

# Introduction to Ray AI Libraries

Emmy Li, Balaji Veeramani, Yunxuan Xiao





# **Welcome!**

**We're happy to have you here.**





# Meet the team!



Emmy



Balaji

TA



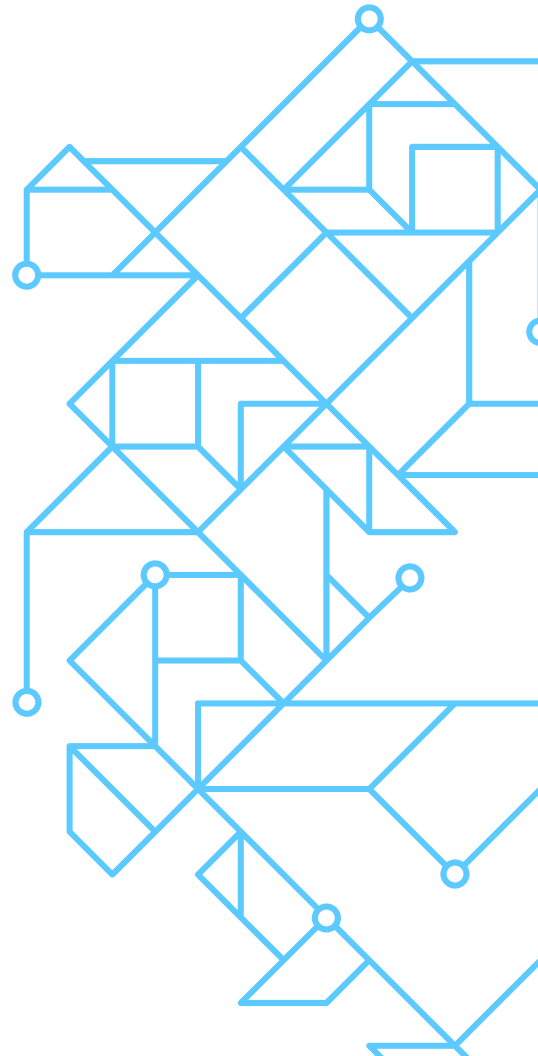
Yunxuan

TA



# The Plan

Here's what to expect today.





# Today's agenda.

<b>9:00am</b> (15 min)	<b>Talk:</b> Introduction to Ray AI Libraries
<b>9:15am</b> (15 min)	<b>Demo:</b> End-to-end mini example
<b>9:30am</b> (60 min)	<b>Coding Lab:</b> HF Vision Transformer + Ray Train for image classification
<b>10:30am</b> (15 min)	Coffee Break
<b>10:45am</b> (60 min)	<b>Coding Lab:</b> Scaling out with Ray Data and Ray Serve
<b>11:45am</b> (15 min)	<b>Talk:</b> Resources for Further Exploration

# Course Set-Up

Tech checks all around.





## Tech check.



Participating via [app.sli.do](https://app.sli.do)

- Join with code **#ray-ailib**
- Ask questions.
  - Pose your own and upvote others.
  - TAs will be answering questions on a rolling basis.



# Tech check.



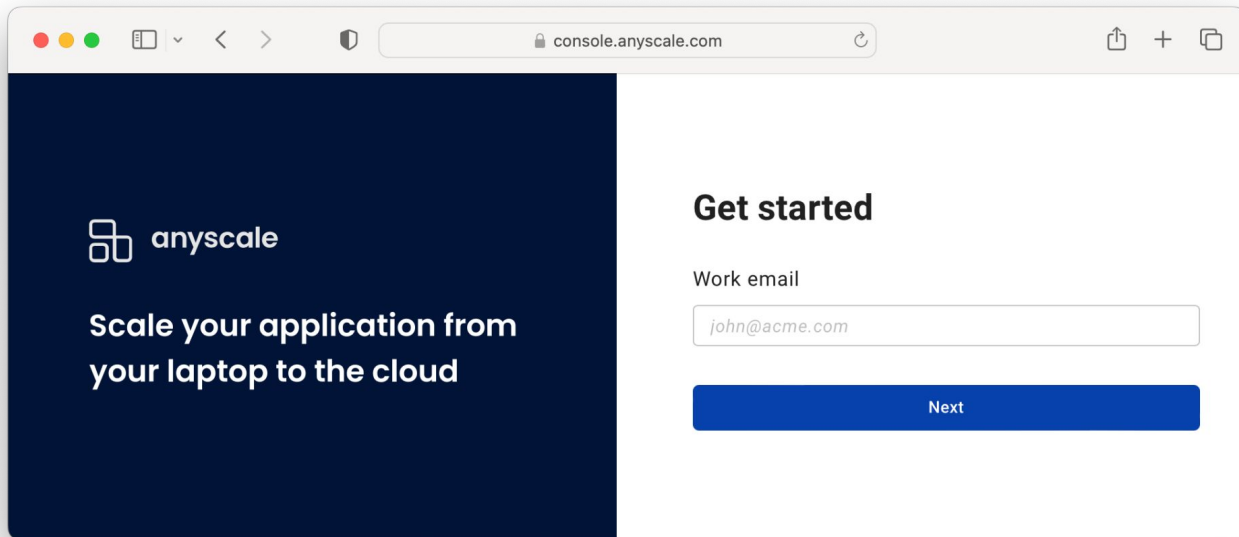
## Accessing Anyscale clusters.

