BUYERS GUIDE FOR EVALUATING MLOPS SOLUTIONS

Table of Content

Overview

There's never been a better time to be involved in machine learning. By all accounts, demand for sophisticated AI/ML solutions continues to increase rapidly. In fact, Gartner estimates that "the AI and data science platform market is due to grow to over $10 billion by 2025 at a 21.6% compounded annual growth rate. This growth in the market mirrors the investments made by organizations in data science and ML initiatives, which are largely turning from strategy to execution." [1]

But we know that rapid adoption of ML & AI is not without its challenges, some of which include:

Data volume and quality

Machine learning algorithms require large amounts of data to learn from and train on, but if the data is of poor quality, your model will be inaccurate and ineffective.

Data privacy and security

The data your ML model uses may contain sensitive information, so you need to ensure that there are appropriate security measures in place to protect this data from unauthorized access.

Lack of expertise

Because machine learning is such a highly specialized field, you may struggle to find staff with the necessary expertise to develop and implement machine learning models.

Integration with existing systems

Integrating ML models with existing systems can be challenging, particularly if the systems were not designed with machine learning in mind.

Cost

Developing and implementing ML models can be expensive, particularly if your organization needs to invest in specialized hardware or software as well as the cost of compute, especially GPUs.

Ethical considerations

Machine learning can potentially perpetuate bias or discrimination if not developed and implemented ethically.

ML readiness

Adopting machine learning may require a significant shift in cross-organizational culture and processes, which employees might resist.

The ability to operationalize at scale

The successful implementation of machine learning at scale cuts across multiple siloed teams and Business Units (especially in large Mid-Market and Enterprise organizations), and requires team collaboration and bridging technical and business communications.

[1] "Market Guide for DSML Engineering Platforms," Gartner, May 2, 2022.

In addition, according to our recent survey, "MLOps in 2023: What Does the Future Hold," which polled 200 U.S.-based machine learning decision makers, 86% of respondents struggled with using ML to generate business and commercial value, with 71% acknowledging that their company was failing to generate value as a result of challenges in implementing ML at scale. Leading analysts have found "Top challenges that customers face with implementing machine learning initiatives in production include lack of expertise, cost, and lack of automation." [2]

The good news is that Machine Learning Operations (MLOps) is an excellent solution to these challenges, and is rapidly transforming the way businesses develop, deploy, monitor, and maintain their machine learning models. MLOps is not an ML tech stack – it goes beyond the sheer technology or tools used to develop ML applications and products. Rather, it is a set of best practices and tools that enable you to automate the entire machine learning lifecycle, from data preparation and ingestion to model development, deployment and maintenance. By leveraging MLOps, you can reduce the time and cost of deploying machine learning models while improving the accuracy and reliability of your ML apps and products at scale.

Here are five additional ways in which MLOps can help you:

1

Improved collaboration:

MLOps practices promote collaboration between data scientists, engineers, and other stakeholders throughout the ML model development and deployment process, ensuring that everyone is working towards a shared goal and minimizing silos of information.

2

Enhanced project management:

MLOps provides a framework for project management that helps teams to better plan, execute, and track ML projects, reducing the risk of delays and other operational issues.

3

Streamlined operationalization:

MLOps helps to automate and streamline the process of operationalizing ML models, reducing the time and effort required to move a model from development to deployment.

4

Better monitoring and maintenance:

MLOps provides tools and best practices for monitoring and maintaining ML models in production, allowing you to quickly identify and resolve issues that may affect performance or model accuracy (e.g., model drift etc.)

5

Improved scalability:

MLOps enables you to scale your ML models and operations to meet the needs of a growing business, without compromising on performance or security.

[2] "IDC MarketScape: Worldwide Machine Learning Operations Platforms 2022 Vendor Assessment," https://www.idc.com/getdoc.jsp?containerId=US48325822

MLOps is seen as a critical approach in addressing the operational challenges that you might be facing when it comes to operationalizing ML at scale, and we expect its adoption will continue to grow as more businesses aim to capitalize on the value of ML.

However, with so many MLOps tools and platforms available, choosing the right solution can be a daunting task. That's where this Buyers Guide for Evaluating MLOps Solutions comes in. This guide is designed to provide you with a comprehensive overview of the key considerations to keep in mind when choosing an MLOps platform or tool for your business.

In this guide, we'll cover a range of topics, including:

- What is MLOps
- The Build-or-Buy Decision
- The Difference Between Point Solutions and End-to-End Platforms
- Open Source versus Managed Platforms
- How To Evaluate an MLOps Solution
    - » Features and Functionality to Look For
    - » Deployment Options
    - » Considerations for Service & Support
    - » The Importance of Scalability, Security, and Governance
- Introducing ClearML's Unified, Open Source Platform for Continuous ML
- How ClearML Compares to Other Solutions
- Checklist for Evaluating MLOps Solutions
- Summary & Next Steps

So whether you're just getting started with MLOps or looking to upgrade your existing ML stack and infrastructure, this guide will help you make an informed decision about the right MLOps solution for your business. Let's dive in!

What is MLOps

MLOps is the practice of integrating machine learning models into an organization’s software delivery process. MLOps is especially important in deep learning, a subset of machine learning that involves the use of artificial neural networks (ANNs), which mimic the working of the brain to process and analyze data. ANNs are extremely good at extracting vital information by filtering it from the noise.

One of the critical advantages of deep learning is its ability to learn patterns and representations from large amounts of data without requiring explicit instructions or labels. This allows deep learning systems to find relationships in the data that may not be immediately apparent to humans. This has led to the development of many successful deep-learning applications, including **image** and **speech recognition**, **natural language processing, generative modeling**, and **machine translation**. That’s where the automation and orchestration inherent in MLOps are particularly valuable.

