White Paper

Accelerating AI Initiatives with HPE and BlueData

Faster Time-to-Value for AI Deployments

By Mike Leone, ESG Senior Analyst; and Leah Matuson, Research Analyst

June 2019

Contents

Introduction

Enterprise-class organizations across all industries are looking to embrace artificial intelligence (AI). They are beginning to recognize the game-changing and transformative benefits they can attain with AI—and they need to arm their data science teams with the tools and data to drive innovation and business value with machine learning (ML) and deep learning (DL). Data science teams want to leverage their preferred frameworks (TensorFlow, PyTorch, Theano, Caffe2, etc.); learning algorithms (decision trees, Naive Bayes, random forest, XGBoost, etc.); and applications and tools (Jupyter notebooks, RStudio Server, AutoML, H2O.ai, Dataiku, etc.) to develop their models—from data selection and training, to model deployment.
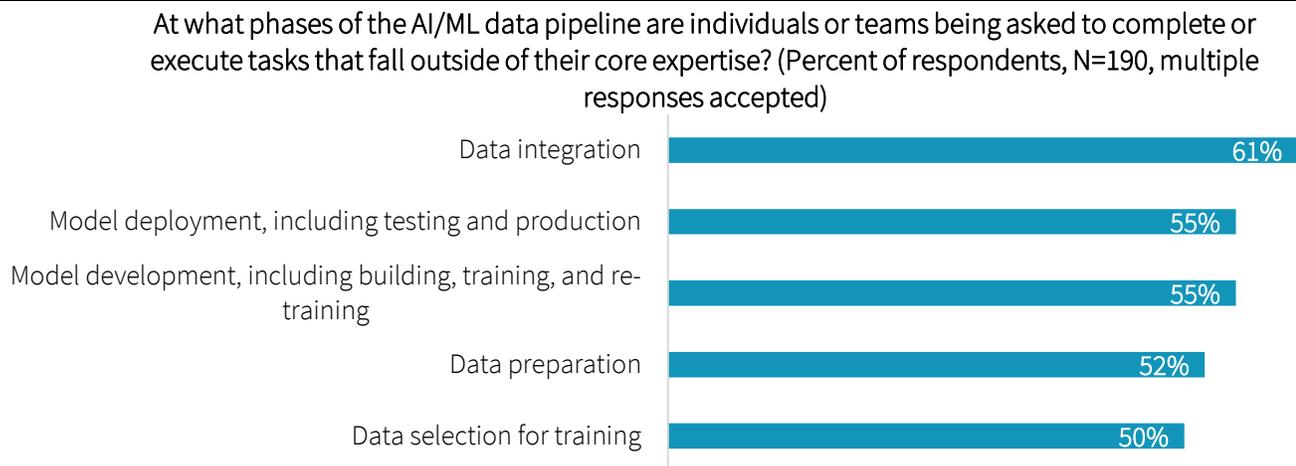
But whether organizations are getting started by leveraging basic learning algorithms through trained data with machine learning or extending initial AI initiatives by embracing multi-layer neural networks with deep learning, they must properly plan to succeed and scale. This means taking the big ideas, breaking them down into realistic goals, and rapidly iterating to gain insights as swiftly as possible. To start, organizations must identify and understand the use cases that matter most, establish an efficient data pipeline, and utilize a performant, reliable infrastructure. To achieve the best possible results, organizational agility is essential—being able to quickly build, iterate, and deploy distributed AI-based applications at scale, while efficiently managing the real-time requirements of the business.

Skills Gaps Across the Data Pipeline

While the data scientist shortage is nothing new, it should be noted that organizations without a data scientist on staff are not deterred in embracing AI—and this situation is driving the need to quickly hire new data scientists and other data-centric staff possessing the right skillsets and experience. Finding the right people with the right skills continues to be a major challenge, and the ramifications of missing personnel across the data pipeline are directly impacting the success of AI initiatives.

Looking across the AI data pipeline, from data integration and data preparation to data selection for training, model development (including building, training, and retraining), and eventually deployment, ESG asked IT and data-focused professionals about where in the data pipeline individuals or teams are asked to complete tasks that fall outside of their core expertise. Based on ESG research, virtually all phases of the data pipeline pose concerns, with at least 50% of enterprise organizations stating they have inappropriately tasked or overworked personnel at each phase (see Figure 1).[1]

Figure 1. Phases of the Data Pipeline Where Staff Are Executing Tasks Outside of Their Core Expertise

At what phases of the AI/ML data pipeline are individuals or teams being asked to complete or execute tasks that fall outside of their core expertise? (Percent of respondents, N=190, multiple responses accepted)

Phase	Percent
Data integration	61%
Model deployment, including testing and production	55%
Model development, including building, training, and re-training	55%
Data preparation	52%
Data selection for training	50%

Source: Enterprise Strategy Group

---

[1] Source: ESG Master Survey Results, *Artificial Intelligence and Machine Learning: Gauging the Value of Infrastructure*, March 2019. All ESG research references and charts in this white paper have been taken from this master survey results set.