- All work will be in Anyscale provisioned clusters.
- Our GitHub repo will be mounted automatically.
- Access begins now.
  - Check your email for login information.
  - Step-by-step instructions to follow.




# Anyscale login

Link to Anyscale cluster: [console.anyscale.com](https://console.anyscale.com)



The screenshot shows a web browser window with the address bar displaying 'console.anyscale.com'. The page is split into two main sections. On the left, a dark blue sidebar contains the Anyscale logo and the text 'Scale your application from your laptop to the cloud'. On the right, a white area titled 'Get started' contains a 'Work email' label, a text input field with the placeholder 'john@acme.com', and a blue 'Next' button.

 anyscale

Scale your application from  
your laptop to the cloud

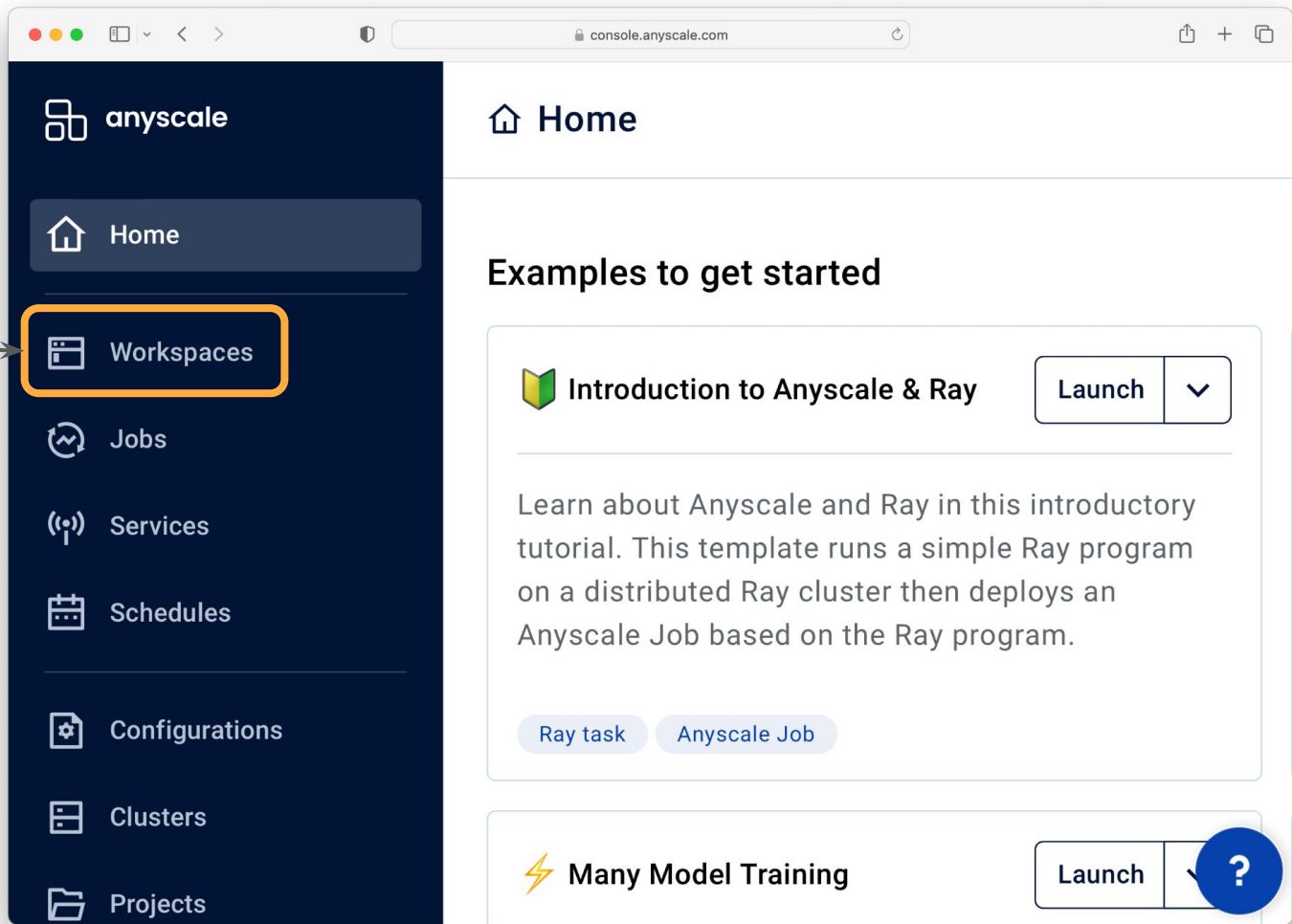
**Get started**

Work email

Next

Enter the  
**unique  
credentials**  
sent to your  
email!

# 1. Select Workspaces



anyscale

Home

Workspaces

Jobs

Services

Schedules


Configurations

Clusters

Projects


Home

### Examples to get started

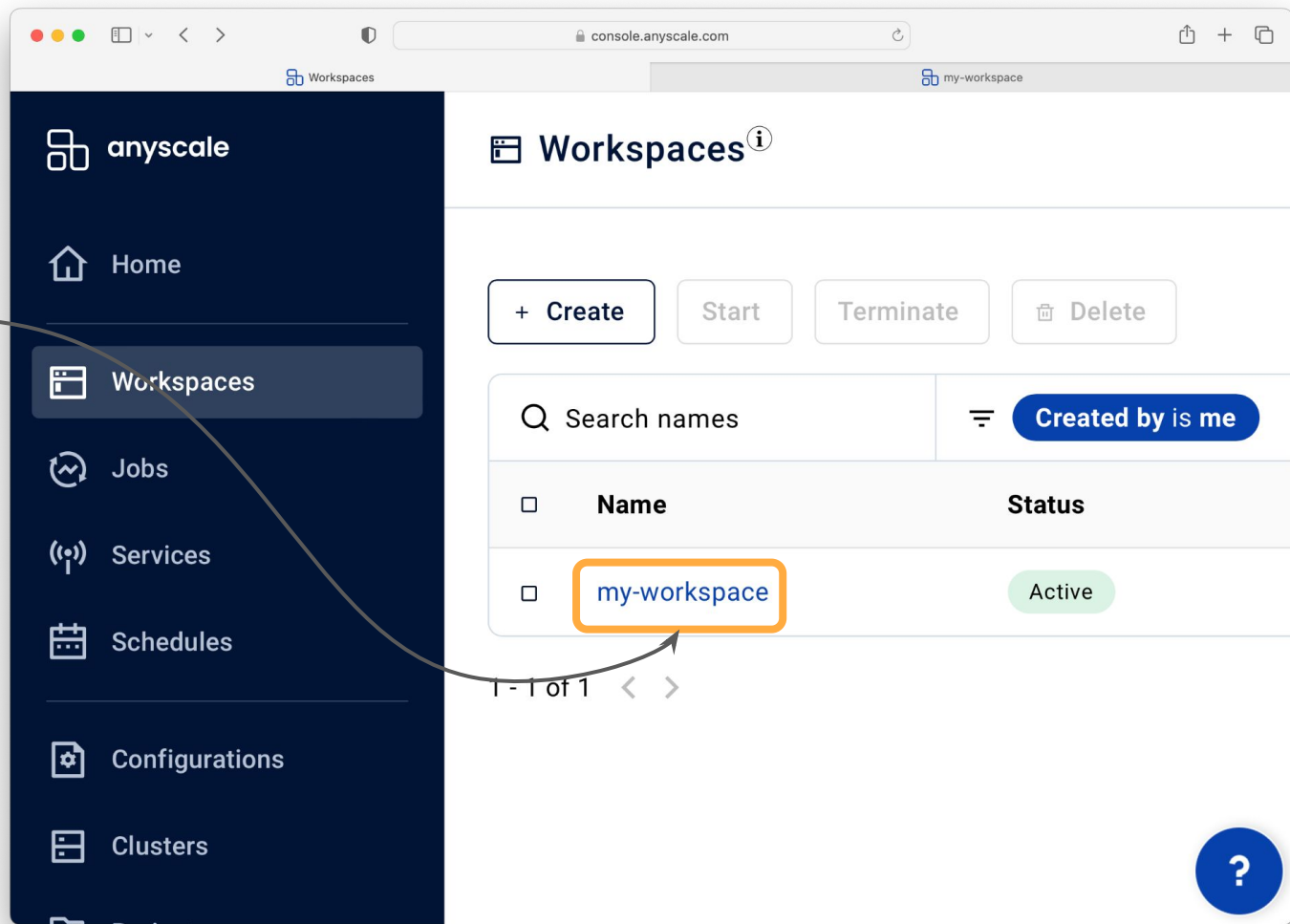
 Introduction to Anyscale & Ray Launch ▼

