PURESTORAGE®

Accelerate Data for AI Models

Victor Alberto Olmedo

Principal Analytics and AI FSA
Pure Storage

@purestorage

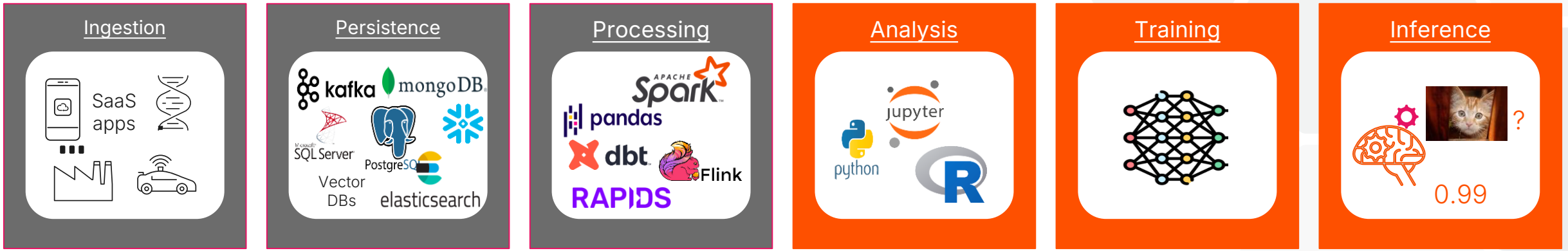
Dec 2024

Data Pipelines Source AI Pipelines

The view inside a typical AI pipeline

Data curation

AI training and inference



Sequential
Writes
Small & Large Files
Varying Concurrency

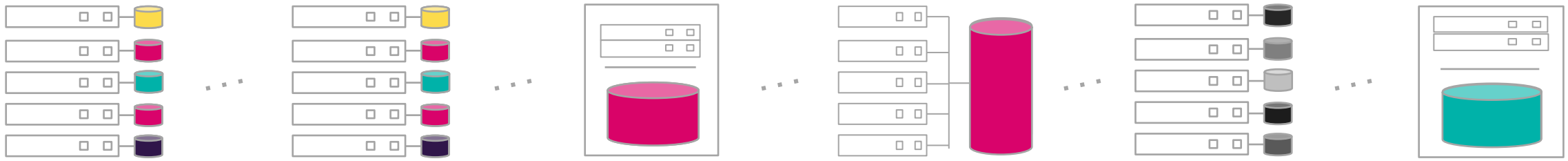
Sequential
Writes
Small & Large Files
Varying Concurrency