Clearly, this is a two-fold issue: Data scientists are spending more time on non-data scientist tasks (think data integration and data preparation, infrastructure deployment and systems configuration), while other data-centric personnel (i.e., data engineers, data stewards, data architects, data analysts, etc.) are being asked to perform data scientist tasks. Due to this frequently occurring situation, adoption of automated machine learning (AutoML) is on the rise, and organizations are looking to embrace automation for tasks such as training and tuning models as well as deployment and operationalization.

Deployment

Interestingly enough, the data science aspects of AI are not the greatest challenges in the data pipeline, especially with AutoML on the rise. Organizations are finding ubiquitous support when it comes to selecting an algorithm that yields high accuracy depending on the specific use case. According to ESG research, when asked which phase of the AI/ML data pipeline creates the most challenges for their organization, nearly one quarter of organizations cited model deployment, including production and testing.

Why is that? Simply ensuring the right data is utilized and the data is high quality is just the beginning. Organizations must ensure data is securely accessed and incorporated from different silos and sources, and confirm all the data is properly integrated. Then, it's a matter of accounting for the number of tools that must be tightly integrated and compatible, including versions, libraries, and drivers. Lastly, it's time to move from a test environment (using sample data, with complete control) to a large-scale distributed environment across on-premises and public cloud environments. This final step is an extremely important stage of the application lifecycle that is often overlooked.

It's not difficult to understand why this often-repeated scenario is driving organizations to find the best means to simplify the deployment process, increasing agility to empower a DevOps approach to AI/ML. With that in mind, many organizations are turning to containers. In fact, of those organizations that already have AI/ML initiatives underway, 17% currently use containers to some extent, 14% plan to use containers in the next month, and 40% are exploring ways to best leverage containers (see Figure 2).

Figure 2. Containers as Part of AI Initiatives

To what extent is your organization leveraging – or planning to leverage - containers as part of its AI/ML initiatives? (Percent of respondents, N=300)
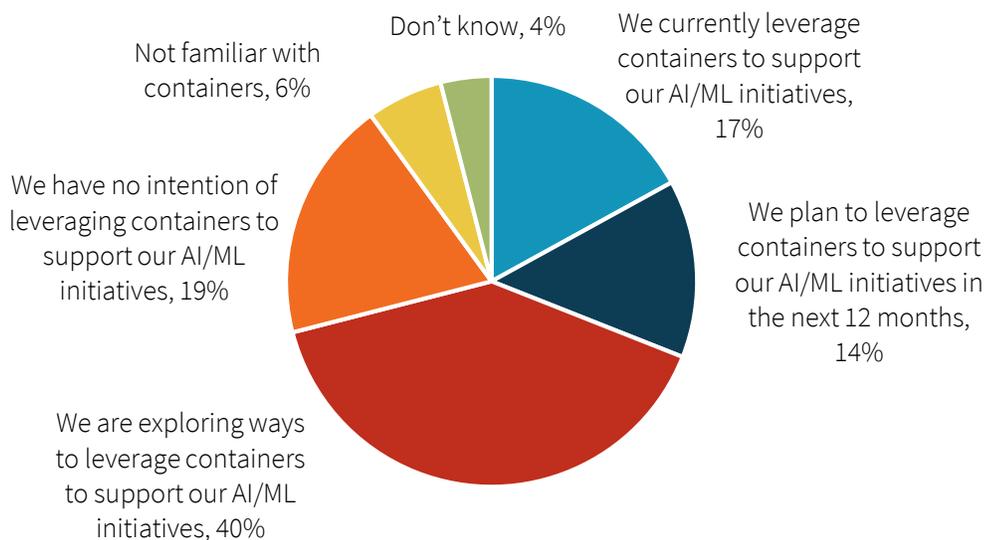
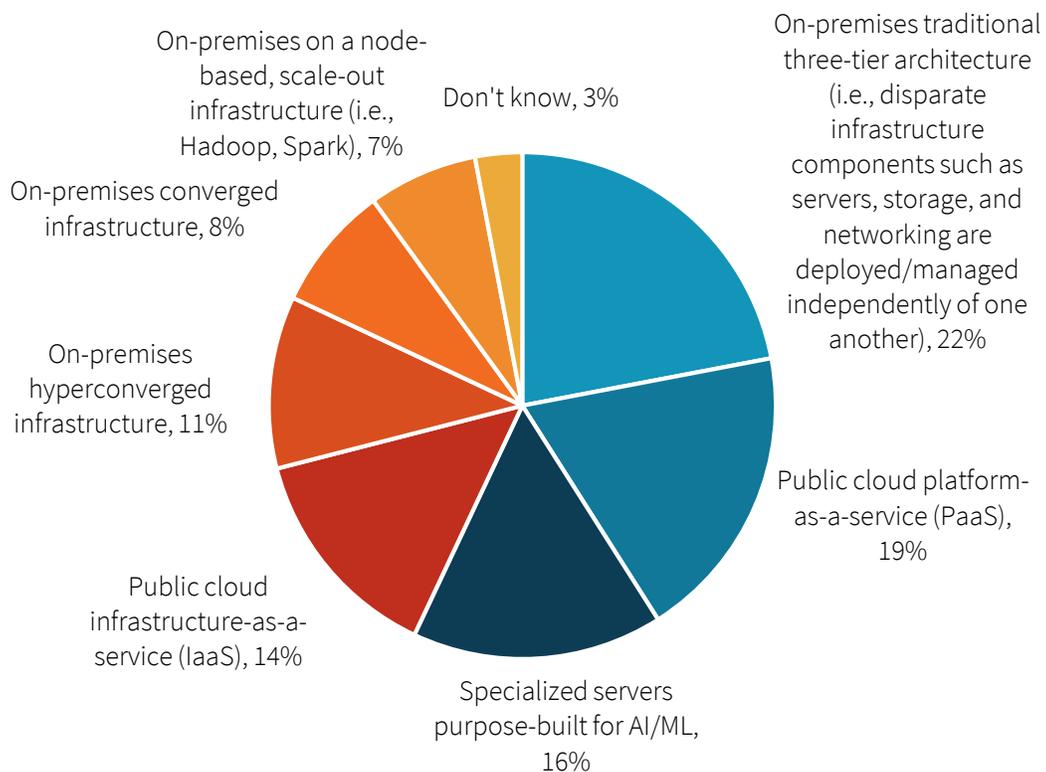*Source: Enterprise Strategy Group*

