

# Gather AI Boosts Developer Velocity with XetHub



## GATHER AI

“ We rely on computer vision and ML to deliver on Gather AI’s mission. XetHub has enabled our ML team to be more productive, collaborate efficiently, and iterate quickly.”

**Daniel Maturana, co-founder and Chief ML Scientist**

### INDUSTRY

Logistics

### LOCATION

Pittsburgh, PA

### APPLICATION

MLOps model development and deployment

Gather AI is the world's leading drone-powered inventory monitoring solution for modern warehouses. Cutting-edge warehouses use Gather AI's services to decrease the cost of inventory accuracy, improve warehouse productivity and revenue.

## THE CHALLENGE

### Brittle workflows create opportunities for improvement

Gather AI drones fly autonomously through a warehouse taking images of pallet locations up to 15x faster than traditional means. The images are then analyzed by a machine learning (ML) algorithm, helping warehouse operators identify and fix inventory errors.

The ML engineering team at Gather AI needed to increase their iterative model development cadence. Like many developer teams, they were frustrated by the need to coordinate multiple tools to do their work, constantly switching between EBS for image management, DVC for model management, and Git LFS for metadata management.

The DVC and Git LFS requirement of manually specifying tracked files added a brittle and error-prone step in an otherwise agile process, resulting in a drag on productivity and accuracy.

Gather AI wanted to invest in a sophisticated model management process that integrated with their code development workflows, without requiring additional management overhead.

## THE SOLUTION

### Git-integrated MLOps enables efficient iteration

Gather AI was able to leverage XetHub as a drop-in replacement for Git LFS and DVC in their MLOps workflow, improving the team's performance, efficiency, and ease of use. By using XetHub to manage code, models, and metadata in lock-step, Gather AI eliminated manual offline processes and sped up time-to-deployment by 40%.

XetHub's automatic block-level deduplication also reduced Gather AI's stored repository size, making it faster than ever to update new models on the server.

## DETAILS

### Streamlined workflows lead to natural development loops

For each iteration of their autonomous inventory models, Gather AI uses labeled drone image data to train models on-premises prior to pushing the model to the cloud for deployment.

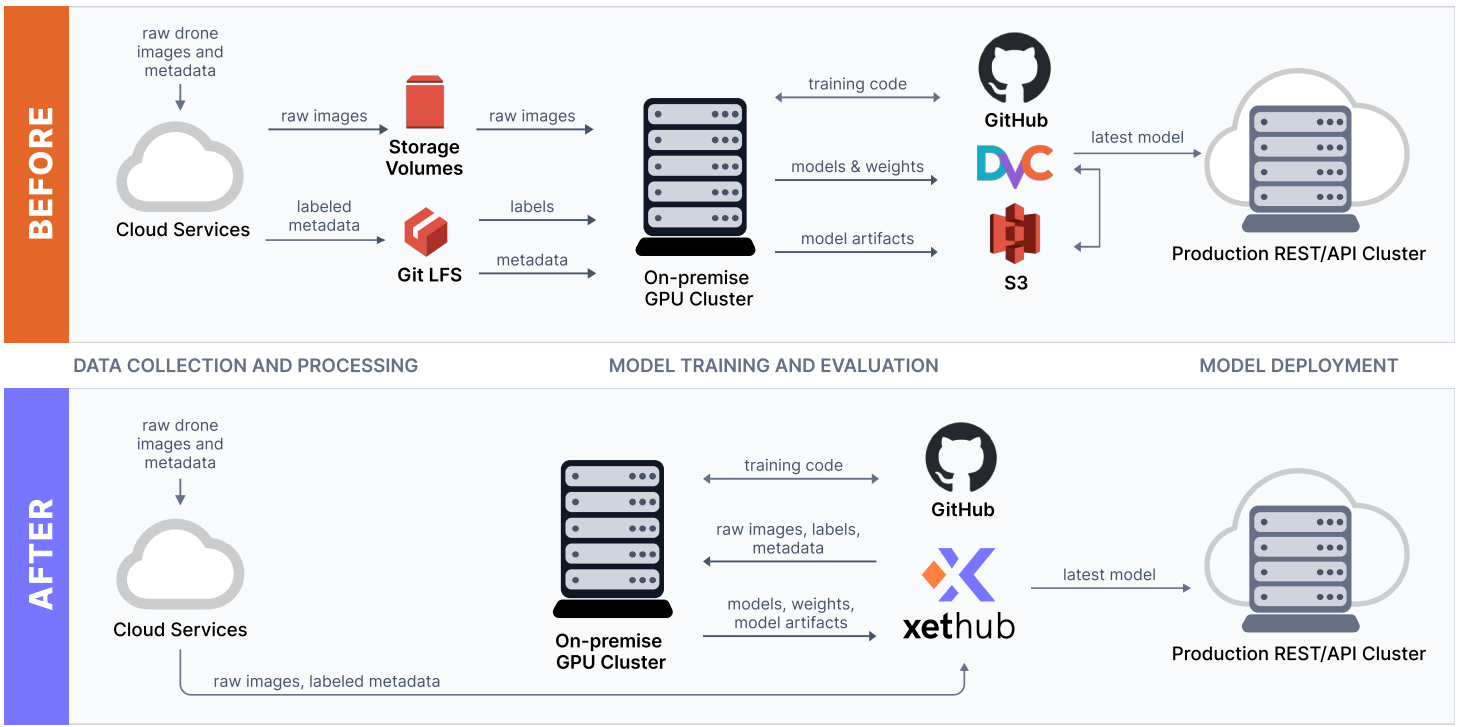
With their original workflow, collected drone images would be stored on cloud storage volumes (EBS and Azure Storage), metadata (including labels) in a Git LFS repo, code in GitHub, and DVC-tracked model artifacts in AWS S3.