Sequential & Random
Reads & Writes
Small & Large Files
High Concurrency

Random Access
Reads
Small & Large Files
Low Concurrency

Random Access
Reads
Small Files
High Concurrency

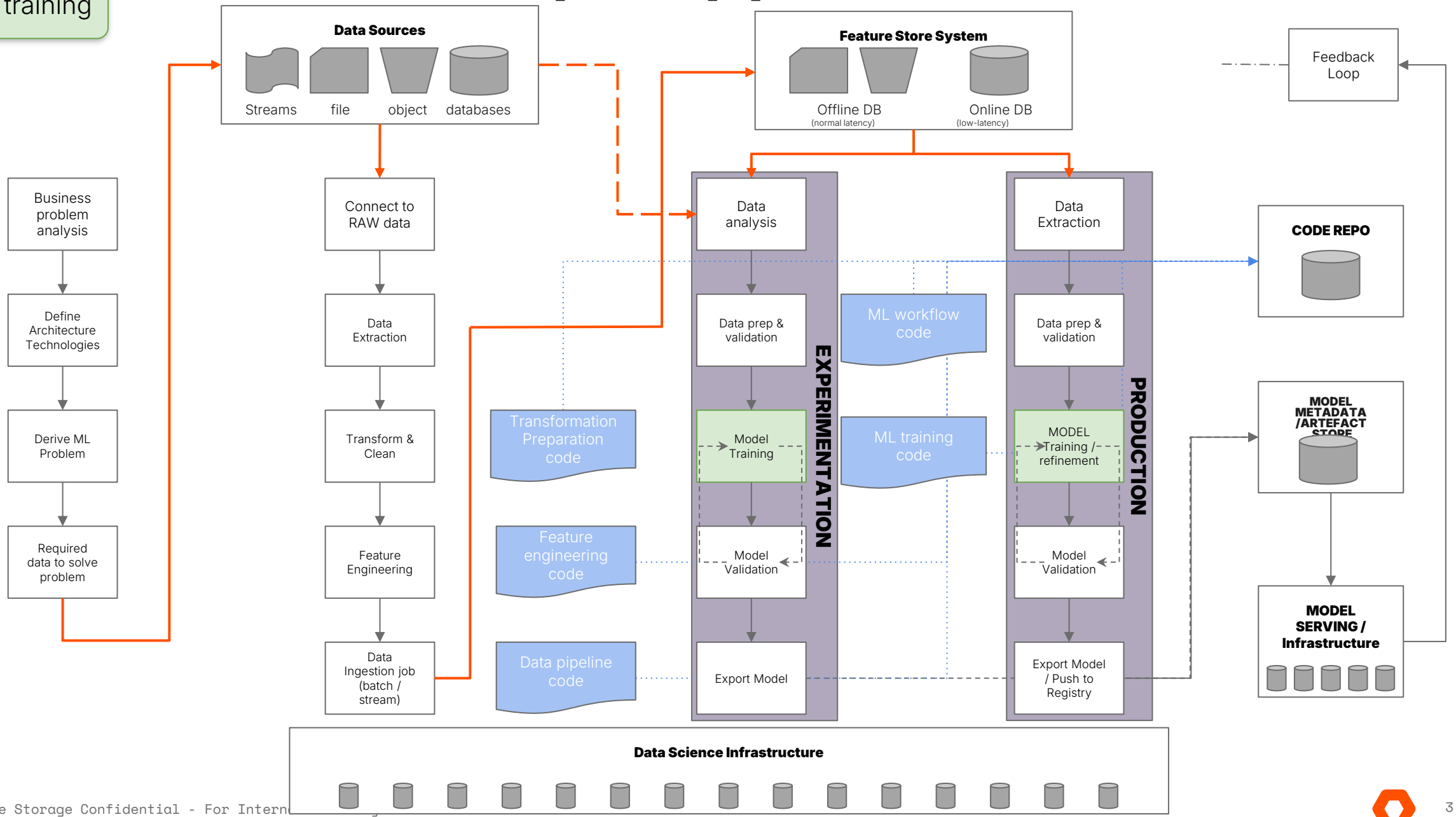
Sequential
Reads & Writes
Small & Large Files
High Concurrency



