

Pachyderm: Enterprise

Engineered to make data science

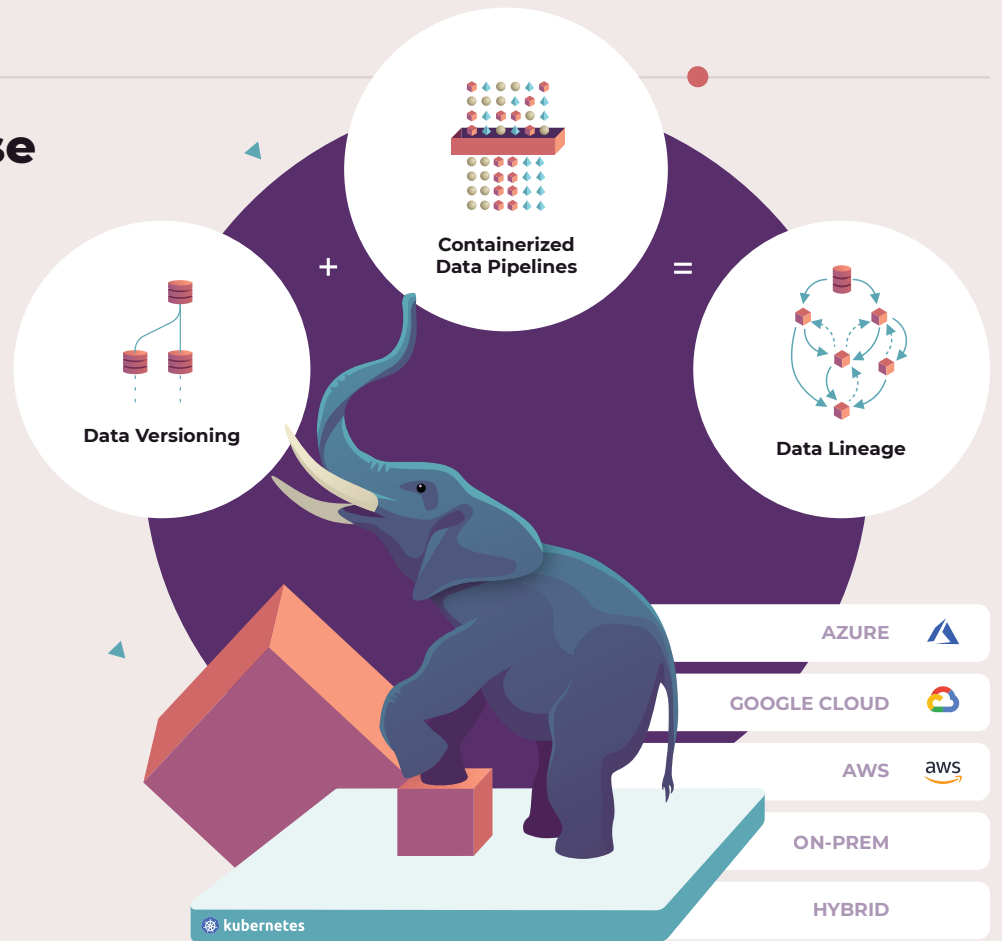
**Explainable.
Repeatable.
Scalable.**

The unspoken truth about real-world data science.

Instead of spending time innovating and using data science as a force-multiplier to the business, most data science teams are stuck between disparate systems, juggling ad-hoc scripts, and hacking their way through complex pipeline jungles.

How Pachyderm can help

Pachyderm is an enterprise-grade, open source data science platform that makes explainable, repeatable, and scalable Machine Learning (ML) and Artificial Intelligence (AI) a reality.



Data Versioning

+

A robust data versioning engine powers the heart of the Pachyderm platform. Every change to your data gets tracked and managed with ease. Adding, removing and modifying files all get meticulously captured and recorded by the Pachyderm engine automatically. With git-like familiarity, teams of data scientists can swiftly track the dependencies and relationships between datasets, their models, and everything in between with confidence. Pachyderm is the first and only data science platform that can version virtually any type of data that exists: structured, unstructured, and even streaming data. It's fast and lightweight too, allowing you to version data all the way up to the petabyte sale and beyond.

Containerized Data Pipelines

=

Pachyderm enables you to build end-to-end data pipelines using any language or framework you want. Together with data versioning, Pachyderm Pipelines intelligently process only changed data, while also efficiently distributing the computation across the cluster for faster results.