By replacing DVC and Git LFS with XetHub for metadata and model management, Gather AI was able to intuitively store data alongside metadata and models alongside model weights for lock-step development, efficient deployment, and built-in provenance.

**40%** reduction in repository size and transfer time

**4** data silos eliminated by switching to XetHub

**51%** cost savings over using EBS, Git LFS, and DVC



Before and after: XetHub replaces cloud storage volumes, Git LFS, DVC, and S3 for a simplified workflow with better performance and lower costs.

XetHub offers a better user experience compared to using DVC or Git LFS to track changes. With XetHub, all files in a repository are tracked by default, removing the manual step of specifying each file to watch. The Git-compatible interface ensures that ML engineers can work in an environment they're familiar with, eliminating the need to learn a new tool.

## RESULTS

### Faster download/upload of new models

Gather AI's models repository is now 40% smaller when stored in XetHub due to block-level deduplication. Smaller repositories means faster deployment as new models are pulled from the model repository by production machines, directly translating into saved time for Gather AI's engineers.

### Easier MLOps deployment

Once the ML team commits and pushes its latest model to their XetHub repository, Gather AI's production application can pull the latest model directly without the need for manual and repetitive offline tasks. This results in faster deployments and simplified operations.

### Smoother user experience compared to DVC

Automatic tracking of all files in a repository reduced Gather AI's manual steps for a more intuitive workflow, while Git's rich change history provides guaranteed reproducibility.

### Seamless collaboration

XetHub allows Gather AI's engineering team to leverage Git workflows (pull requests, branches, forks) for working with data, metadata, and code. All without the need to learn any new commands or tools.

### More efficient and affordable

Gather AI saves money by replacing EBS, Git LFS, and DVC with XetHub, with no overhead needed to track files across systems. Today, using XetHub, Gather AI saves 51% and as their usage grows 10x, they are projected to save 70%.

“As we performed our technical evaluation of XetHub, we found that it scaled well as our repo sizes got larger. It was easy to adopt and required almost no training for the engineers on the team.

The usage-based pricing model makes it easy to align our costs with system utilization, unlike some other models based on team size.”



**DANIEL MATURANA**  
CO-FOUNDER & CHIEF ML SCIENTIST  
GATHER AI