MLOps uses a set of practices to ensure ML model reliability. MLOps involves collaboration between data scientists and programmers, who build and train the models, and IT professionals, who handle infrastructure and deployment of the models. MLOps workflows cover the entire machine learning model lifecycle, from producing ML models to continued maintenance after deployment. In other words, MLOps bring together the development and operations of machine learning (and DL) models.

Why is MLOps Important?

Adopting an MLOps approach can positively benefit machine learning model creation by making the process of model development more streamlined. When used effectively, MLOps helps you develop, deploy, and maintain machine learning models quickly and easily. Without MLOps, the process of developing and deploying machine learning models at scale might be slow and error-prone, which could lead to delays in shipping the model, resulting in increased costs and delayed time to value.

Here are just a few additional ways that MLOps can help the ML-driven organization:

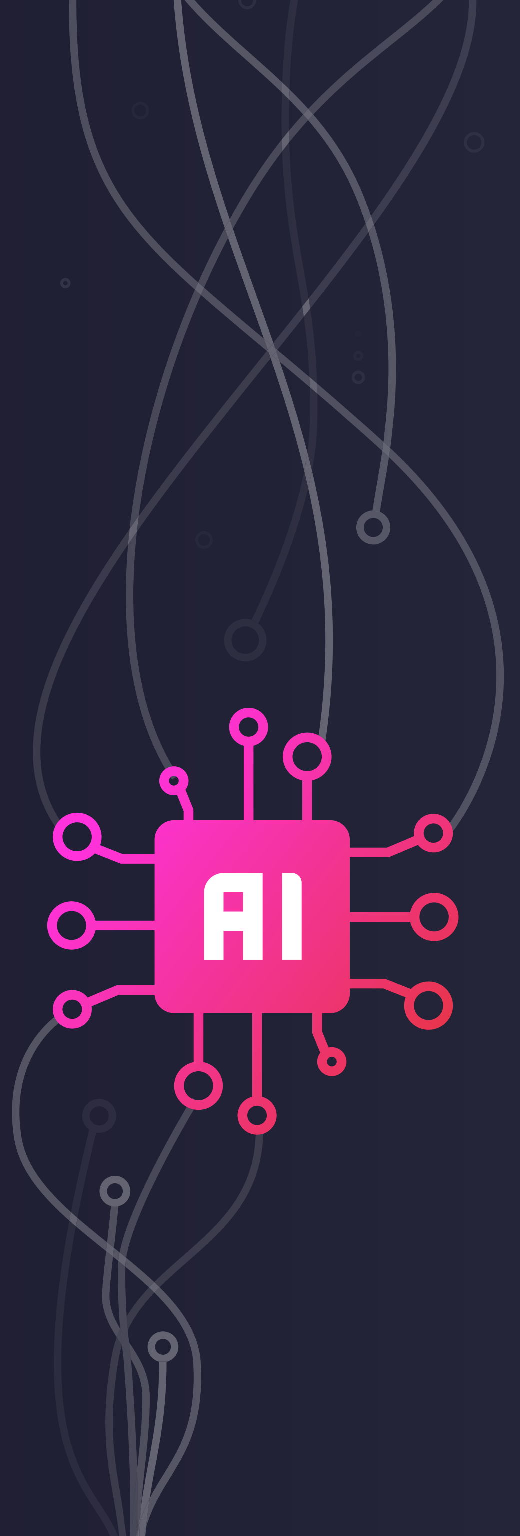
- **MLOps can help ensure the reproducibility and reliability of machine learning models:**
  MLOps practices such as version control and continuous integration can help ensure ML models are developed in a reproducible way and that they are reliable in production.

- **MLOps increases overall production productivity:**
  Low-code environments that provide access to organized, focused data sets reduces the amount of data scientists needed to ship an ML model and allows them to move faster, which reduces time and costs.

- **Improving collaboration between data scientists, ML practitioners, IT professionals and business owners:**
  MLOps helps bridge gaps between these groups by providing a common set of practices, processes and/or tools that they can use to work together more effectively.

- **Improving the monitoring and management of ML models:**
  MLOps practices, such as monitoring and alerts, can help you more effectively monitor the performance of your machine learning models, identifying and resolving any issues that arise.

- **Increasing sales for new applications or products by producing ML models:**
  Machine learning models can be used to:
  - a) analyze historical data and forecast future trends
  - b) provide personalized recommendations to customers
  - c) segment customers for targeted marketing campaigns
  - d) optimize prices based on consumer behavior and trends
  - e) analyze transaction data to identify fraud and reduce risk
  - f) forecast sales to help influence inventory, resource allocation, and marketing strategies

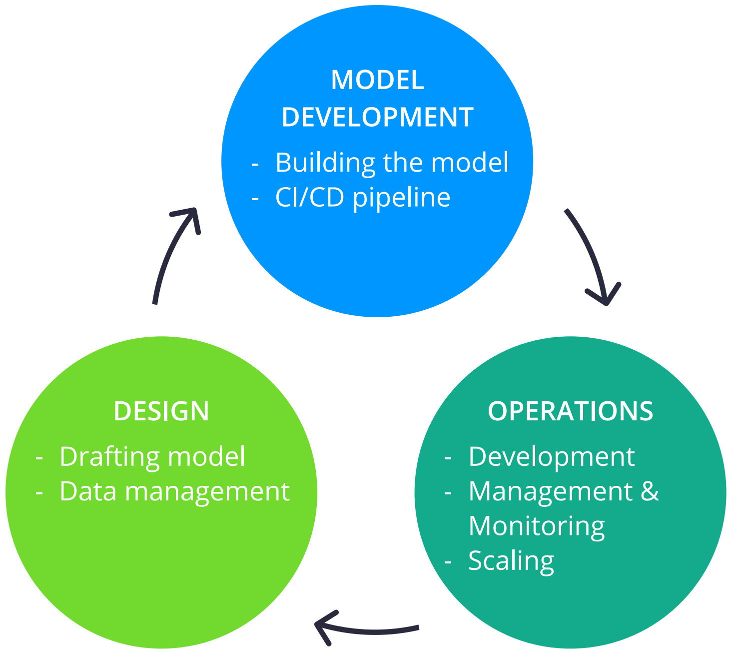


All of these factors work together to accelerate time to deployment for ML models, allowing you to realize the benefits of your machine learning investments more quickly, and reducing time-to-market and time-to-value for new products and applications.