True Data Lineage

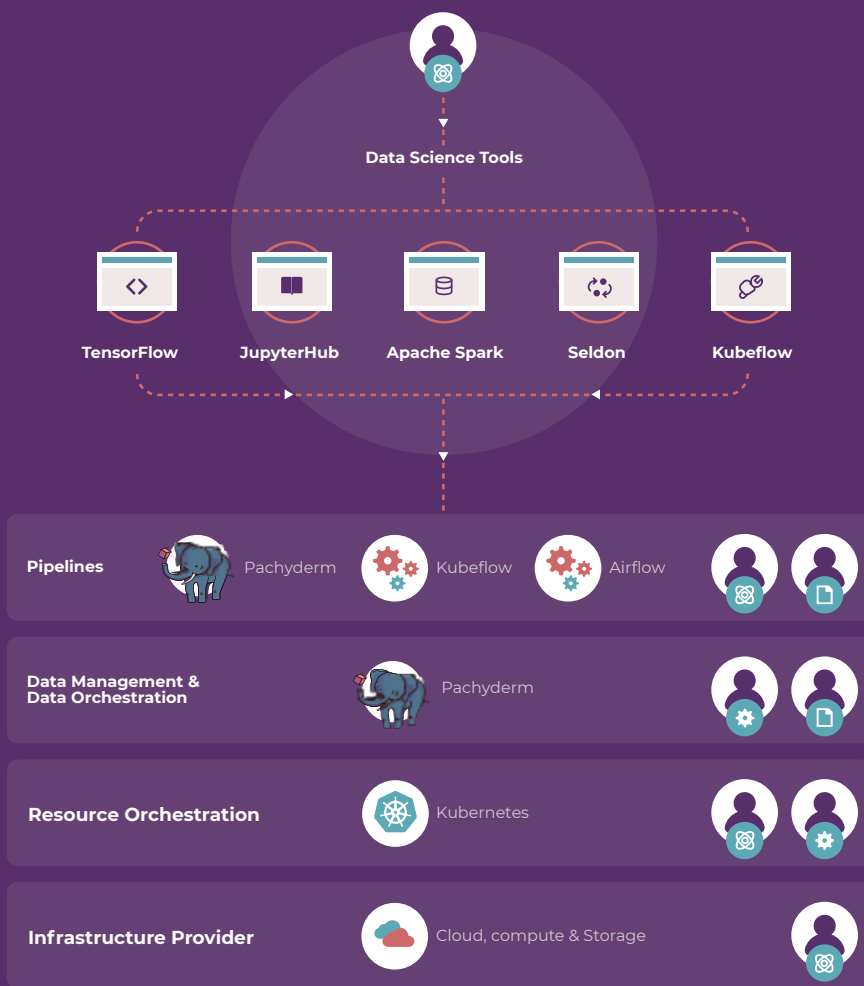
Pachyderm provides you with true data lineage. Our solution helps you capture the entire journey of your data, code, models, and results in one unified system. With that power at your finger tips, your data team can reason about the complex relationships between them and drive business decisions efficiently.

“Pachyderm reduced our data processing time from 7 weeks to **just 7 hours**”

Eyal Heldenberg
Voice AI Product Manager
LogMeIn AI Center of Excellence

“Pachyderm helped us convert our existing data science pipelines from manually managed scripts to scalable, repeatable end-to-end workflows”

Mauricio Borgen
Director of IT & Scientific Compute
AgBiome



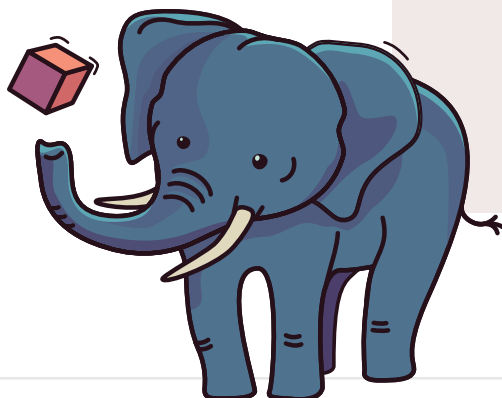
Scalability

Pachyderm on Azure gives you the freedom to build, manage, and deploy end-to-end data science workloads on a massive, global network using your preferred tools and frameworks. Built on Kubernetes (K8s) using docker containers, Pachyderm allows for fully automated and scalable workflows while remaining totally agnostic to your choice of tools or hardware.

It doesn't matter if you're developing mission-critical ML models on your laptop or on a super-cluster, Pachyderm can provide your data science workloads with an explainable, repeatable path to production at any scale.

Book your own personalized demo of Pachyderm: Enterprise today at:

pachyderm.com/request-a-demo/



Pachyderm In the Wild

Pachyderm lets you and your team deploy and manage multi-stage, language-agnostic, data pipelines while maintaining complete reproducibility. Make data and results more accessible for your organization without sacrificing functionality or security. Pachyderm is the ideal platform for anyone looking to make machine learning successful in their organization.

Pachyderm: Enterprise

Pachyderm Enterprise is our fully featured data science platform designed for large-scale collaboration in highly secure environments. Trusted by global Fortune 100 companies across the globe, Pachyderm has been proven to solve real-world data science problems regardless of their size or complexity.

Enterprise Features:

- World-class Support
- Pachyderm Dashboard
- Advanced Statistics
- TLS Encryption
- User Access Controls
- Custom Deployments
- Hands-On Training
- JupyterHub Integration