On-premises and the Cloud

When it comes to the stages of the data pipeline, it's evident that both on-premises infrastructure and the cloud are utilized to support all initiatives. While digital-native companies lean more toward utilizing the cloud for integration and deployment, those organizations that have previously invested in on-premises infrastructure to satisfy data-driven initiatives currently complete their data pipeline tasks close to where the data itself is generated—whether in a core data center or at an edge location. This speaks to the significant impact data gravity has on organizations, illustrating that data is key in determining where tasks are carried out.

While the cloud has been a great launching pad for AI initiatives—especially for learning and experimentation into interconnectivity of AI services and technologies—ESG research shows that organizations are divided in their primary infrastructure deployment model preferences, between on-premises and the public cloud. Figure 3 highlights the primary infrastructure deployment model surveyed organizations currently use or expect to use in the next 12-18 months.

Figure 3. Primary Infrastructure Deployment Model to Support AI Initiatives

What is the primary infrastructure deployment model your organization uses–or expects to use–
to support AI/ML? (Percent of respondents, N=300)

- On-premises on a node-based, scale-out infrastructure (i.e., Hadoop, Spark), 7%
- On-premises converged infrastructure, 8%
- On-premises hyperconverged infrastructure, 11%
- Public cloud infrastructure-as-a-service (IaaS), 14%
- Don't know, 3%
- On-premises traditional three-tier architecture (i.e., disparate infrastructure components such as servers, storage, and networking are deployed/managed independently of one another), 22%
- Public cloud platform-as-a-service (PaaS), 19%
- Specialized servers purpose-built for AI/ML, 16%

Source: Enterprise Strategy Group

The importance of a hybrid cloud strategy should not be understated, especially for large enterprises that already have or will have application and data footprints across environments. While certain applications live in the cloud, and others live on-premises, the data generated and used by these applications will likely follow suit—making application and data portability essential. And containerization can help: their inherent portability enables deployment agility and flexibility across any infrastructure, including on-premises and in the cloud.

Accelerating AI Initiatives with BlueData (recently acquired by Hewlett Packard Enterprise)

BlueData (recently acquired by HPE) provides container-based infrastructure software to simplify the deployment of AI/ML and big data infrastructure and applications on-premises and in the cloud. The software platform uses containers to easily and quickly deploy multiple AI/ML tools, data science notebooks, and other big data analytics applications. This provides an as-a-service solution in an organization's preferred environment (whether on-premises, in the cloud, or both), leading to improved agility, elasticity, and flexibility. HPE's BlueData software enables the self-service deployment of on-demand environments while also satisfying enterprise requirements for AI initiatives, including security and performance.

