



**Hewlett Packard  
Enterprise**

Solution overview

Check if the document is available  
in the language of your choice.



# EDGE-TO-CORE DATA ENGINEERING PLATFORM

Accelerate and automate the data engineering processes with a  
platform solution





## LAST MILE OF MACHINE LEARNING DEPLOYMENT

Over the last several years, the explosion of data and the simultaneous reduction in computing and storage costs, combined with the introduction of novel methods for machine learning (ML), have provided the ability to both uncover insights and unlock new capabilities across customer verticals. Organizations have started building out data science and ML teams that focus on model development. Yet 90% of ML models never make it into production.<sup>1</sup> Worse, highly demanded and skilled data scientists spend about 45% of their time on data preparation tasks, including loading and cleaning data.<sup>2</sup>

The challenge for data teams is beyond the ML code because of:

- **Infrastructure complexity:** In production deployments, ML code is surrounded by many contributory systems, including training, deployment, and serving infrastructure, monitoring tools, stream processing frameworks, databases, and data storage systems.
- **Skills gap:** There is also a significant skills gap in both data engineering (the process of acquiring, transforming, storing, and serving data in a manner accessible to ML teams) and in ML operations ([ML Ops] the process of operationalizing ML models themselves, including their deployment, monitoring, training, and such).
- **Edge deployment:** When data and ML models are geographically spread out to so-called edge sites even more contributory systems need to be successfully deployed.
- **Open-source integration:** The data engineering and ML Ops tools that exist in the open-source space today may be hard to deploy and require significant effort in integrating them in an end-to-end workflow.
- **Cloud experience:** End-to-end public cloud solutions can end up being expensive at scale, and many organizations may prefer the benefits of on-premises hardware and data retention if provided a solution that rivals the ease of use and flexibility of the cloud such as HPE GreenLake cloud services.

## ACCELERATING TIME TO PRODUCTION OF ML/AI SOLUTIONS THROUGH A PREINTEGRATED, TESTED DATA ENGINEERING PLATFORM

Leveraging the vast experience gained over the last several years working directly with organizations to build and deploy data platform solutions to production, the HPE experts have created a preintegrated set of pluggable components to serve as a starting point for an end-to-end edge-to-core platform solution. A data platform is a technology solution or a service that allows you to ingest, process, store, access, analyze, and present huge volumes of data. The solution addresses both the data engineering and ML Ops skills gap and the complexity in integrating these systems by delivering a composable edge-to-core platform blueprint for an organization's modern data-intensive workloads addressing two possible deployment scenarios:

1. On-premises deployment that is mostly confined to the data center (core). This is predominantly for customers where the production and consumption of AI (including analytics) happen at the core.
2. Deployments both at the core and multiple edge sites. The models are typically trained at the core by aggregating data from the edge sites; however, the trained models are deployed at the edge for real-time inference. This introduces new challenges, including but not limited to the management of edge nodes and deployment/tracking of multiple simultaneous models, and represents substantially increased complexity.

<sup>1</sup> [towardsdatascience.com/why-90-percent-of-all-machine-learning-models-never-make-it-into-production-ce7e250d5a4a](https://towardsdatascience.com/why-90-percent-of-all-machine-learning-models-never-make-it-into-production-ce7e250d5a4a)

<sup>2</sup> The State of Data Science 2020—Moving from hype toward maturity  
[anaconda.com/state-of-data-science-2020?utm\\_medium=press&utm\\_source=anaconda&utm\\_campaign=sods-2020&utm\\_content=report](https://anaconda.com/state-of-data-science-2020?utm_medium=press&utm_source=anaconda&utm_campaign=sods-2020&utm_content=report)