Learn about Anyscale and Ray in this introductory tutorial. This template runs a simple Ray program on a distributed Ray cluster then deploys an Anyscale Job based on the Ray program.

Ray task Anyscale Job

 Many Model Training Launch ?

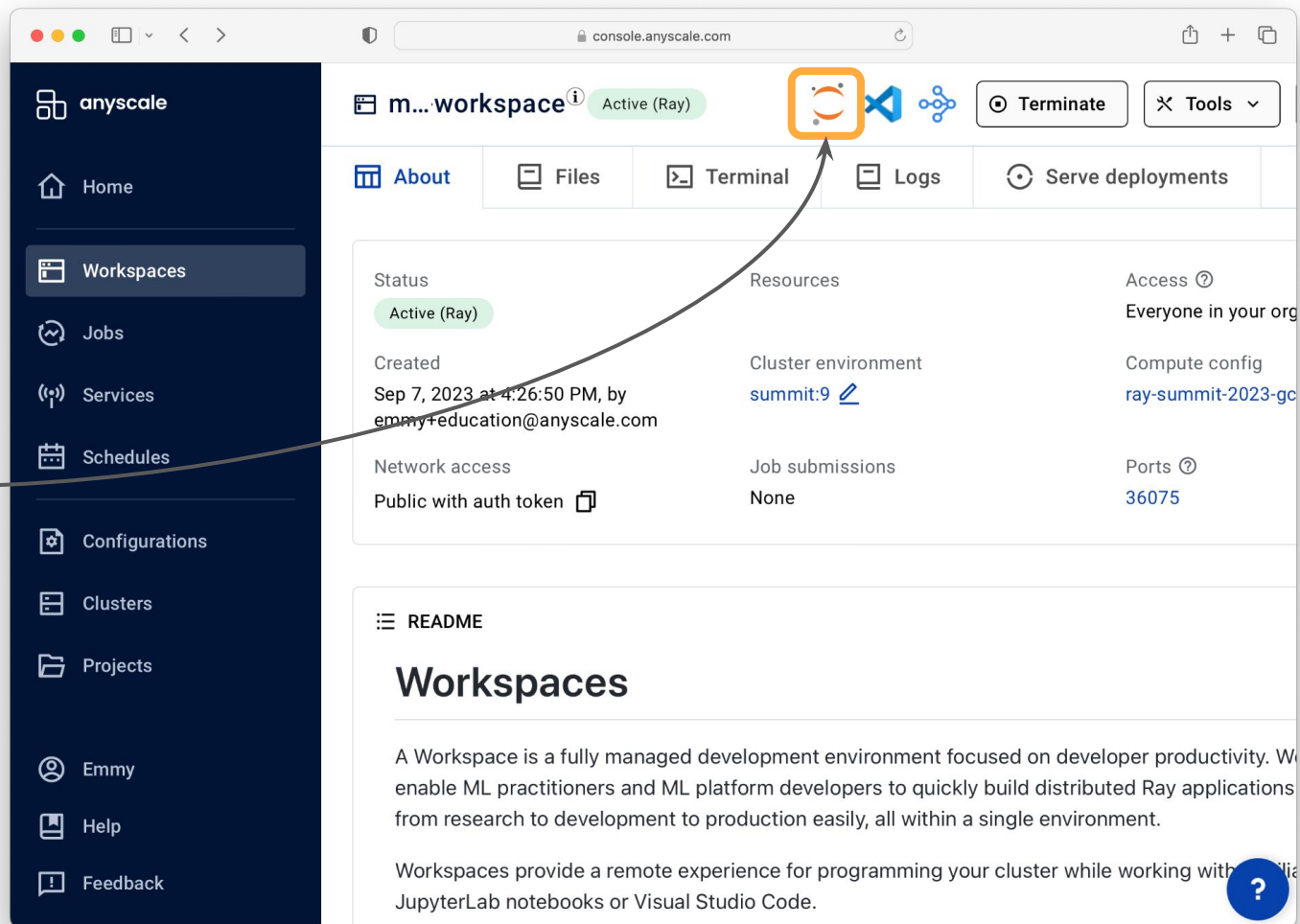
## 2. Select Your Workspace



The screenshot shows the Anyscale console interface. The left sidebar contains a navigation menu with the following items: Home, Workspaces (highlighted), Jobs, Services, Schedules, Configurations, and Clusters. The main content area is titled 'Workspaces' and includes a '+ Create' button, 'Start', 'Terminate', and 'Delete' buttons. Below these is a search bar labeled 'Search names' and a filter button labeled 'Created by is me'. A table lists the workspaces with columns 'Name' and 'Status'. The table contains one entry: 'my-workspace' with a status of 'Active'. An orange box highlights the 'my-workspace' entry, and a curved arrow points from the 'Workspaces' menu item in the sidebar to this entry. The bottom of the table shows '1 - 1 of 1' and navigation arrows. A blue help button with a question mark is in the bottom right corner.