Enterprises can effectively use containerized clusters sharing the same underlying resources (storage and compute), which are decoupled to enable independent resource scaling depending on specific application and workload requirements. With BlueData's unique technology innovations, organizations can tap into data from any enterprise storage system, thus reducing data duplication and the need to move data.

Through optimal and efficient resource utilization, as well as streamlined management and automation, organizations can recognize substantial capital and operational cost savings. Additionally, by reducing the effort required to deploy distributed AI/ML and analytics, manage the infrastructure, and access data, both technical architects and data science teams can realize productivity gains. In doing so, enterprises can make the most of their data science teams—freeing up their time and helping to address the skills gap in the organization.

IT can easily manage multiple clusters for different use cases, achieve the desired levels of multi-tenancy, and adhere to infrastructure requirements of a specific business unit or project—assuring non-IT personnel, from data analysts to data scientists, that they will always have access to the right level of resources to satisfy their individual requirements. Control and self-service are enhanced by enabling faster spin up of resources within minutes, without having to deal with the delay of involving IT. This scenario then leads to improved operational efficiency, faster time to value and, ultimately, more AI-driven innovation and competitive differentiation for the business.

Faster AI/ML Deployment and Operationalization

To further simplify adoption of AI/ML within enterprises, HPE's BlueData solution enables data scientists and developers to quickly deploy sandbox environments for rapid innovation and then operationalize their models in production deployments. With BlueData, they can instantly spin up multi-node, containerized environments and automate model creation through model deployment with a wide range of different data science tools and AI/ML technologies (i.e., TensorFlow, Spark ML, H2O, Anaconda, Jupyter, etc.).

By standardizing the user experience, consistency and repeatability in pipeline creation and management are provided across all stages of the AI/ML model development and deployment lifecycle. Organizations can simplify and streamline AI/ML model operationalization, while providing security and control—through integration with user authentication (AD/LDAP/SAML) and commonly utilized distributed file systems such as HDFS, NFS, and S3. The multi-tenant architecture can easily scale to support higher levels of end-user accessibility and concurrency.

The BlueData platform includes an App Store with out-of-the-box container-based images for one-click deployment of a wide range of AI/ML, data science, and analytics tools. However, if the data science team's preferred tools are not available out of the box, the extensibility of the solution enables organizations to easily add new AI/ML applications and frameworks as container images.

BlueData and HPE

With the acquisition of BlueData, HPE is looking to accelerate AI/ML deployments for existing and future customers with an innovative container-based software platform. HPE hopes to assist organizations in their AI initiatives with a powerful solution that increases agility, speeds deployment times, and reduces costs along the way.

For organizations looking for a cloud-like experience, BlueData enables an "as-a-service experience" for on-premises and hybrid IT deployments, regardless of the underlying infrastructure. It's important to note that HPE intends to retain the infrastructure-agnostic capability of BlueData to run on any hardware and any public cloud—a wise move, as many organizations place HPE in the role of a trusted advisor for its customers embarking on AI initiatives and a hybrid IT journey. At the same time, HPE is also offering the ability to easily run BlueData on HPE infrastructure, including HPE Apollo systems. And together with HPE Pointnext, HPE and BlueData are providing the education, training, and support required to accelerate AI initiatives and ensure that customers achieve their business goals.

With the hardware leadership of HPE, as well as the HPE Pointnext services organization, and the pay-as-you-go resource consumption model of HPE GreenLake, organizations across all industries can benefit from the combination of BlueData with HPE's AI experience and expertise to ensure the success of their AI initiatives.

The Bigger Truth

With major AI initiatives effecting game-changing impact across every industry around the world, enterprises must begin to focus on investing in AI and incorporating it into their businesses—or they will be left behind. As with anything worthwhile, those investing in this technology have experienced and will continue to experience challenges associated with the underlying infrastructure, the complexity of tools and data pipelines, a lack of skilled staff, and high costs. Regardless of the challenges, the benefits of AI outweigh any shortcomings, and enterprises are fearless in their pursuit of AI.

With their differentiated combination of software, infrastructure, expertise, and services, HPE can help address these challenges for organizations across all industries, enabling them to be successful as they embark on their AI journey and scale their deployments. HPE's container-based BlueData software solution enables organizations to quickly spin up their AI/ML and analytics tools of choice for a variety of use cases, and across the application lifecycle—from sandbox environments to large-scale, distributed environments in production. With containers anchoring the platform, organizations gain the desired DevOps agility in the AI/ML model development and deployment process to jumpstart their AI initiatives, operationalize their AI/ML models, develop new AI-driven innovations, and deliver business value.

Enterprise Strategy Group is an IT analyst, research, validation, and strategy firm that provides actionable insight and intelligence to the global IT community.