

[Careers](#)

Distributed Training Engineer, Sora

San Francisco, CA — Sora

[Apply now ↗](#)

About the Team

The Sora team is working on making video a key capability of OpenAI's foundation models. We are a hybrid research and product team that seeks to understand and expand the capabilities of our video models, while ensuring their reliability and safety. We accomplish this both through directly studying and experimenting with the models, as well as deploying them into the real-world to distribute their benefits widely.

About the Role

As a Distributed Systems/ML engineer, you will work on improving the training throughput for our internal training framework and enable researchers to experiment with new ideas. This requires good engineering (for example designing, implementing, and optimizing state-of-the-art AI models), writing bug-free machine learning code (surprisingly difficult!), and acquiring deep

knowledge of the performance of supercomputers. We're looking for people who love optimizing performance, understanding distributed systems, and who cannot stand having bugs in their code.

This role is based in San Francisco, CA. We use a hybrid work model of 3 days in the office per week and offer relocation assistance to new employees.

In this role, you will:

- Collaborate with researchers to enable them to develop systems-efficient video models and architectures
- Apply the latest techniques to our internal training framework to achieve impressive hardware efficiency for our training runs
- Profile and optimize our training framework

You might thrive in this role if you:

- Have experience working with multi-modal ML pipelines
- Love diving deep into systems implementations and understanding their fundamentals in order to improve their performance and maintainability
- Have strong software engineering skills and are proficient in Python.
- Have experience understanding and optimizing training kernels
- Are passionate about understanding stable training dynamics

About OpenAI

OpenAI is an AI research and deployment company dedicated to ensuring that general-purpose artificial intelligence benefits all of humanity. We push the boundaries of the capabilities of AI systems and seek to safely deploy them to the world through our products. AI is an extremely powerful tool that must be created with safety and human needs at its core, and to achieve our mission, we must encompass and value the many different perspectives, voices, and experiences that form the full spectrum of humanity.

We are an equal opportunity employer and do not discriminate on the basis of race, religion, national origin, gender, sexual orientation, age, veteran status, disability or any other legally protected status.

For US Based Candidates: Pursuant to the San Francisco Fair Chance Ordinance, we will consider qualified applicants with arrest and conviction records.

We are committed to providing reasonable accommodations to applicants with disabilities, and requests can be made via this [link](#).