Components of an MLOps Workflow

Another way to think about MLOps delves into what it means to operationalize data science and machine learning solutions. MLOps utilizes code and best practices that promote efficiency, speed, and robustness.

MLOps works hand-in-hand with DevOps to integrate practices with ML systems and therefore produce better ML models, faster. Part of producing better ML models at a faster pace is creating the right environment to make that happen. The right environment is data-driven, catalyzes cross-team collaboration, and uses feedback (both data feedback and real-life feedback) to get better. So how can we do this? We can either build a solution in-house, buy point solutions and cobble them together, or invest in an end-to-end platform.



Components of an MLOps Workflow

The Build-or-Buy Decision

In MLOps, there are various build-or-buy decisions that you need to make when it comes to the tools and platforms you use to build, deploy, and manage ML models. For example, when it comes to model development tools, model training infrastructure, deployment and serving platforms, or monitoring and management tools, you can either build your own or use existing tools. Building your own tools can give you more control and customization options, but it can also be time-consuming and expensive to both build and maintain.

For example, building a homegrown MLOps tech stack includes development time and resources, infrastructure and ever-growing maintenance costs, budget related to scaling the tech stack, and, of course, training. On the other hand, buying existing tools or platforms can save time and money, but they may not be as customizable and give you the same level of control as in-house tools. According to "MLOps in 2023: What Does the Future Hold" nearly 58% of respondents said they are very satisfied with their current MLOps platform/tool or MLOps stack of multiple tools. However, when asked why they may not be satisfied with their current MLOps platform/tools, 41% reported friction with needing to pair it with multiple other tools to get the most value, stating that they don't natively work together which creates problems.

Ultimately, the decision to build or buy depends on your organization's specific needs, resources, and priorities. You may prefer to build in-house tools and infrastructure to have more control and customization options, while others may opt for existing tools and services to save time and money. The right MLOps solution should be able to help you with all facets of the ML lifecycle and workflow, but first we need to make sure we are taking the right steps to evaluate the solution, and to do that, we need to start with the difference between point solutions and end-to-end platforms.

The Difference Between Point Solutions and End-to-End Platforms

Point solutions and end-to-end platforms are two different approaches to building MLOps systems. The main difference between point solutions and end-to-end platforms is their scope and integration. Some end-to-end platforms are also designed to be highly integrated, allowing you to manage all aspects of MLOps within a single system, while point solutions may require you to integrate multiple tools and systems to achieve the same level of functionality. According to ClearML's recent survey, 92% of respondents said they would prefer to use one, unified MLOps platform that does everything, while only 8% prefer using multiple platforms and tools as part of an MLOps tech stack.

Point solutions are tools that address specific aspects of the machine learning lifecycle, such as data management, experiment management, orchestration, compute, scheduling, deployment, or monitoring. These tools are designed to solve a specific problem or provide a specific feature, and they may not be integrated with other tools or systems in the MLOps pipeline.

End-to-end platforms, on the other hand, provide a complete suite of tools and services that cover the entire machine learning lifecycle, from data preparation and model development to deployment and monitoring. These platforms aim to provide a seamless and integrated experience for MLOps, where you can manage all aspects of the pipeline within a single system.

MLOps platforms make the orchestration and deployment of machine learning workflows both frictionless and unified. They provide services for the different components of MLOps workflows through a single pane of glass and most of the time, through a unified user interface that allows end users end-to-end visibility and control. This makes it much easier to incorporate an ML workflow at scale in a new or existing product because much of the continuous ML lifecycle development, deployment and monitoring is done within a centralized unified platform.

Open Source Platforms Versus Managed Platforms

If you decide to go the platform route, you have a choice between open source platforms and managed platforms, each with its own set of pros and cons:

OPEN SOURCE

Pros:

- ✓ Open source platforms are free to use, making them an attractive option for anyone with limited budgets.
- ✓ They provide a high degree of customization and flexibility, allowing you to tailor the platform to their specific needs and requirements.
- ✓ They have a large community of developers and users who contribute to their development, providing a rich ecosystem of tools and resources.
- ✓ They can be deployed on-premises or in the cloud, providing a high degree of control over the infrastructure.

Cons:

- Open source platforms can require a high degree of technical expertise to set up and manage, which can be a barrier for some organizations.
- They may not offer the same level of support or reliability as managed platforms.
- They may require significant time and resources to maintain and upgrade as new versions of the software are released.
- You may need to hire or train staff with specific technical skills to work with open source platforms.

MANAGED PLATFORMS

Pros:

- Managed platforms provide a fully managed and supported solution for MLOps, making them easy to set up and use.
- They provide a high level of reliability and availability, with service level agreements (SLAs) that guarantee uptime and performance.
- They often provide built-in security and compliance features, such as data encryption and access controls.
- They can provide a wide range of additional services and features, such as data storage, analytics, and visualization.

