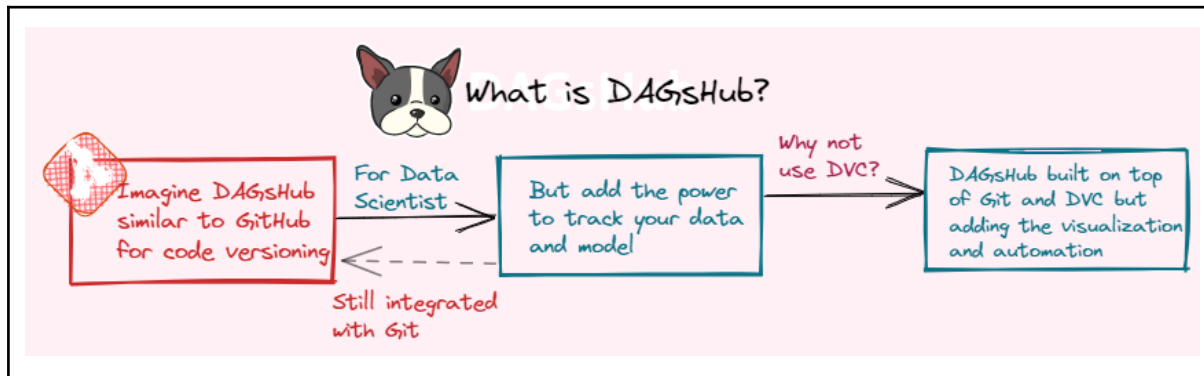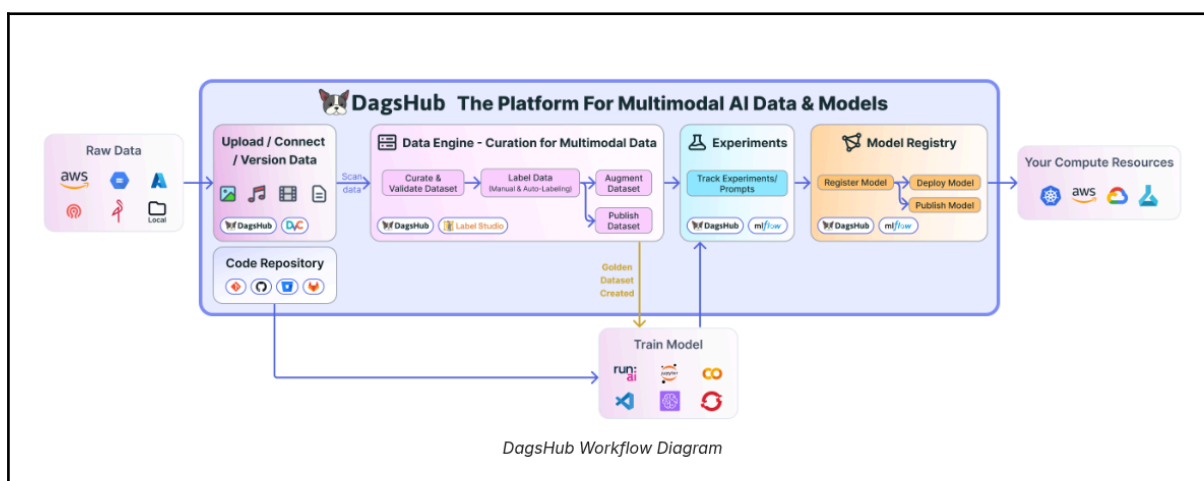# DagsHub

---

DagsHub is an AI platform for version control and collaboration on data science and machine learning projects. It combines Git (code), DVC (data/models), and MLflow (experiments) in one place to make ML projects reproducible , trackable, and collaborative.

DagsHub key features include:
1. Git repository hosting (similar to GitHub/GitLab).
2. Data & model versioning using DVC (large files handled efficiently).
3. Experiment tracking via built-in MLflow UI.
4. Dataset and model storage with remote backends.
5. Visualization of metrics, plots, and data diffs.
6. Access control for teams and private projects.

DagsHub Workflow Diagram

The above graph is a simplified ML/AI workflow diagram and where DagsHub fits in. DagsHub integrates with our storage and compute providers. We can host our data on

the platform with DagsHub Storage. DagsHub doesn't currently provide compute resources, so we should rely on the compute infrastructure we have, which ranges from running training and deploying locally in the cloud or on edge devices.