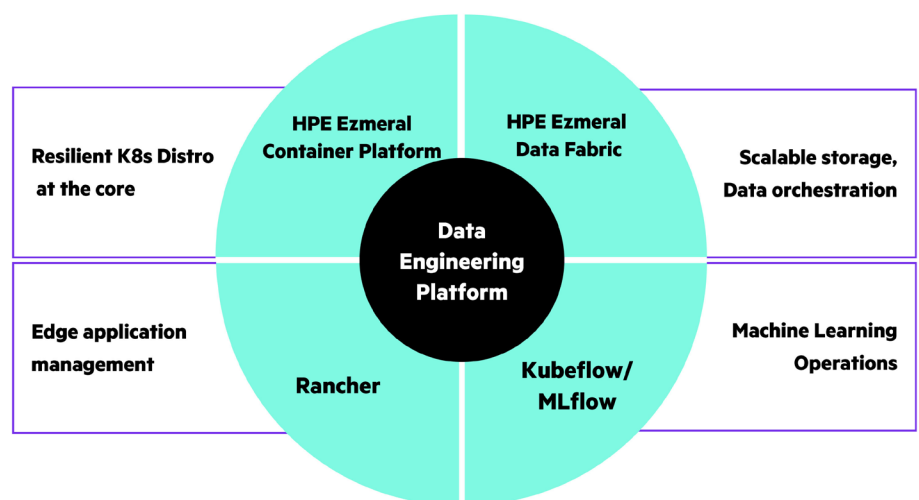




## ARCHITECTURE

Hewlett Packard Enterprise offers end-to-end suite of products services for your data platform needs including core (HPE ProLiant DL and HPE Apollo series), edge hardware (HPE Edgeline, Aruba), software (HPE Ezmeral), HPE Pointnext Services (Advisory and Professional Services), and cloud option (HPE GreenLake). The platform is a sum of four distinct, integrated components:

1. The core platform is based on industry-standard Kubernetes via the [HPE Ezmeral Runtime](#), which builds on open-source Kubernetes, adding features demanded by enterprises, such as enterprise-grade security and controls, an ML-focused App Store, and multicluster management.
2. [HPE Ezmeral Data Fabric](#) provides a scalable, proven, and state-of-the-art file system with a unified global namespace, allowing data access across the edge and core.
3. Edge application deployment and management is orchestrated via Rancher and K3s, enabling continuous delivery to a very large number of edge sites.<sup>3</sup>
4. Development, deployment, and monitoring of ML workflows are orchestrated via Kubeflow, a prevailing industry-standard set of tools purpose-built to solve contemporary ML problems at scale.<sup>4</sup>



**FIGURE 1.** Edge-to-core data engineering platform overview

<sup>3</sup> [rancher.com/](https://rancher.com/)

<sup>4</sup> [kubeflow.org/docs/started/kubeflow-overview](https://kubeflow.org/docs/started/kubeflow-overview)



## GETTING STARTED WITH THE RIGHT APPROACH AND EXPERTISE

Whether you are exploring your ML/AI use case, experimenting it through pilots, or scaling with a modern data engineering platform, you can work with AI and data experts of HPE Pointnext Services. Refer Table 1 to find out what services are on offer:

**TABLE 1.** Advisory and Professional Services for data platforms

Workshop	Capability assessment	Design	Agile implementation
Align the business and IT; gain stakeholder support and common understanding	Evaluate the future data platform and assess what to retire, what to modernize, and what to adopt	Design the data platform and plan the implementation on-premises, at the edge or in hybrid cloud	Customized implementation including infrastructure, software, and integrations with operational guidelines
<ul style="list-style-type: none"> <li>• Understand the scope, scale, and critical success factors</li> <li>• Identify quick wins that generate momentum</li> <li>• Leverage best practices from consolidation, automation, cloud like experience, and containerization</li> </ul>	<ul style="list-style-type: none"> <li>• Assess the maturity of your data platform</li> <li>• Evaluate the effort for achieving the future data platform</li> <li>• Understand what to discard, what to save, what to modernize, and what new to adopt</li> </ul>	<ul style="list-style-type: none"> <li>• Design based on target use cases, functional and non-functional requirements</li> <li>• Get an integration plan, standards to support advanced analytics (AI/ML), new use cases, new data flows</li> <li>• <b>Follow</b> infrastructure and operations guidelines</li> </ul>	<ul style="list-style-type: none"> <li>• Implementation of the designed data platform including infrastructure, software, and integrations, operational guidelines</li> <li>• Trainings and knowledge transfer</li> </ul>

HPE is a leader in worldwide AI IT services.<sup>5</sup> Our data and AI experts play a key role in leading complex AI and data projects, as well as managing and integrating solutions with the broader AI ecosystem. They help operate AI and data infrastructure as a service, consumption-based models with [HPE GreenLake](#).

## LEARN MORE AT [hpe.com/services/ai-data](https://hpe.com/services/ai-data)

<sup>5</sup> IDC MarketScape: Worldwide Artificial Intelligence IT Services 2021 Vendor Assessment, Jennifer Hamel, IDC #US46741821, August 2021

Make the right purchase decision.  
Contact our presales specialists.



Chat



Email



Call



Get updates