Cons:

- Managed platforms can be more expensive than open source platforms, particularly if you have large datasets or high compute requirements.
- They may not provide the same degree of customization and flexibility as open source platforms.
- You may have less control over the infrastructure, which can be a concern if you have strict data privacy or regulatory requirements.
- They may require vendor-specific skills or knowledge, which can be a barrier for organizations that want to switch vendors or use multiple vendors for different parts of their MLOps pipeline.
- Cloud providers tend to have their own managed platforms that dictate how you work, with very little flexibility to customize.

How To Evaluate an MLOps Solution

Evaluating an MLOps solution involves assessing the platform's ability to manage the end-to-end lifecycle of a machine learning model, from development to deployment and maintenance. Some of the most common steps in evaluating solutions include:

1. **Defining your business requirements:**
   Before evaluating any MLOps platform, it's important to define your specific commercial requirements. Consider your business goals, the size of your data, the complexity of your machine learning models, your team's technical expertise, scale and other factors that might influence your decision.

2. **Evaluating the platform's features:**
   Once you have defined your requirements, evaluate the platform's features against them. Look for features such as data preparation, model training, orchestration, model deployment, monitoring, and automation between the various development steps. Ensure that the platform's features align with your needs and integrate with your current stack seamlessly.

3. **Assess the platform's ease of use:**
   Ease of use is a crucial factor in any MLOps platform. Evaluate the platform's user interface and see how easy it is to navigate, set up workflows, and monitor models. Consider how easily your team can use the platform without extensive training.

4. **Check the platform's scalability and performance:**
   Consider the platform's ability to scale to accommodate growing data and models. Evaluate the platform's performance and latency in deploying and monitoring models.

5. **Evaluate the platform's security and compliance:**
   Assess the platform's security measures, data privacy, and compliance with industry regulations such as ISO 127000, GDPR, and HIPAA.

6. **Evaluate the platform's cost:**
   Finally, evaluate the platform's cost and compare it with your overall ML stack budget. Consider the pricing model, subscription fees, and any hidden costs.

By evaluating an MLOps platform based on these factors, you can determine which platform is best suited for your needs.

Features and Functionality to Look For

Ideally, if you are investing in an end-to-end platform you will want to ensure that it supports, automates, and optimizes every step of the ML lifecycle workflow. For example:

1. Design

The first phase of MLOps workflows is dedicated to designing the right framework for the ML models to be incorporated and used in a business context. There are a lot of business-oriented questions asked during this first phase. For example, what is the goal of this machine learning model? How will it help drive sales, and what kind of data is going to be necessary to train it? How will that data be acquired?

In this step, data scientists work with various internal teams and stakeholders to define and outline the desired needs and scope of the ML model before engineers create it. The design component aims to develop an understanding of the problem statement, the availability of data, and your business objectives. Then, teams might proceed to perform exploratory data analysis, feature engineering, or model selection to design the best machine learning model to address the challenge.

Drafting the Model

During the design phase, it is important to consider things like:

- How will the model be trained and evaluated?
- What kind of data is needed?
- How will the model be deployed in production?
- What performance metrics will be used to evaluate the model?
- What kind of monitoring and alerting will be set up?

You'll want to create a "draft" for the machine learning model by considering questions like these and then developing answers that are as in-depth as possible – the more that's figured out before jumping into creating the model, the more productive and efficient creating the model will be when the project is passed to engineers.

DataOps & Management

Data management can be included in the design component when building an MLOps workflow. The data management component of MLOps covers all aspects of data acquisition and containment. Data needs to be preprocessed, stored, and managed before creating models — data doesn't have to actually be handled as of yet, but the guidelines for how data management will be done should be figured out during this step.

2. Model Development

This step of the MLOps workflow is where we create the machine learning model. Here, the ML model is engineered and data is polished using the steps configured in the design phase. The focus is on building, training, and testing ML models.

Model development is done in notebooks such as Jupyter or Google Colab for lower-level projects, or IDEs that can handle robust Python (ex. PyCharm, VS Code). Common frameworks that you might see used to create ML models in the model development phase include Tensorflow or Pytorch.

During the model development phase, data scientists and engineers will:

- Develop the model architecture
- Train the model
- Test the model
- Perform model validation
- Iterate if needed

Continuous Integration & Continuous Deployment (CI/CD)

In a more developed MLOps workflow (in basic MLOps workflows, sometimes CI/CD isn't incorporated), CI/CD is used to automate the process of building, testing, and deploying machine learning models as code changes are made. For example, a data scientist might make a change to an ML model and push the change to a version control system like Git. The CI/CD pipeline would then automatically build the model, run tests on the change to make sure the change is valid and doesn't cause issues, and then deploy the model to production if the tests pass. This saves a lot of manual labor as well as time.

3. Operations

The operations phase in an MLOps workflow handles the ML model after the bulk of it is already configured and built out. This phase uses these common practices:

- Deployment
- Monitoring
- Model management
- Model serving
- Scaling

Of course, this isn't an all-inclusive list, and every MLOps workflow is different, so you might not even use all these parts – however, these are the most common practices seen across many MLOps workflows.

Deployment

Once the ML model is built, it needs to be efficiently deployed. Cloud-based platforms or on-site servers are used for deployment. CI/CD is also sometimes considered part of this step of the operations phase.

Model Management & Monitoring

Often there are multiple versions of models in production, so this step handles the management of those versions. Monitoring handles the tracking of an ML model's performance in production, and negative changes such as model degradation specifically sought out, as well as:

1. **Model versioning:** Keeping track of model versions and ensuring the right one is being used. Model versioning allows you to switch between previous and new model versions and easily compare their performance.

2. **Model performance monitoring:** Metrics are collected – accuracy, precision, and recall are stored and monitored. This step helps identify degradation.
3. **Model drift detection:** Continuously re-fit (periodically re-train the model to learn from historical data) models according to past model measurements.

