DagsHub

DagsHub: Empowering Collaboration in Machine Learning ⚡
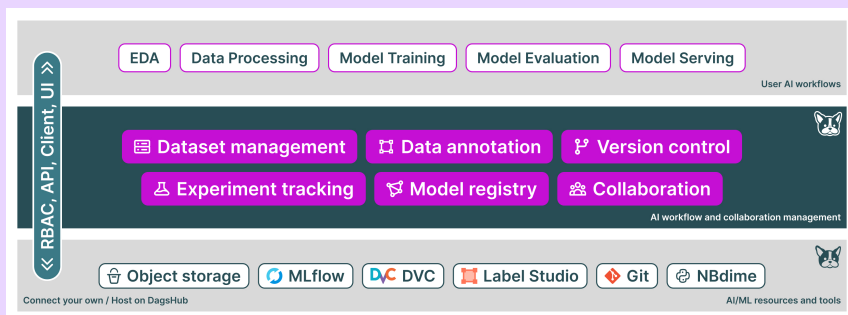
38,000 +
Users

26,000 +
Projects

750+
Organizations

DagsHub is a pioneering platform in the realm of data science and machine learning collaboration. Tailored specifically to meet the demands of ML practitioners and teams, DagsHub simplifies and enhances the way teams manage and share their machine learning models, data, and experiments.

What is DagsHub?

DagsHub is a fully reproducible single source of truth for ML projects, integrating code, data, models, pipelines, experiments, and annotations — into a unified platform. This integration not only streamlines the workflow but also enhances the reproducibility and manageability of projects.

Unique Features of DagsHub

DagsHub is packed with features designed to address the multifaceted challenges of building machine learning models and managing data:

Model Management & Sharing

Enhances collaboration by enabling the sharing and version control of models, including integration with Hugging Face and other popular ML frameworks.

Data Annotation

Supports the annotation of various data types, including images, text, audio, and video, directly within the platform.

Experiment Tracking Server

An MLflow-compatible server is included with each project, allowing for efficient tracking of experiments with zero DevOps.

Dataset Management

Offers tools for filtering and visualizing datasets, facilitating better understanding and management of data.

Data Versioning

Enables precise tracking and control over dataset versions, ensuring that every change is recorded and retrievable.

Built-in Integrations with Other Parts of the ML Stack

Seamlessly integrates with Git servers, storage solutions, Jenkins, and other essential tools, providing a cohesive environment for project management.

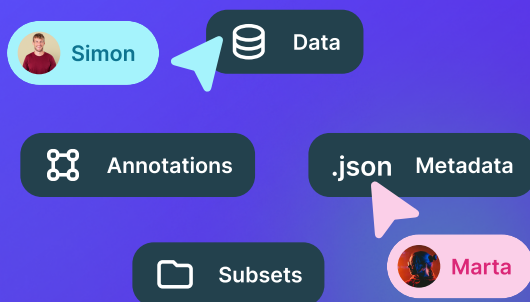Contact us: sales@dagshub.com
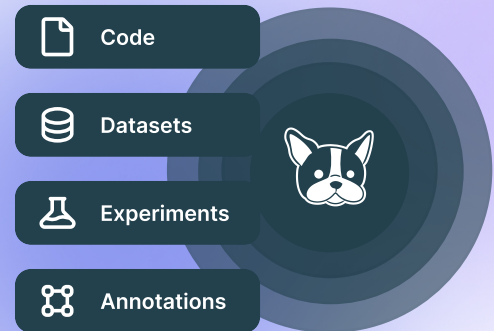
Visit our website: dagshub.com

DagsHub

DagsHub Advantages

Centralized Management of ML Project Components

The platform shines as a single place for managing all aspects of ML projects — data, code, experiments, and models. This centralized approach simplifies project oversight and enhances team productivity by providing a unified view of project components. It streamlines workflows and ensures that all team members have access to the latest project elements, enabling better coordination and faster progress.

Code

Datasets

Experiments

Annotations

Simon

Data

Annotations

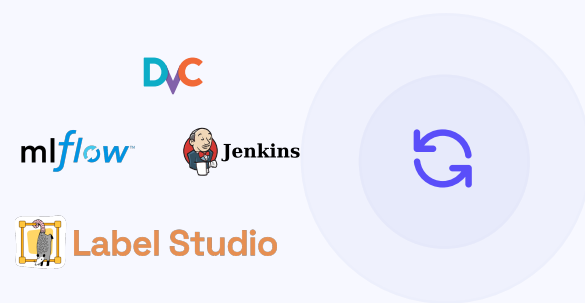.json Metadata

Subsets

Marta

Enhanced Collaboration and Sharing

DagsHub amplifies collaborative efforts by offering a centralized hub for models and datasets, that can be easily shared throughout the organization. This not only streamlines the discovery and reuse of resources but also significantly speeds up the innovation process by making it easier for teams to build upon each other's work.

Integrating Leading Open-Source Tools

DagsHub seamlessly combines leading open-source tools in the ML ecosystem, such as MLflow for experiment tracking, Label Studio for data annotation, and Data Version Control (DVC) for data management. This approach allows users to leverage the strengths of each tool within a cohesive environment without having to deal with the complexity and DevOps of setting up and managing each tool independently. By combining these tools, DagsHub provides a flexible and robust platform that adapts to a wide range of use cases and data types.

DVC

mlflow

Jenkins

Label Studio

DagsHub stands out as a specialized platform that adeptly meets the needs of the data science and machine learning teams. By offering a comprehensive suite of features that enhance collaboration, manageability, and efficiency, DagsHub is positioned as a crucial tool for anyone involved in ML projects. As the platform continues to evolve and expand, it reinforces its commitment to advancing the field of data science through innovation and collaboration.