©2024 Pure Storage Confidential - For Internal Use Only

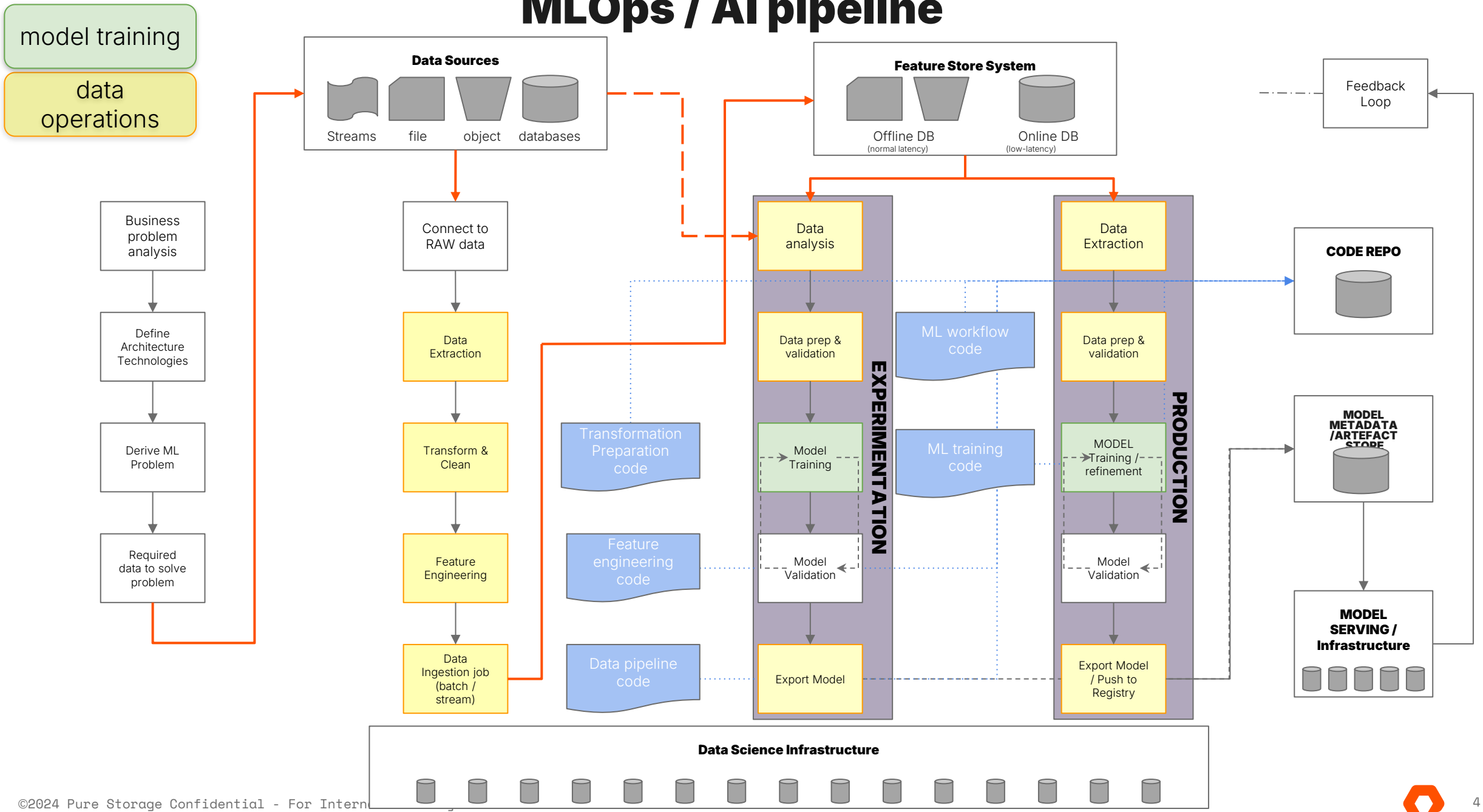
MLOps / AI pipeline

model training



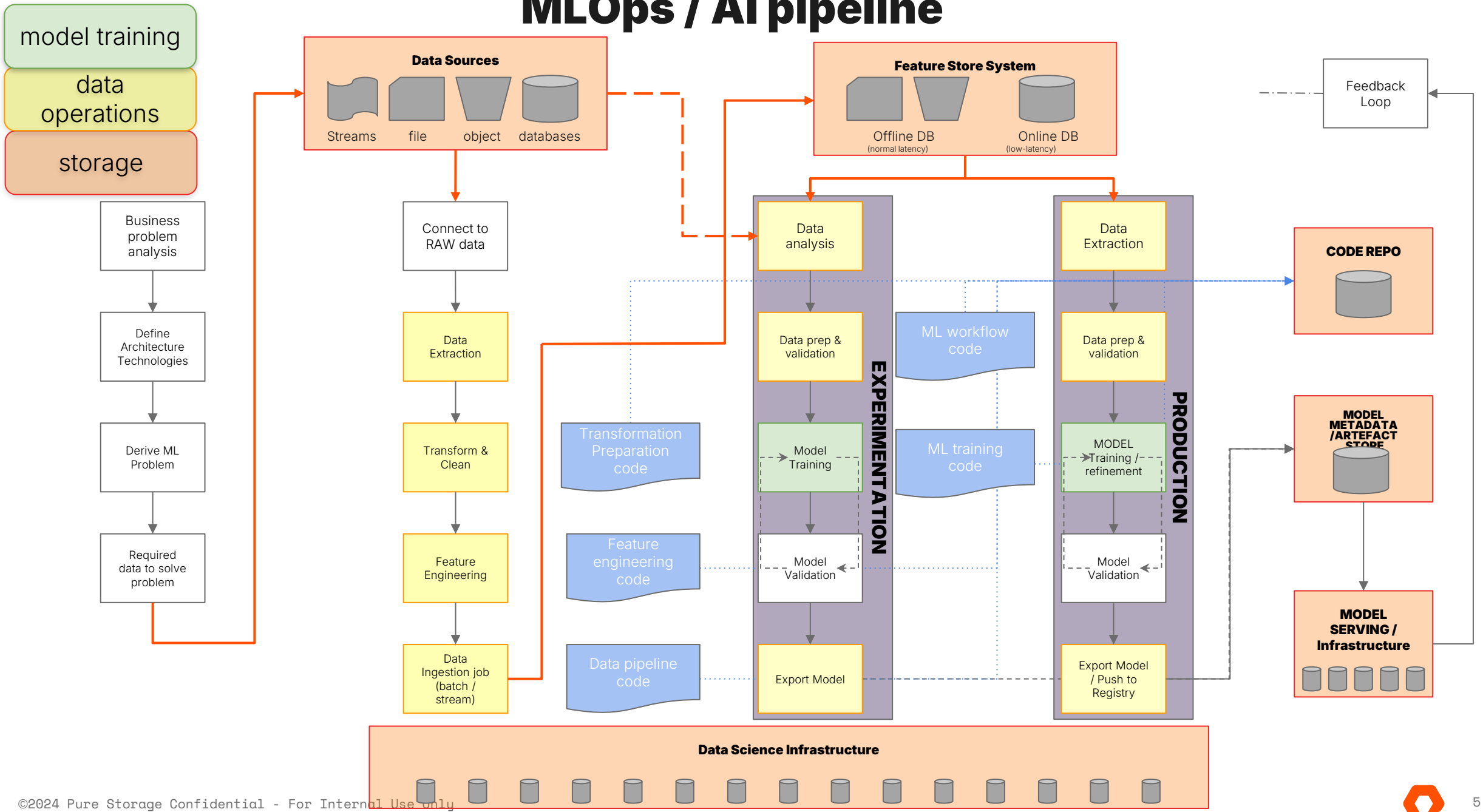
©2024 Pure Storage Confidential - For Intern