Scaling

Many MLOps workflows accommodate scaling changes as part of the management and monitoring phase. This includes scaling resources up or down as needed to handle changes in workload. This is done by changing parameters like CPU and GPU usage, using container systems (such as Kubernetes) to scale up or down data-processing tasks, as well as vertical/horizontal scaling.

Deployment Options

Another consideration when going the MLOps platform route are the various deployment options, each with its own advantages and disadvantages. The four main options are cloud, on-premises, virtual private cloud (VPC), and hybrid:

Cloud:

Cloud-based MLOps platforms are hosted by third-party providers. These platforms are accessible over the internet and can be accessed from anywhere with an internet connection. They offer high scalability, flexibility, and reliability, as the underlying infrastructure is managed by the provider. Cloud-based platforms are ideal if you require rapid scalability or want to avoid the upfront costs and maintenance associated with on-premises infrastructure. However, they can be more expensive in the long run, particularly if you have large datasets or high compute requirements.

On-premises:

On-premises MLOps platforms are deployed on an organization's own hardware and infrastructure. This approach provides maximum control over the infrastructure, as well as data privacy and security. On-premises platforms are ideal if you have strict data privacy or regulatory requirements, as well as those organizations with limited internet connectivity or high data transfer costs. They can be more expensive upfront, as you must purchase and maintain your own infrastructure, but when it comes to GPU cost, especially as it pertains to deep learning use cases, it often makes more sense to purchase machines than to rent them.

Virtual Private Cloud (VPC):

VPC-based MLOps platforms are hosted by third-party providers, but they are deployed within a private network that is isolated from the public internet. This approach provides the benefits of cloud-based infrastructure, such as scalability and flexibility, while also ensuring greater security and data privacy. VPC-based platforms are ideal if you need a high degree of security and privacy, but still want the benefits of cloud-based infrastructure. However, they can be more complex to set up and manage, and may require specialized networking knowledge.

Hybrid:

Hybrid MLOps platforms combine two or more deployment options, such as cloud and on-premises infrastructure. This approach provides maximum flexibility and allows you to leverage the benefits of both cloud-based and on-premises infrastructure. Hybrid platforms are ideal if you have varying compute and storage requirements, or those that want to use existing on-premises infrastructure while gradually transitioning to the cloud. However, hybrid platforms can be more complex to set up and manage, and may require specialized skills and knowledge.

Considerations for Service & Support

When evaluating an MLOps platform vendor, there are several key considerations to take into account when it comes to service and support. These include:

- **Technical support**:
  It's important to understand the level of technical support that the vendor provides, including their hours of availability, response times, and the types of issues that they will help you troubleshoot. You should also consider whether the vendor provides onboarding support to help you get started with their platform.

- **Training and documentation:**
  Look for vendors that provide comprehensive documentation and training resources, including online tutorials, user guides, and technical documentation. This will help you get up to speed with the platform quickly and ensure that you can take full advantage of its features and functionality.

- **Maintenance and upgrades:**
  Consider the vendor's approach to maintenance and upgrades. Will they provide regular updates to the platform to address bugs and add new features? How frequently are updates released, and how are they communicated to users? You should also consider whether upgrades are included in the cost of the platform, or if they are an additional cost.

- **Data security and compliance:**
  Make sure that the vendor has robust security measures in place to protect your data and ensure compliance with relevant regulations, such as GDPR or HIPAA. You should also consider whether the vendor has undergone independent security audits or obtained relevant certifications

- **Customer references:**
  Ask the vendor to provide customer references, and take the time to speak with these references to get a sense of their experience with the platform and the vendor's service and support. You can also look for reviews and feedback online to see what other customers are saying about the vendor.

- **Vendor stability:**
  Finally, consider the vendor's stability and longevity in the market. Look for vendors that have a proven track record of success and a strong customer base. You should also consider whether the vendor is well-funded and has a clear roadmap for the future development of their platform.

The Importance of Scalability, Security, and Governance

Lastly, scalability, security, and governance are critical considerations when implementing MLOps. Here's why:

SCALABILITY

MLOps involves managing large amounts of data and complex machine learning models, which can be resource-intensive. Scalability is important because it ensures that you can scale your infrastructure to handle increasing amounts of data and models, without compromising performance or reliability. This means that as the organization grows, they can continue to develop and deploy machine learning models efficiently and effectively.

SECURITY

MLOps platforms often deal with sensitive data, such as personally identifiable information or financial data, and can be vulnerable to attacks. Security is important because it ensures that data is protected against unauthorized access, and that models are protected against attacks such as data poisoning or adversarial attacks. Implementing security measures such as encryption, access controls, and user authentication can help prevent data breaches and protect against malicious attacks. Another way to protect data is to ensure that the platform never actually sees the data – sensitive data should be abstracted before passing it through encrypted channels for use in model training.

GOVERNANCE

MLOps involves managing complex machine learning models that can be difficult to understand, monitor, and manage. Governance is important because it provides a framework for managing and tracking data, models, and processes. By implementing governance processes, organizations can ensure compliance with relevant regulations and standards, mitigate risks such as model bias or errors, manage data effectively, and facilitate collaboration between different teams.

Introducing ClearML

Our founders, Moses Guttmann, who serves as CEO, and Gil Westrch, who serves as CTO, are long-time colleagues and friends who initially met in the IDF's (Israel Defense Force) Unit 81 – its elite technology unit – and have been working together ever since.

With the rapid rise of machine learning (ML) and deep learning across all sectors of the economy – academia, non-profits, innovative start-ups, and mid-market and enterprise companies – Moses and Gil saw the need for automating, optimizing, and scaling ML processes and set to work creating tools that could be easily integrated into existing workflows.

