SOLUTION BRIEF

The training data solution for production ML

Manage, annotate, and iterate on training data for data scientists and machine-learning teams

NetApp | Labelbox

The challenge

Production AI teams spend roughly 80% of their time building and maintaining training data infrastructure. Data science teams shouldn't have to build their own expensive and incomplete tools for managing this data. They need a platform that acts as a central hub for creating and managing training data with internal or external labeling teams. Better ways to input and manage data result in higher-quality training data and more accurate ML models. Teams need to be able to seamlessly manage, annotate, and iterate training data for production AI. ML engineers and labelers need a fast, powerful, and intuitive solution that gives them full visibility into the real-time operations of labelers and the quality and accuracy of labels.

The solution

NetApp and Labelbox have partnered to deliver an integrated training data solution that is streamlined and creates new productivity and efficiency metrics. Data scientists rely heavily on iterating on training data, but they need a central place to store and house all of an organization's training data. With the NetApp® ONTAP® AI proven architecture, you can fully realize the promise of AI and deep learning (DL) by simplifying, accelerating, and integrating your data pipeline. With Labelbox, the same datasets can be reused multiple times with less effort. Using a training data solution has the following business advantages:

- **Productivity gains.** Feature-level analytics and a streamlined design enable faster iteration cycles.
- **Sharing and collaboration.** Advanced workflows across distributed labeling teams and automated task distribution, team provisioning, and dataset management increase collaboration across business and technical teams.
- **Cost savings.** Companies save money by not having to build in-house labeling systems that are brittle, hard to maintain, and often lack the required features to scale.

Key Benefits

Training data infrastructure fused together for fast labeling and iteration by ML teams

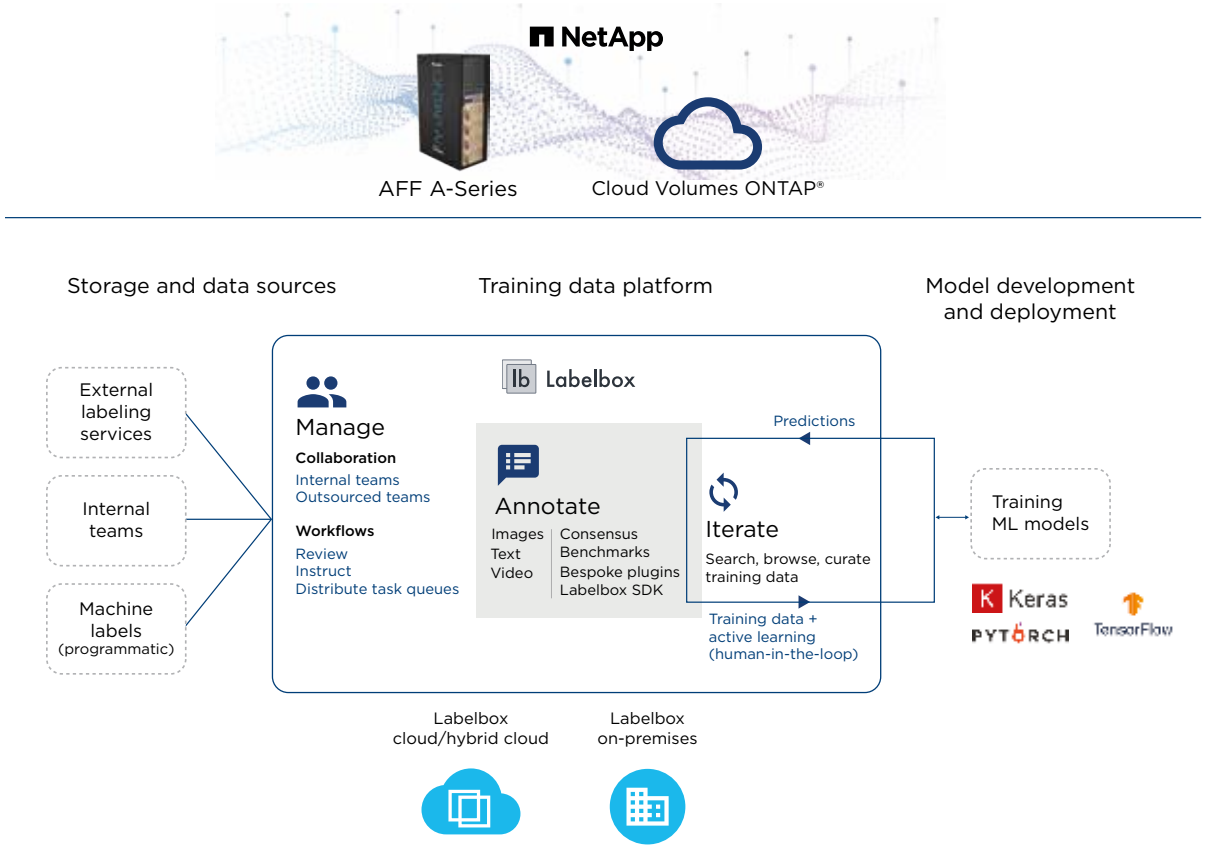
- Use high-quality annotation tools that support complex annotations; nested classifications; and image, video, and text datasets.
- Search, browse, and curate all of your training data in one place.
- Stream data into Labelbox and push labeled data into training environments.

Industry-leading solution for ML training data team collaboration

- Create high-quality training data by using an optimal and organized workflow.
- Prevents your data science and labeling team from 'flying blind' by allowing seamless review workflows.
- Works both in the cloud and on-prem.

Universal ontology management and training data curation

- Set up accurate, universal, and nimble ontologies to create performant AI models.
- Reuse classes organized by global ontologies.
- Projects using the same ontology will reference the same schema, so you can run queries across projects and make programmatic edits.



NetApp and Labelbox training data solution for production ML.

About Labelbox

Dramatically reduce the costs and time it takes to create high-quality training data for your computer vision projects. Labelbox is the leading training data platform for AI applications. We help teams annotate, manage, and iterate on labeled data at scale so that they can focus on building what their customers want. For more information, visit www.labelbox.com.

About NetApp

In a world full of generalists, NetApp is a specialist. We're focused on one thing, helping your business get the most out of your data. NetApp brings the enterprise-grade data services you rely on into the cloud, and the simple flexibility of cloud into the data center. Our industry-leading solutions work across diverse customer environments and the world's biggest public clouds.

As a cloud-led, data-centric software company, only NetApp can help build your unique data fabric, simplify and connect your cloud, and securely deliver the right data, services and applications to the right people—anytime, anywhere.

NetApp