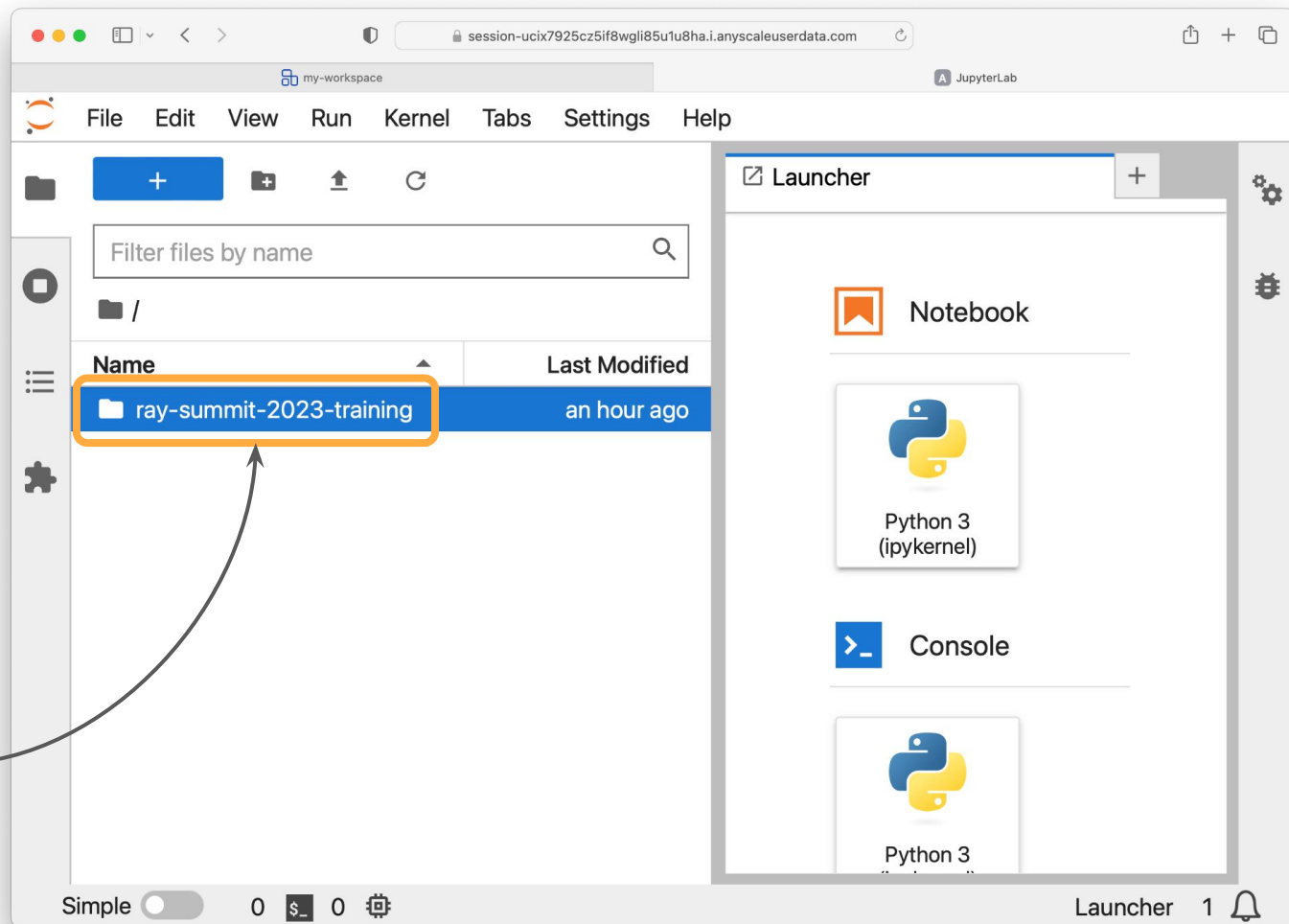
Name	Status
my-workspace	Active

### 3. Click on Jupyter icon



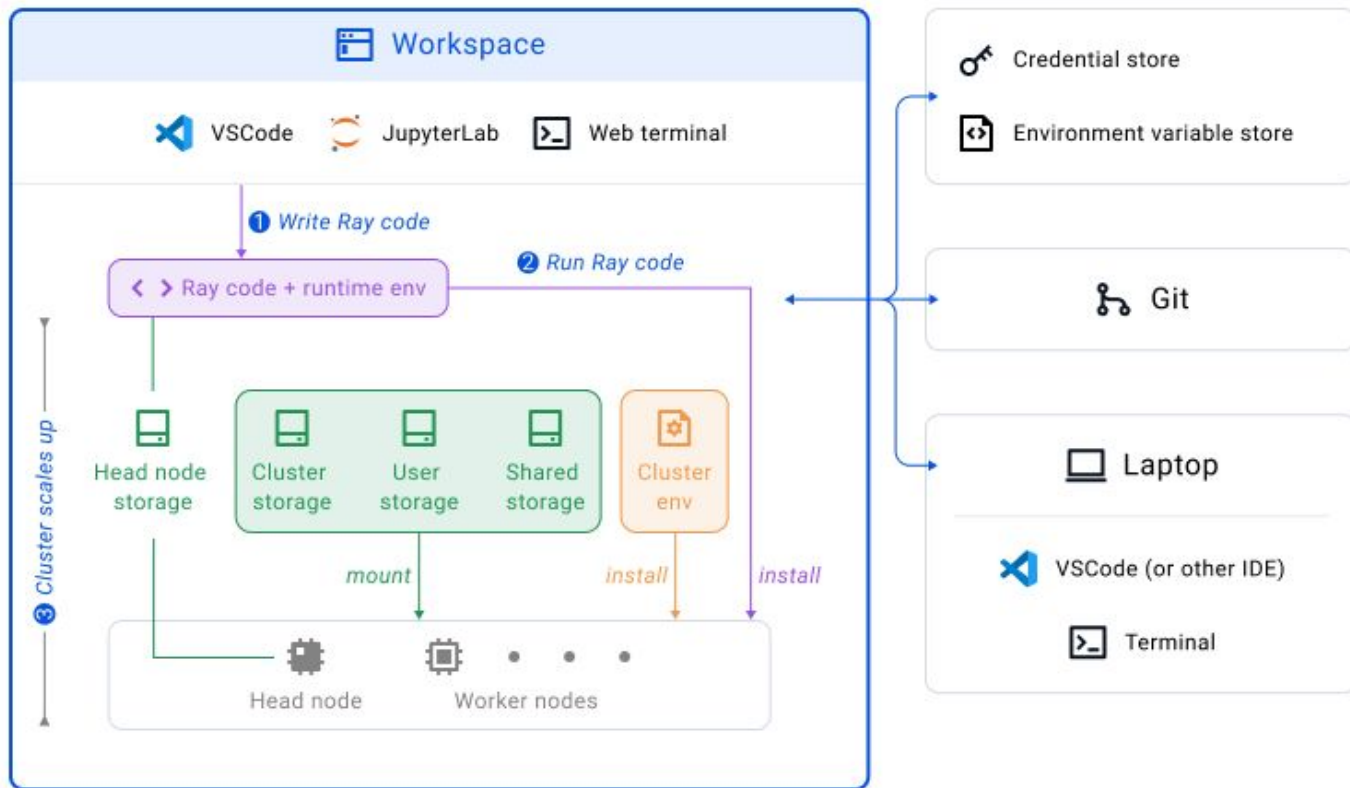
The screenshot shows the Anyscale console interface. On the left is a dark blue sidebar with navigation links: Home, Workspaces (selected), Jobs, Services, Schedules, Configurations, Clusters, Projects, Emmy, Help, and Feedback. The main content area displays details for a workspace named 'm...workspace' which is 'Active (Ray)'. At the top of this workspace view, there is a row of icons: a Jupyter icon (highlighted with an orange square), a VS Code icon, and a Ray icon. To the right of these icons are buttons for 'Terminate' and 'Tools'. Below the icons is a tabbed interface with 'About', 'Files', 'Terminal', 'Logs', and 'Serve deployments'. The 'About' tab is active, showing metadata like 'Created' (Sep 7, 2023 at 4:26:50 PM, by emmy+education@anyscale.com), 'Network access' (Public with auth token), 'Resources' (Cluster environment: summit:9), 'Job submissions' (None), and 'Access' (Everyone in your org). A 'README' section titled 'Workspaces' explains that a workspace is a fully managed development environment. A blue question mark icon is visible in the bottom right corner of the console.