ClearML is now the leading open source, end-to-end MLOps platform used by more than 150,000 ML Practitioners and more than 1,300 enterprises worldwide, that helps data science, ML practitioners, product management, and DevOps teams easily develop, orchestrate, and automate ML workflows at scale. It is designed as a frictionless, unified, end-to-end MLOps suite enabling users and customers to focus on developing their ML code and automation, ensuring their work is reproducible and scalable.

Among its many accomplishments, the company is a founding member of the AI Infrastructure Alliance (AIIA), the holder of more than 10 patents (with more on the way), and is certified to run NVIDIA AI Enterprise, an end-to-end platform for building accelerated production AI. Our motto is "Powered by Open Source. Free Forever. Made with ❤️ by ClearML" – and we live and breathe every word of it, every day.

ClearML is a unified, open source end-to-end platform that supports the entire ML lifecycle from research to production. The platform's open architecture lets you choose whether to use our best-of-breed modules or swap in your existing tools, making it an easy way to plug into, as well as augment, your existing processes and systems.

Explore ClearML's End-to-End Platform for Continuous ML

Easily develop, integrate, ship, and improve ML models at any scale with only 2 lines of code. ClearML delivers a unified, open source platform for continuous ML. Use all of our modules for a complete end-to-end ecosystem, or swap out any module with tools you have for a custom experience – we already work with everything you use:

DataOps

Data Management, Catalog, Versioning, & Hyper-Datasets

Experiment

Experiment Management & Visualization

Train

Model Training & Lifecycle Management

Reports

Collaborative, Dashboards & Reporting

Modelstore

Model Management, Repository, & Versioning

Pipelines

Automation (CI/CD) & Pipelines

Deploy

Model Serving & Monitoring

Schedule

Complete Visibility Over Your Infrastructure Usage

Orchestrate

Automagically Package & Ship Environments Into Remote Machines

Compute

Reduce Compute, Hardware, and Resource Spend to Optimize Cost and Performance

Available as a Unified Platform or Modular Offering

DataOps Data Management, Catalog, Versioning	Hyper-Datasets	Experiment Experiment Management & Visualization	Train Model Training & Lifecycle Management
Use ClearML to abstract your data, make it traceable and accessible from anywhere, and usable by anyone. Interface with your data from any machine, catalog your data to make it more accessible, and easily integrate with your codebase.	ClearML's feature store for unstructured data, allowing you to explore, visualize, and query your data. With ClearML, enterprise organizations can share data across the organization with Hyper-Datasets' visual exploration tool. Automate ingestion pipelines with full versioning control capabilities, creating complete, traceable ML workflows.	Get a single source of truth for any experiment at any point in time from any machine. ClearML logs everything: codebase, uncommitted changes, Jupyter Notebooks, Python packages, containers, arguments, configuration files, metrics, outputs, plots, and even datasets.	Work more collaboratively and get feedback to help your models reach production-ready status faster. ClearML lets any team member launch any shared experiment on your compute infrastructure, change arguments, and review results, automating the process.
Reports Collaborative Dashboards & Reporting	**Modelstore** Model Management, Repository, & Versioning	**Pipelines** Automation [CI/CD] & Pipelines	**Deploy** Model Serving & Monitoring
Showcase and discuss your findings with colleagues, managers, or your future self. Summarize top-performing models, data, or experiments in interactive, and real-time dashboards Embed live graphs, charts, and plots into a ClearML Report or your favorite third-party knowledge-sharing tool.	Catalog and share models with full traceability and provenance. Interface with any model using the queryable model catalog, add performance metrics, and trace lineage.	Scale workloads with multi-node pipelines straight from your code. Create multi-node, optimized pipelines with cached, logic-driven execution flow.	ClearML gives you two options to deploy and service any model on your infrastructure: 1) Single-click batch model deployment on your infrastructure. Launch and schedule batch deployment from the UI or iterate into your CI/CD GitOps workflow. 2) Expose any model via REST API with a single CLI command. Get real-time blue/green model deployment including canary support with enterprise-grade scaling.

Orchestrate	Schedule	Compute	Autoscalers
Package and ship environments into remote machines. Launch from anywhere (code, CLI, Git, Web) into your infrastructure (cloud, on-prem, Docker, Kubernetes, Slurm) – no need to manually containerize every job.	Expose and control computer resources. Simplify priority and job order with execution queues. ClearML lets you monitor and optimize utilization of GPU and CPU resources by user group, order, priority, and budget.	Minimize overhead, simplify maintenance, and maximize compute infrastructure utilization. with a more accessible, flexible, and utilized compute infrastructure. Open up controlled access to compute resources by abstracting workflow requirements from machines, whether Kubernetes, bare metal, cloud, or a mix.	Dynamically launch jobs on your cloud with elastic budget control. Transparently orchestrate and schedule jobs in the cloud with budget-controlled, dynamically provisioned instances.

Enterprise-Grade Security & Governance

With ClearML for Enterprise, you can turn your ML models into enterprise-grade production systems that run reliably and automate your machine learning pipelines, from data ingestion to the generation of business insights. Deploy on your VPC or on-prem to ensure the highest standards of security, customize your own security policies, and retain full control of accessing data.

When it comes to security, role-based access controls, authentication with SSO and LDAP, and audit trails for full traceability are just a few of the ways ClearML can limit and track user access to critical systems and data assets.

The ClearML control plane offers federated out-of-the-box support for object storage and NFS/CIFS on-prem solutions. We never see your data, ensuring maximum data governance inside your organization. ClearML is ISO 27001-certified so you can be assured you've taken the necessary steps maximizing secure development, auditing, and governance.

ClearML seamlessly integrates with your current architecture and maps to your pre-existing workflows, processes and ML stack, so you can enjoy lightspeed deployment and adoption across teams and business units and rapidly drive value from your ML investments.

