

Tackling AI and Machine Learning

Simplify Orchestration and Acceleration for MLOps with Domino Data Lab and WEKA Data Platform

Summary

- Find, reproduce, and reuse work to maximize productivity and compound knowledge.
- Provide tools, data sources, compute resources, etc.
- Support hybrid cloud deployments
- Drive faster time to production and even more immediate outcomes

Use Cases

- Machine Learning & Deep Learning
- Genomics
- High Tech
- Manufacturing

MLOps is a comprehensive approach to managing the deployment and performance of machine learning models in production. It addresses the challenge of bridging the gap between data scientists focusing on model building and operations teams focusing on deployment and maintenance. Enterprise MLOps aims to enhance the efficiency and quality of model deployment while reducing risk and fostering collaboration among teams. This requires a robust and scalable approach that encompasses the entire data science lifecycle and meets the needs of various teams now and in the future. Enterprise MLOps incorporates practices like continuous integration and deployment, automated testing and monitoring, and version control for machine learning models.

As your organization expands the scope and scale of its machine learning efforts, success hinges on the ability to deploy the end-to-end enterprise MLOps stack quickly—both on-premises and in the cloud—while ensuring that you deliver:

- Optimized application and data performance and management
- Access to self-service tools and powerful compute and storage infrastructure
- Publish/deploy work faster and easier to drive business impact

But the reality of today's MLOps can be a long way from this ideal. When companies scale their MLOps operations enterprise-wide, they often struggle with the following challenges:

- **Complexity:** Machine learning models can be complex, and scaling them to handle large amounts of data and perform well in production environments can be difficult.
- **Data Quality:** Maintaining the quality of data used to train and test models is critical for MLOps; it can be challenging to ensure that data is consistent, accurate, and up-to-date at scale.
- **Model Versioning:** Keeping track of different models and their respective performances can be difficult, especially as the number of models deployed increases.
- **Collaboration:** Collaboration between data scientists and operations teams can be complex, as they may have different priorities and use different tools and technologies.
- **Monitoring and Debugging:** Monitoring the performance of machine learning models in production environments can be complex, and debugging issues that arise can be time-consuming and complex.
- **Security and Compliance:** Ensuring that machine learning models are deployed and managed securely, and compliant is crucial, but it can be challenging to achieve at scale.
- **Lack of Automation:** MLOps requires a lot of manual processes, which can be time-consuming and lead to errors. Automating these processes is essential for scaling MLOps.
- **Limited Resources:** Organizations may need more resources to devote to MLOps, making scaling difficult.

