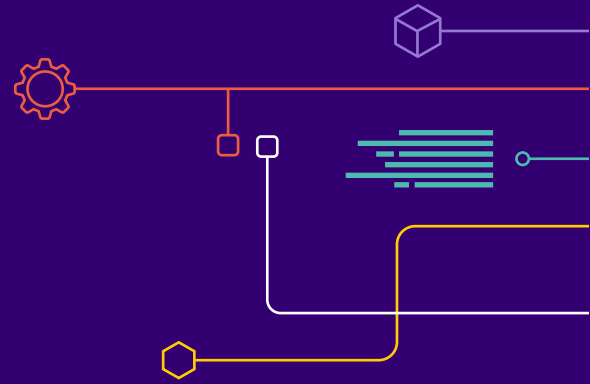




SOLUTION BRIEF

Kaptain AI/ML



Overview

Kaptain AI/ML is an artificial intelligence (AI) and machine learning (ML) platform designed to simplify AI/ML operations to enable data scientists to focus on furthering business goals rather than configuring complex AI/ML infrastructures.

Using the notebooks-first approach with which data scientists are most comfortable, Kaptain AI/ML provides best-of-breed tools, libraries, and frameworks to enable data scientists to harness the scalability and flexibility of Kubernetes without having to struggle with its complexity.

Kaptain AI/ML in the D2iQ Kubernetes Platform (DKP) enables data scientists to be separated from the underlying Kubeflow and Kubernetes operations, infrastructure, and lifecycle management to enable their organizations to quickly and easily attain the benefits of AI/ML, get to market faster, and boost business results at a significantly reduced cost. This enables data scientists to focus on their jobs rather than managing a complex underlying environment.

Increased Productivity at Lower Cost

Kaptain AI/ML breaks down operational barriers to streamline and accelerate the process of getting models into production and improve the success of AI/ML projects, while reducing the total cost of ownership (TCO) of running AI and ML at scale.

Kaptain AI/ML, with the full support of D2iQ, ensures that investments in AI/ML deliver a strong return on investment (ROI) for your organization by ensuring that models become production-ready rapidly.

End-to-end AI/ML requires a curated and tested set of components that extends beyond notebooks to include model frameworks, libraries, tools, observability, and security required for AI/ML model development and deployment. Kaptain AI/ML meets these needs by bundling best-of-breed open-source technologies to build, automate, deploy, track, and monitor models across their entire lifecycle.

Fully automated deployment with integrated observability, rich security, and cost management make Kaptain AI/ML enterprise-ready for Day-2 operations out of the box.

Platform + People = Kubernetes Success

D2iQ enables customers to successfully deploy DKP and Kaptain AI/ML through a combination of simplified platform and expert support and training services.

D2iQ's technical support experts are masters in troubleshooting and diagnosing technologies such as Kubernetes and accompanying open-source components from the Cloud Native Computing Foundation (CNCF) ecosystem included with DKP.

D2iQ's support experts have extensive experience supporting the data technologies, certified operational frameworks, and AI/ML components required to support enterprise deployment of AI/ML initiatives in production at scale. Our full stack support offers the lowest TCO in the industry.

D2iQ offers end-to-end expert support from basic to robust. These include Premium and Signature support options with 24x7 coverage and some of the strongest service-level agreements (SLAs) in the industry to ensure that your business and AI/ML initiatives have the dedicated support and response time they require.

Reduced Complexity, Increased Business Results

AI/ML are incredibly complex, resource-intensive processes with high failure rates. It takes a tremendous amount of time and effort to gather, aggregate, and process the required data, and to build and train successful AI/ML models.

At the core of Kaptain AI/ML is Kubeflow, an open-source framework for developing, managing, deploying, and running scalable and portable AI/ML workloads on Kubernetes. While Kubeflow and Kubernetes would seem to be an ideal way to address AI/ML challenges, the steep learning curve of Kubeflow and Kubernetes can introduce complexities for data scientists and data engineers, many of whom do not have the technical know-how to design or manage these technologies.

Kaptain AI/ML reduces the complexity of Kubeflow and Kubernetes through automation, tooling, and integration of a select group of services. Kaptain AI/ML supplements Kubeflow with a carefully curated set of Kubeflow projects (20+ Kubernetes operators), a software development kit (SDK) that simplifies and abstracts Kubeflow functionality, and additional components (like Spark and Horovod) that are not part of the base Kubeflow ecosystem.

These operators are defined with Helm, an open-source Kubernetes deployment tool for automating creation, packaging, configuration, and deployment of applications and services to Kubernetes clusters. Helm integration provides full lifecycle support, which is absent from the upstream version of Kubeflow.

Likewise, Kaptain AI/ML provides the enterprise-level security that organizations require in production environments, which also is absent in the open-source Kubeflow version. These enhancements make Kaptain a significantly more capable and consumable enterprise-grade distribution of Kubeflow.

Kaptain AI/ML is the only Kubernetes-based AI/ML platform that has made Kubeflow enterprise-ready through these engineering modifications and extensions. Through these Kubeflow enhancements, Kaptain AI/ML enables customers to go from prototype to production in a matter of days or hours, rather than weeks or months, and to manage entire AI/ML pipelines on any infrastructure.