Open Architecture & Extensibility

We're built with open source, run on Python, and integrate out-of-the-box with any Python framework and existing tools. Our flexible and open architecture means you can work the way you want, with no vendor lock-in. We're easy to use and built from the ground up to enable all of your stakeholders, whether they're data scientists, ML engineers, or DevOps. For example, data scientists can self-serve (reducing the burden on DevOps) to instantly spin up new environments and begin working immediately, saving an average of 85 hours per instance.

There's no need to rip and replace; with ClearML, you'll optimize the infrastructure you already have (and maximize those prior investments) by connecting your existing workflows and point solutions with our solution. And, you'll accelerate your end-to-end workflows with ClearML's orchestration and automation capabilities that seamlessly integrate with your in-house solutions as well as external point solutions and services.

Deployment Options

ClearML offers four deployment options; choose the configuration that best supports the way you want to work:

Managed ClearML SaaS

Cloud deployment with dedicated single tenant server. All data and compute stays behind your firewall / VPC.

Virtual Private Cloud (VPC)

Self-managed with full remote support. All data stays behind your firewall.

On-Prem

Self-managed with full remote support. Run ClearML on your own on-prem servers and personal devices.

Managed VPC

Fully managed by ClearML on your VPC sub-account.

World-Class Services & Support

Benefit from unmatched performance, scalability, and access control for production ML with all the white glove support you need. That means multiple channels of communication and support with our MLOps experts and support specialists – email, phone, Slack, you name it. We're here when you need us.

Accelerate deployment and decrease time to onboard new team members: ClearML's intuitive interface and support for existing tools in your stack means that little or no training is needed. Whether it's white glove support and customer service, lightspeed onboarding, or the creation of custom enterprise apps and integrations, our expert MLOps team provides all the DevOps, product, and technical support you need.

How ClearML Compares to Other Solutions

ClearML vs. Data Versioning Point Solutions

Managing your data on ClearML is painless and easy. Our automated data logging tracks every change for every run, making experiments repeatable without eating up a lot of storage or managing tons of credentials. The intuitive features built into the UI, such as tags, sampling, and visualization, make understanding and sharing your data catalog much easier. And ClearML gives complete visibility of your data's lineage and where it is being used in tasks and pipelines. Once the data is configured, it can be accessed by anyone (with the right permission settings) from anywhere.

Connected with the larger ClearML continuous machine learning platform, your data is connected to experiments and jobs as well as outputs. Reproduce results with 100% accuracy and progress faster by running cloned experiments with minor argument adjustments.

Enterprise customers leveraging our Hyper-datasets feature can access elegantly abstracted metadata that enables oversampling and debiasing without the difficulties of working with cumbersome large files. Metadata and annotations are automatically versioned into a database, letting you slice the data in real time and change the streaming data sources going into tasks, experiments, jobs, etc. ClearML Hyper-datasets is a powerful tool for making data easy to access, regardless of where it physically resides.

ClearML vs Experiment Management Point Solutions

Experiment Management solutions are the most common MLOps tools on the market today. While there are superficial differences between interfaces, performance evaluation dashboards, and different defaults for logging to the cloud or a local drive, the big differences show themselves when you are trying to compare multiple runs, execute tasks faster, or manipulate very large datasets.

Fundamentally, ClearML's unified MLOps platform connects orchestration into every part of experiment management, automatically logging configs, parameters, and settings for every run and facilitating faster progress overall through the use of hashing, caching, and metadata. Experiments are immediately reproducible and can be re-run using different parameters with just one click.

ClearML was designed to make life easier for data scientists and ML engineers. Building pipelines is intuitive and dare we say, enjoyable. We support project hierarchies with RBAC, track literally everything (including Git diffs), and allow you to compare as many experiments as you want against each other (up to 100, unlike other tools that only allow you to compare up to 8 experiments!). For Enterprise customers, our Hyper-datasets feature easily supports use cases for computer vision, NLP, or any other deep learning using huge, heavy, unstructured data files.

Deploying trained models into production is seamless on ClearML, as all parts of the operation already "speak the same language" across the platform. There is no need to reformat or reconfigure anything, and setting up a true CI/CD feedback loop is not only possible, it's easy. Your DevOps team will really appreciate that.

ClearML vs Scheduling and Orchestration Point Solutions

ClearML's built-in capabilities let you fully schedule and orchestrate and automate your machine learning workflows and pipelines without paying for any other software. Working on ClearML is easy and flexible – you can deploy straight from your development environment and launch from anywhere (code, CLI, Git, or Web UI) into your infrastructure (cloud, on-prem, Docker, Kubernetes, or Slurm). Your infrastructure could be fully cloud-based or on-prem, it doesn't matter. ClearML is cloud-agnostic and also supports installation on bare metal (no Kubernetes required).

ClearML eliminates the need to containerize every job, saving time, storage, and storage cost.

Build transparent, prioritized job queues that let team members see the full breadth of current work, re-order their jobs, and collaborate more often and easily. Max out your compute utilization! ClearML is certified to run the NVIDIA AI Enterprise Software Suite and allows customers to fully optimize the GPU power of NVIDIA DGX™ systems and NVIDIA-Certified Systems™. Take advantage of NVIDIA's new capability for GPU splicing and create up to 7 partitions per GPU and match the correct computing power per workload. Fractional and virtual GPUs created by ClearML can be accessed by containers, allowing different workflows to run in tandem on a single GPU. All out of the box.

Need more muscle? ClearML Autoscalers enable you to spin up or spillover compute on AWS, GCP, or Azure on an as-needed basis, and usage can be controlled with team-level budget caps. ClearML enables more people to access compute power with less overhead.

