments, where stability, performance, and cost aren't tradeoffs, they're requirements. His background spans enterprise IT leadership in financial services and media, startup engineering in global markets, and the hands-on design of systems built to last. He holds a B.E. in Electronics & Communications from NIT Trichy and an MBA in Finance from Drexel University.

Email: [lakshmi@easltech.com](mailto:lakshmi@easltech.com)

### **ABOUT JOHN P. DERHAM**

John Derham is a co-founder of EASL Tech, Inc. and a longtime architect of advanced data and analytics platforms. With deep experience spanning media, financial services, and e-commerce, he's known for building high-performing teams, bridging the gap between technical and executive leadership, and developing technology that delivers measurable impact. He holds a degree in Financial Analytics from the Villanova School of Business.

Email: [jderham@easltech.com](mailto:jderham@easltech.com)