

## Solution Brief

# Operationalize Data Science at Scale with NetApp and Iguazio

One-click deployment for all your AI, ML, and DL needs

### Key Benefits

#### Improve Operational and Cost Efficiencies

- Simpler and faster delivery of AI applications to production.
- Unify all your data science workloads—analytics, AI, ML, DL, and APIs—on one self-service platform.
- Open set of tools and frameworks.

#### Maintain Your Competitive Edge

- Reduce time to develop and deploy new AI services by 6 to 12 times.
- Meet the most extreme workload demands with the ability to automatically scale compute and capacity resources as needed.
- Focus your team on the business apps, not the infrastructure.

#### Harness the Power of the Cloud

- Rapidly scale computational resources with enterprise multicloud bursting capabilities.
- Give AI and analytics applications unfettered access to data, no matter where it resides.
- Future-proof your AI infrastructure with an integrated and managed platform that enables you to continue using all your familiar open source tools—as well as any tools that are developed in the future.

#### Are You Struggling to Deliver on Your AI and ML Initiatives?

It's no secret—the data science market is exploding. IDC forecasts that spending on artificial intelligence (AI) and machine learning (ML) will reach \$57.6B by 2021. Yet, according to leading analysts, 87% of AI and ML projects fail to deliver on their goals and get to production.

In many cases, Hadoop with Spark is used to build data pipelines, which involves recoding and manual labor on an outdated architecture. Hadoop environments can be expensive, complex, and difficult to scale. They often deliver poor utilization, which means that you may be paying for more compute resources than you use, or you may waste storage on duplicate data. Hadoop doesn't fit online applications, which forces you to add Kubernetes clusters alongside to run additional microservices and modern ML frameworks.

#### Unleash Your Data with Iguazio and NetApp

The Iguazio Data Science Platform, combined with NetApp® ONTAP® AI or FlexPod AI, delivers the performance and scalability to support your most extreme workloads. Iguazio offers native integration with NetApp ONTAP solutions to enable building highly available end-to-end ML pipelines that can process and react to massive amounts of data. In addition, NetApp StorageGRID, combined with ONTAP solutions, brings in the power of S3-connected AI tools and frameworks, further empowering Iguazio's capabilities and enriching ecosystem tools.

The solution enables the full data analytics and AI and ML lifecycle on one ready-to-use platform. Users can ingest data from a variety of sources, explore and prepare data at scale, run distributed training, and automatically deploy models or complex applications to production. Data scientists, data engineers, DevOps, and app developers can collaborate efficiently on one secure self-service platform.

#### Open and Integrated Architecture

The NetApp and Iguazio joint solution offers the leading data-science tools in a managed and easy to use package, which includes:

- Managed Jupyter Notebooks
- Managed data-science and machine learning operations services over Kubernetes (Spark, Presto, KubeFlow, etc.)
- Serverless functions and jobs, using Nuclio and MLRun
- Data and ML pipeline automation with built-in job and artifact tracking
- Automated model deployment and serving
- GPU as a service and out-of-the-box GPU acceleration for analytics, AI, ML, and deep learning (DL) workloads

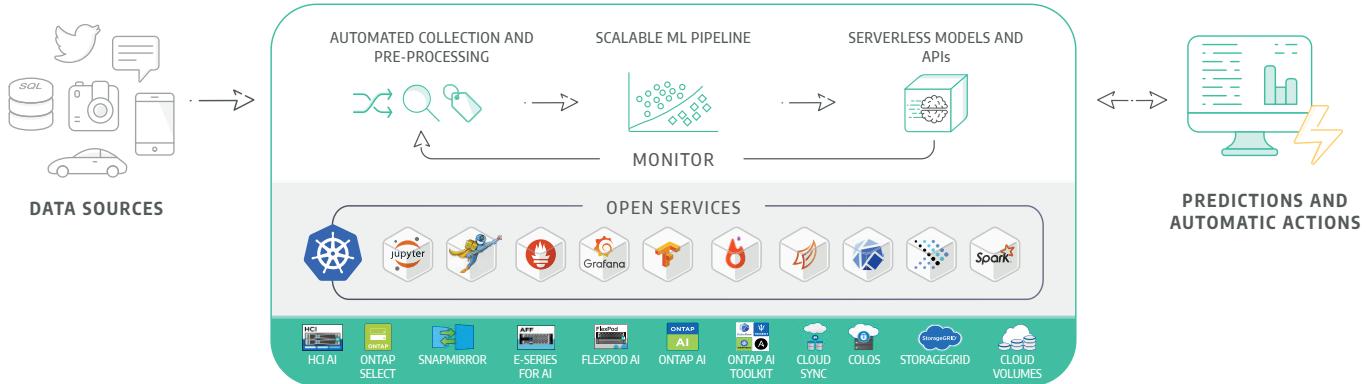


Figure 1) Iguazio data science platform.

### Improve Operational Costs and Efficiencies

With the unified solution from NetApp and Iguazio, you can save 40% to 50% in infrastructure costs, with 16 times the storage capacity reduction and 3 times fewer compute nodes. ONTAP data management features, such as data reduction, tiering, and NetApp Snapshot™ copies drive huge infrastructure efficiencies. Automated, serverless microservices reduce costs and accelerate deployment, so you only pay for what you use. Separate compute and capacity allow you to right-size the resources for your data lake, eliminating overprovisioning and underutilization.

### Maintain Your Competitive Edge

Self-service and one-click deployment features help you reduce time to develop and deploy new AI services by 6 to 12 times. And a high-performance, multimodel data engine for active data means that you can process more data, faster.

Using improved resource management features via Kubernetes, you can also achieve elastic scalability to make sure that your training and inference workloads always get the performance they need, when they need it. You can even set your resources to scale automatically, so you don't need to spend time or resources managing your infrastructure.

### Harness the Power of the Cloud

Modern AI and ML workloads require the utmost flexibility and speed. They also need data—the more, the better. For the best results, you need a solution that can unify your data pipeline across edge, core, and cloud.

With NetApp and Iguazio's joint solution, you can rapidly scale compute resources with enterprise multicloud bursting capabilities in an infrastructure-agnostic configuration, so you don't need to worry about data lock-in. You can even repurpose existing compute nodes for bursting. By enabling an integrated data science pipeline built on an open platform, NetApp and Iguazio provide you with a gradual path to modernization.

### About Iguazio

The Iguazio Data Science Platform enables enterprises to develop, deploy and manage AI applications at scale. With Iguazio, enterprises can run AI models in real time, deploy them anywhere (multi-cloud, on-prem or edge), and bring to life their most ambitious AI-driven strategies. Enterprises spanning a wide range of verticals, including financial services, manufacturing, smart mobility and telecoms use Iguazio to create business impact through a multitude of real-time use cases such as fraud prevention, self-healing networks and location-based recommendations. Iguazio brings data science to life. Find out more on [www.iguazio.com](http://www.iguazio.com)

### About NetApp

NetApp is the data authority for hybrid cloud. We provide a full range of hybrid cloud data services that simplify management of applications and data across cloud and on-premises environments to accelerate digital transformation. Together with our partners, we empower global organizations to unleash the full potential of their data to expand customer touchpoints, foster greater innovation and optimize their operations. For more information, visit [www.netapp.com](http://www.netapp.com). #DataDriven