Through ClearML, administrators can set up users with their permission levels once, without the stress of provisioning more machines, managing credentials, or tracking compute usage. Our modular architecture and open source nature make integrations with other tools easy. Managing your entire MLOps infrastructure overall becomes a much lighter burden for your engineering team.

Summary

MLOps is an ever-growing field in AI and machine learning/deep learning that helps AI/ML-driven solutions go to market faster. If you're looking for an MLOps platform that can help your product or business incorporate machine learning, consider ClearML – an open-source platform that automates MLOps solutions for thousands of data science teams across the world. As you've seen, ClearML enables Data Scientists, ML Engineers, and DevOps to automate and optimize machine learning work on a single, open source platform that does all your heavy lifting. However, whether you prefer one MLOps platform that handles all of the work or prefer many specialized tools is ultimately a decision your organization has to make. The choice is yours!

Checklist for Evaluating MLOps Solutions

If you are evaluating MLOps solutions, either for PoC or purchase, here is a handy checklist to see if your solution has the features and functionality you need.

Feature/Functionality	CLEAR ML	How Does Your Solution Compare?
Data Management, Catalog & Versioning	◔	
Experiment Management & Visualization	✓	
Model Training & Lifecycle Management	✓	
Collaborative Dashboards & Reporting	◔	
Model Management, Repository & Versioning	◔	
Automation (CI/CD) & Pipelines	✓	
Model Serving & Monitoring	◔	
Orchestration	✓	

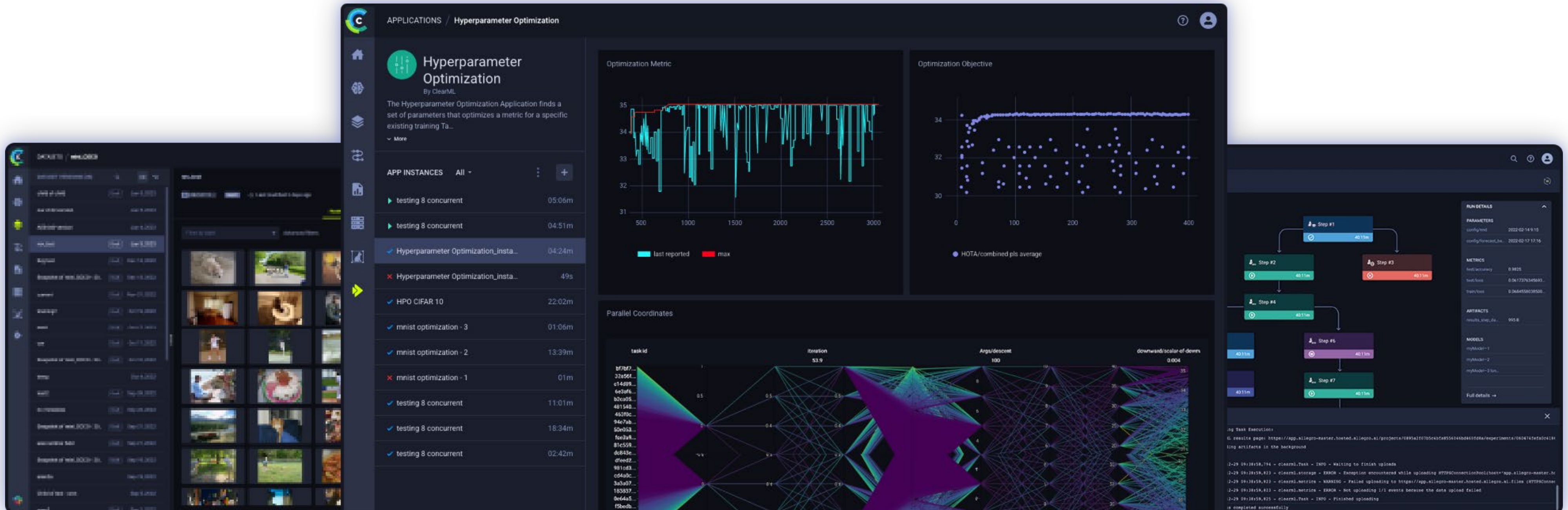
Scheduling	
Compute Resources	
Point Solution or End-to-End Platform?	End-to-End Platform
Open Source or Managed?	Fully Open Source
Deployment Options	Saas, On-prem, VPC & Hybrid
Service & Support	Enterprise customers receive white glove support and customer service, lightspeed onboarding, and the creation of custom enterprise apps and integrations for all the DevOps, product, and technical support you need.
Security & Governance	Role-based access controls, authentication with SSO and LDAP, and audit trails for full traceability. ISO 27001-certified.

Next Steps

- Get started with ClearML by using our free tier servers or by hosting your own.
- Read our documentation here.
- You'll find more in-depth tutorials about ClearML on our YouTube channel.
- We also have a very active Slack channel for anyone that needs help.
- Read how our customers have been successful in using ClearML as their platform for continuous machine learning. There are more customer success stories here.

If you need to scale your ML pipelines and data abstraction or need unmatched performance and control, please

REQUEST A DEMO



About ClearML

ClearML is a unified, open source platform for continuous machine learning (ML), trusted by forward-thinking Data Scientists, ML Engineers, DevOps, and decision makers at leading Fortune 500 companies, enterprises, academia, and innovative start-ups worldwide. We enable customers to build continuous ML workflows -- from experiment management and orchestration through data management and scheduling, followed by provisioning and serving -- to achieve the fastest time to ML production, fastest time to value, and increased performance. In this way, ClearML accelerates ML adoption across business units, helping companies reach their revenue potential and materialize their ML investments. With thousands of deployments and a vibrant, engaged community, ClearML is transforming the ML space -- bridging software, machine learning, and automation. To learn more, visit the company's website at https://clear.ml.