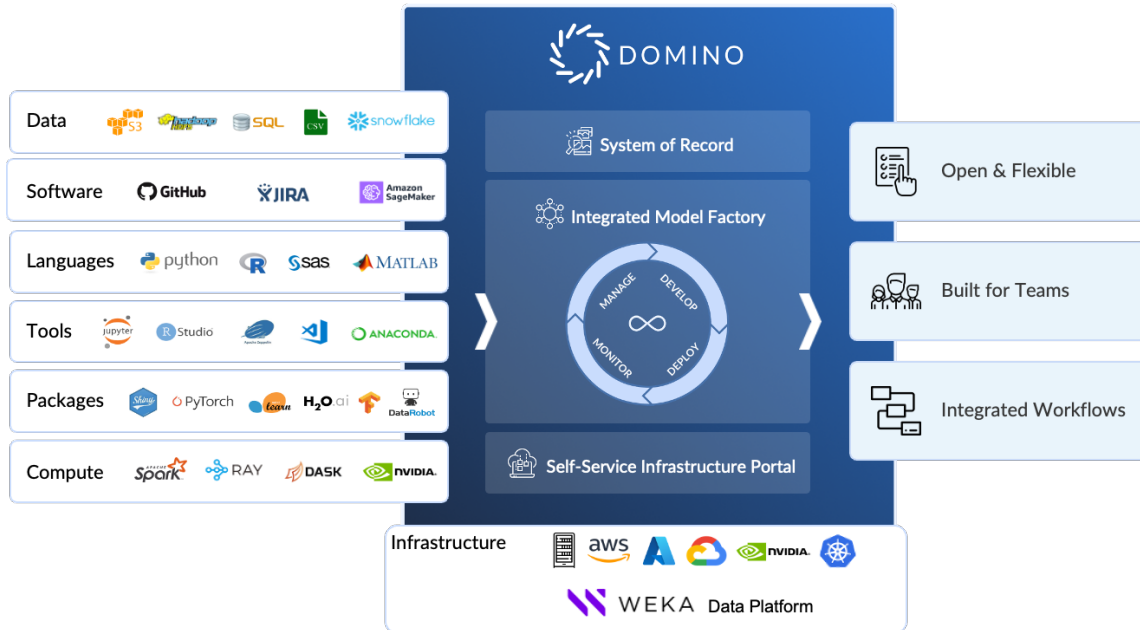
The Enterprise MLOps Solution

WEKA has joined forces with Domino Data Lab to help your data scientists and organization overall address these challenges, increasing the success of your machine learning and deep learning efforts. Research teams can gain on-demand access to resources for their entire AI workflow, from building the model to training to inference to maintenance.

Domino Data Lab Advantages

Domino Data Lab is a platform for data science teams that provides the tools for the entire MLOps lifecycle while facilitating collaboration, version control, and reproducibility.

Domino Data Lab allows data scientists to easily share and collaborate on code, data, and models. The platform includes version control, easy rollbacks and reproducibility, and a built-in environment for running code and experiments. This makes it easy for data scientists to work together on projects while ensuring their work is easily reproducible.



Domino Data Lab's new hybrid cloud offering Nexus provides a single pane for data science across all an enterprise's regions and environments—whether on-prem, in the cloud, or in multi-cloud settings.

Core features include:

- **One-click access.** Nexus lets you launch new jobs and workspaces on-prem or on different cloud platforms and allows you to select hardware tiers based on your cost and performance requirements.
- **Unified data access control.** With Nexus, you can designate which environments can access specific data sets. It thus allows you to restrict access to data by region, helping you enforce compliance with data localization and sovereignty regulations.

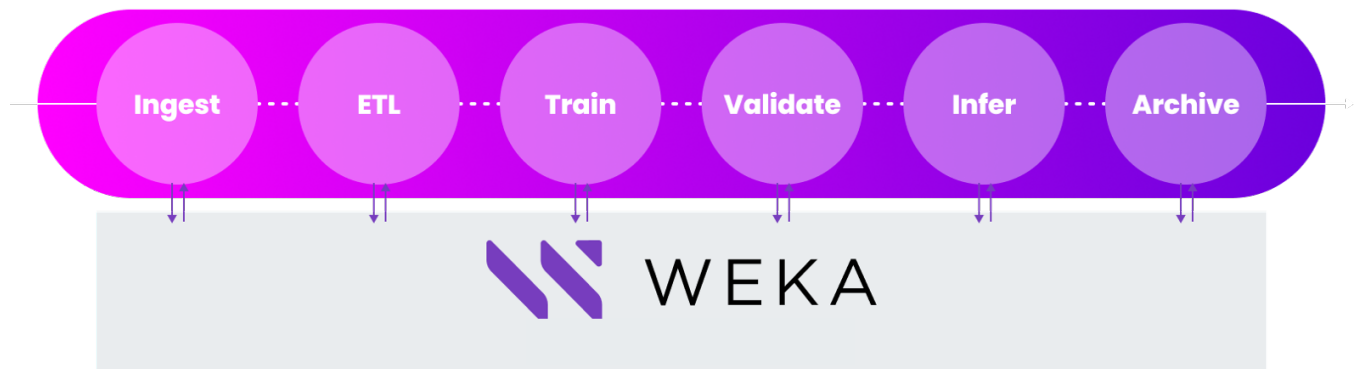
Benefits include:

- **Data Science productivity:** Domino Data Lab supports the data science market's broadest ecosystem of infrastructure and tools. Data scientists can use their preferred tools, languages, and packages without worrying about how to configure them.
- **Collaboration and Re-use:** Capture and share collective wisdom from all corners of your business to turbocharge data scientist productivity. You will always retain valuable IP when data scientists change jobs, and new team members can become productive faster.
- **Governance and Standardization:** Centralize and orchestrate all data science work on one platform with enterprise-grade security, governance, compliance, and policies to scale safely and universally across your organization.
- **Model Velocity:** Accelerate innovation by removing friction from every aspect of the data science lifecycle. End-to-end workflows create common patterns and practices used by everyone regardless of their role or favored data science tool so you can get more models into production faster.

WEKA Advantages

WEKA has built a software-defined Data Platform that leverages cutting-edge cloud, compute, storage, and fast networking technologies to unleash the value of your data.

WEKA Data Platform delivers consistent lightning-fast access to data at terabytes to exabytes scale when needed across the AI workflow. Our patented architecture starts small and scales to extremely large deployments without compromising performance. The WEKA Data Platform eliminates the need for multiple storage options and data copies across MLOps workflows, reducing operational complexity, enhancing pipeline efficiency, and increasing GPU utilization.



With WEKA, a single data platform supports all popular data access methods, including the POSIX-compliant file system, NFS, SMB, S3, CSI for Kubernetes, and GPU Direct Storage (for direct data movement between GPUs and storage). In addition, because the platform is software-defined, it enables both on-prem, private clouds and cloud deployment across popular clouds. WEKA also makes moving data across different clouds a breeze allowing organizations to take advantage of multiple or hybrid clouds like our partner Domino Data Lab.

Additional benefits include:

- **Move and back up data easily.** Advanced data management capabilities enable data movement quickly and efficiently between different clouds and simplify backups to local or remote cloud regions.
- **Tier automatically.** WEKA can automatically tier cold data to low-cost object storage, on-prem, or cloud for better economics. All data remains in the namespace, and metadata stays on the flash tier for fast access.
- **Ensure security.** The WEKA data platform was architected to ensure the security of your data with advanced authentication, in-flight and at-rest encryption, and flexible key management.

Read more in the WEKA Data Platform [architecture white paper](#).

What Does the Domino Data Lab and WEKA Solution Mean for Your Organization

Together, Domino Data Lab and WEKA provide a complete solution for data science teams to manage and accelerate their MLOps workflows and collaborate on projects while ensuring that data is accurate, protected, and easily accessible.

A hybrid cloud approach has many benefits and recognizes the reality of the on-prem systems and regulations companies face. However, to take full advantage of hybrid, companies must move on from the manual processes and disconnected platforms they have today.

Domino Data Lab and WEKA can make it available for projects no matter where your data lives. IT can consolidate siloed stacks onto a single governed data science platform running on-prem or in the cloud while supporting distributed compute networks such as Spark, Ray, and Dask. With our joint solution, you capture and track all project artifacts, including code, package versions, and parameters, to establish full visibility, repeatability, and reproducibility at any time across the end-to-end data science lifecycle.

Data science teams can collaborate better and be more productive with a true hybrid platform that enables them to access data, compute resources, and code in every environment where the company operates in a secure, governed fashion. The alternative is spiraling cost, wasted effort, suboptimal models, and higher risk.



Domino Data Lab and WEKA allow organizations to create a hybrid- and multi-cloud MLOps platform

Get Started Today

If you are an organization investing in artificial intelligence, machine learning, and deep learning initiatives, Domino Data Lab and WEKA can help simplify your journey with excellent ROI, better model quality, and an excellent time to production. We are excited to offer a complimentary technical consultation looking at your current and future needs, including a proposal for a proof of content. Email us at DDL-solutions@weka.io or contact your authorized Domino Data Lab and WEKA representatives to learn more about this offer and more information.



weka.io

844.392.0665