4. Find the  
content for  
your class  
here.





# What are Anyscale Workspaces?



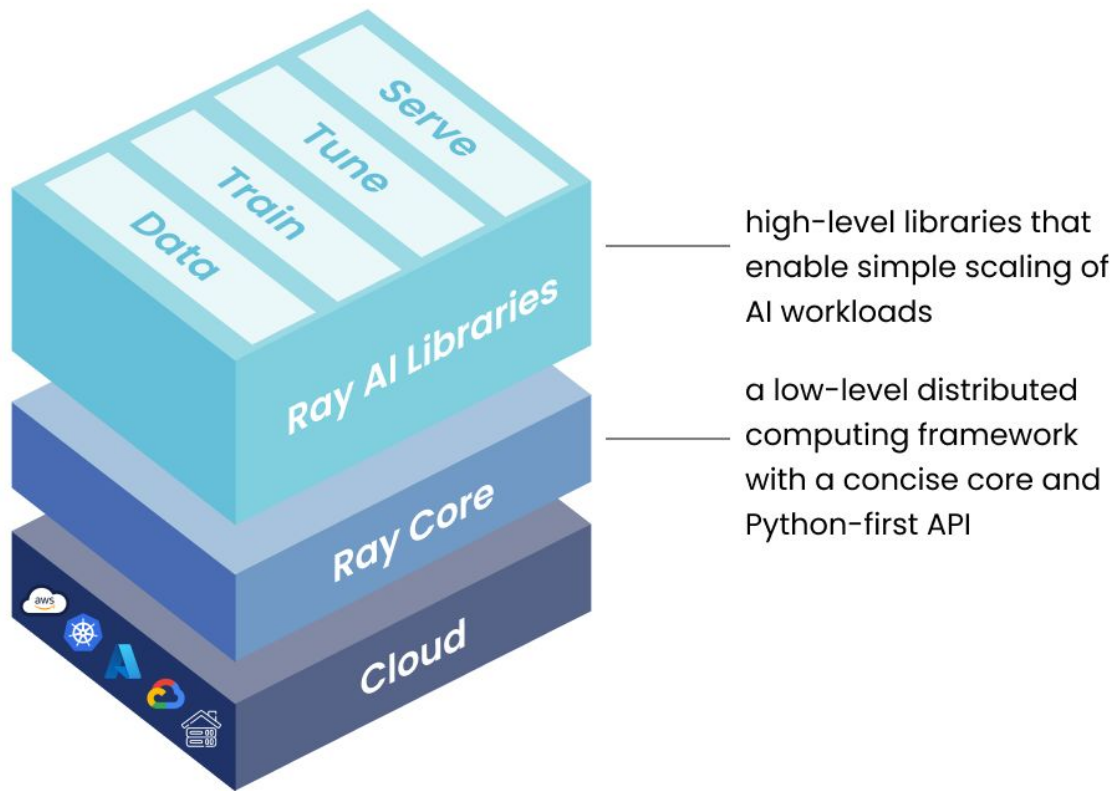
# Ray AI Libraries

Your morning briefing.





# Overview of Ray AI Libraries







# Ray Data

A scalable data processing library well-suited for distributed data ingest, preprocessing, and batch inference.

## Highlights

- + Streaming execution across CPUs and GPUs
- + World-record breaking scalability.
- + Peak performance for offline batch inference



# Ray Train

A library for distributed training and fine-tuning with integrations with PyTorch, Tensorflow, XGBoost, and much more.

## Highlights

- + [Training for LLMs](#) across 1000s of GPUs.
- + [Fine-tuning LLMs](#) with DeepSpeed.
- + Cost-savings on [heterogeneous hardware](#).



# Ray Tune

Easily scalable **hyperparameter tuning.**

## Highlights

- + Integrations with Hyperopt, Optuna, Nevergrad and more with access to algorithms like PBT and HyperBand/ASHA.
- + Many model training for millions of workloads.



# Ray Serve

Scalable **model serving** library for building online inference APIs.

## Highlights

- + Optimizations for LLMs like response streaming, dynamic request batching, multi-node serving, etc.
- + Framework-agnostic serving from huge models to Python business logic.



# The essentials.



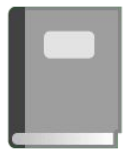
**Scalable ML** – Targeted libraries for large machine learning workloads.



**Ease of use** – Pythonic distributed computing primitives designed for smoother dev-to-prod.



**Flexibility** – Integrations with popular machine learning frameworks (e.g. PyTorch, Tensorflow, LightGBM etc.) and infrastructure (e.g. Kubernetes, AWS, GCP, Azure, etc.)



**Let's make our way over to the notebooks!**



**Time for a  
Break!**