MLOps / AI pipeline



©2024 Pure Storage Confidential - For Intern

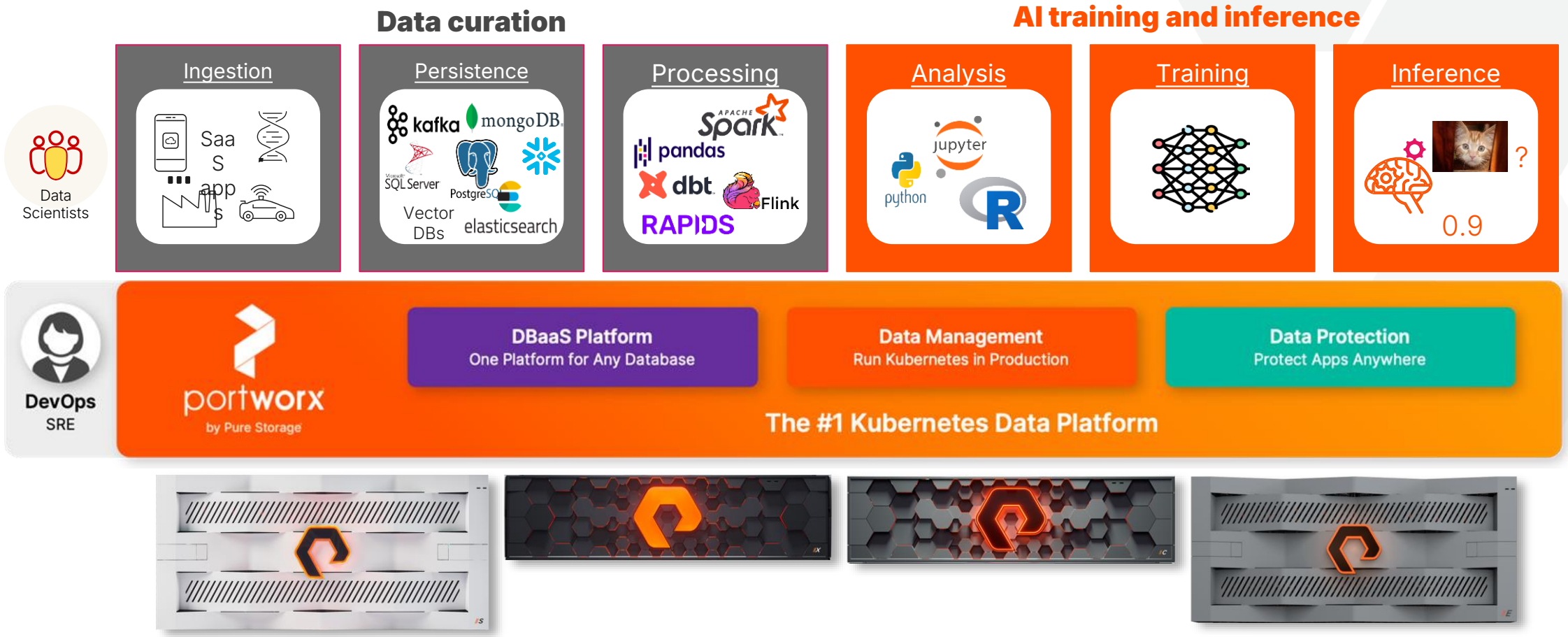
MLOps / AI pipeline



©2024 Pure Storage Confidential - For Internal Use only

Accelerate and Unify Data

Single Storage Platform for Unstructured Data



©2024 Pure Storage Confidential - For Internal Use Only