Fully Loaded to Harness AI Power with Greater Ease

Kaptain AI/ML is an add-on to D2iQ's DKP Essential and DKP Enterprise. It comes with an independent SDK, which is a Python API that abstracts away underlying Kubeflow details, enabling data scientists to focus on building, tuning, and deploying models.

Kaptain AI/ML comes with everything an organization needs to hit the ground running, such as hyperparameter tuning, model tracking, military-grade security, real-time cost management, and other necessary enterprise-grade functionality at no additional cost.

The Jupyter Notebooks that data scientists favor are preinstalled with many top libraries such as SciPy and Keras, and with deep learning frameworks such as PyTorch, Tensorflow, and MXNet. The Kaptain AI/ML add-on has out-of-the-box graphics processing unit (GPU) support and is certified on Nvidia's DGX platform.

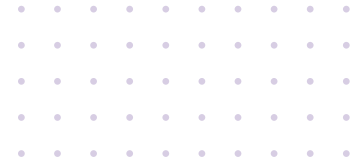
Built-in and fully tested support of GPUs enables analysts to focus on conducting deep learning operations without having to manually deploy drivers. In-depth notebook tutorials help speed up onboarding, time to value, and production readiness of AI/ML models at scale.

Complete Deployment Flexibility and Agility

DKP, from which Kaptain AI/ML is deployed and managed, is built on pure upstream open-source components, giving you the freedom to leverage continual open-source innovation with immunity from lock-in to proprietary solutions, with the added benefit of yielding the lowest TCO.

Kaptain AI/ML can be deployed on any cloud, multi-cloud, hybrid-cloud, on-premise, and air-gapped environment, wherever required for mission-critical AI/ML initiatives. This gives you the agility to deploy and manage AI/ML operations in public clouds, private clouds, and secure on-premise locations, and to easily change and migrate deployments without incurring excess cost.





Further Improvements in DKP 2.2

The already powerful capabilities of Kaptain AI/ML have been enhanced in the latest release of the D2iQ Kubernetes Platform (DKP). The enhancements include integration with DKP Essential and DKP Enterprise that make Kaptain AI/ML even easier to deploy and manage.

The integration of Kaptain AI/ML with DKP enables Kaptain AI/ML to run across multi-cluster environments on any CNCF certified Kubernetes platform, including Amazon Elastic Kubernetes Service (EKS), Microsoft Azure Kubernetes Service (AKS), and Google Kubernetes Engine (GKE). Kaptain AI/ML also offers easier installation via Helm and DKP catalog integration. These new capabilities provide more productivity, flexibility, and choice when deploying Kaptain in Kubernetes environments.

Kaptain AI/ML also gives data science teams expanded control for mounting data volumes and increased visibility into idle notebooks, so they can spend more time developing and less time debugging resource capacity. The enhanced user experience enables data scientists to more effectively manage the lifecycle of AI and ML models without the need for infrastructure knowledge and skill sets.

For more information about the DKP Essential and DKP Enterprise offerings, see the [DKP Essential and DKP Enterprise solution briefs](#).



Kaptain AI/ML Eliminates Barriers to AI/ML Success

Kubernetes adoption is growing at an explosive rate as organizations adopt Kubernetes platforms to improve their agility and deliver new and innovative digital services, scale to meet increasing customer demands, and achieve 99.999% uptime. With the increasing maturity of Kubernetes, its ability to scale, and improvements to support stateful workloads, organizations are steadily expanding the use cases for the platform.

Kubernetes is a natural fit for AI/ML because of its ability to meet the scalability needs of AI/ML workloads as well as the continuous development nature of AI/ML models.

There are, however, several significant challenges to overcome, including:

High risk: Up to 87% of AI/ML initiatives are abandoned before they reach production.

Long time to value: For those initiatives that do make it to production, it can take more than 3 months for a single model to be deployed. Software provisioning at enterprises can take weeks or even months, which adds time and delays to obtaining value.

New technology: AI/ML platforms for big data and deep learning have only been around since 2016, with few technologies that are cloud-native.

Complexity: Build-or-buy decisions for scalable platforms require immense knowledge of cloud-native infrastructure as well as the entire AI/ML landscape.

Investments in AI/ML projects that don't make it to production are completely lost. The longer it takes to get models into production, the longer it takes to reach a meaningful ROI.

There are also organizational challenges in assembling the teams required to move AI/ML models from development to production. Typical roles to fill include data engineers responsible for managing data acquisition and cleansing, data scientists responsible for creating, testing, and tuning models, machine learning engineers for deploying and monitoring models, and infrastructure teams for creating and managing the platform on which all the work needs to be done.

Kaptain AI/ML is designed to reduce complexity and remove operational barriers to enable your organization to seamlessly move AI/ML models to production and improve the success rate of AI/ML projects, while also reducing the cost of running AI/ML at scale.



To learn more about how D2iQ can be your partner in the cloud native journey, visit [D2iQ.com](https://d2iq.com)