15 minutes.

# More Resources

For further exploration with  
Ray, Anyscale, and GenAI.







# Today we learned...



## Overview of Ray AI Libraries

*Getting acquainted with each library for distributed ML.*



## Hugging Face ➡ Ray

*Converting a vision transformer to run distributed.*



## Exploring Ray Train, Data, Serve

*Constructing an end-to-end ML pipeline with Ray.*



# Sneak Peek: Self-Paced Ray & Anyscale Education



Online at [training.anyscale.com](https://training.anyscale.com)



Preview special technical content releases from the whole team!



# Fill out the survey.



Go to [bit.ly/ray-summit-feedback](https://bit.ly/ray-summit-feedback)





# Reading list.



## Self-Paced Ray & Anyscale Education

*Access bonus notebooks and scripts about Ray.*



## [Ray documentation](#)

*API references and user guides.*



## [Anyscale Blogs](#)

*Real world use cases and announcements.*



## [YouTube Tutorials](#)

*Video walkthroughs about learning LLMs with Ray.*



# Upcoming events



## Bay Area AI + Ray Summit Happy Hour

***Today at 5:00p.m.***

*Cap off an exciting conference with lightning talks, new friends, and good times!*

***[bit.ly/bayai\\_ray\\_meetup](https://bit.ly/bayai_ray_meetup)***





# Connect with the community.



## Join the community

[Attend events](#), [subscribe to newsletter](#), [follow on Twitter](#).



## Get support

[Join Ray Slack](#), [ask questions on forum](#), [open an issue](#).



## Contribute to Ray

[Read contributor guide](#), [create a pull request](#).

# Thank you!

We hope to meet again.

