

How OpenAI and Weights & Biases Work Together

The OpenAI + Weights & Biases collaboration enables developers to access GPT-3 fine-tune experiments on Weights & Biases' best-in-class platform, bringing experiment tracking, dataset versioning, and model management into their workflows and building powerful AI applications faster.

Sync your OpenAI fine-tune runs in one line of code

We made visualizing and understanding your GPT-3 models as simple as possible. Nothing changes with how you're currently fine-tuning those models: you'll still use OpenAI's API and have all the same, robust controls you're used to. To see and compare runs, visualize and version your datasets, analyze your results, and a whole lot more, all you need is one line of code. It's really that simple:

```
$ # one line command  
$ openai wandb sync
```

A central dashboard and source of truth

Weights & Biases is perhaps best known for our experiment tracking dashboard. Once you've synced your fine-tunes with that one line of code above, you'll be able to dig into performance and compare different metrics and predictions across your various model training runs.

artifact

“With the new W&B and OpenAI integration, we’re able to consolidate experiment tracking of our in-house ML projects and OpenAI fine-tuning projects in one place. This gives us huge time savings towards delivering insights from unstructured data for our customers.”

PREM VISWANATHAN
ML Lead at Artifact.io

Our dashboards are customizable so you can create the views and look into the metrics that matter most for your project. You can share dashboards across your team to increase collaboration or pin certain notable insights into our Reports to share with interested stakeholders.

Weights & Biases is an entire MLOps toolbox in 5 lines of code. Here's what we offer:

EXPERIMENT DASHBOARD

- W&B lets teams track, compare, log, and visualize every step of model development.
- That means teams have everything they need to reproduce their models, debug performance in real-time, capture code, see live updates on model development, check for overfitting, and more.

ARTIFACTS

- Data and model pipelines are complex. They have deep interlocking dependencies where a single change in just one place can affect models downstream.
- Artifacts logs and visualizes those pipelines and records every version of your datasets so you can effortlessly understand your training flow and track your models.
- Great for companies with strict regulatory or compliance requirements

SWEEPS

Sweeps provides a scalable, customizable tuning system that allows teams to optimize their hyperparameters.



Log & explore predictions

A great way to evaluate generative models is by looking at their predictions on data they've not yet seen (a.k.a. validation data). Weights & Biases' Tables is a great way to both judge your models performance and share insights with your team.

For example, you could easily create a Table with a prompt, a target, and a generation from your fine-tune, then sort them like you would a typical csv. This is an easy and intuitive way to make comparisons and find patterns in your predictions, helping you understand what your model is excelling at and where it might need a little extra work.

Additional resources

If you'd like to get a little more information about the collaboration, you can check out our docs on how it works, as well as a couple blog posts that walk you through some workflows and tips about how best to leverage W&B with your GPT-3 fine-tunes.

- DOCS:**
<https://wandb.me/openai-docs>
- TUTORIAL:**
<http://wandb.me/openai-report>
- TEXT GENERATION ON DR. WHO**
<http://wandb.me/dr-who-gpt-3>

TABLES

- Tables makes it easier to iterate on and evaluate models.
- Log, version, organize, data in intuitive visualizations that help teams find patterns in their datasets.
- Tables lets you see and interact with your data and models in an interactive way, leading to faster insights and deeper understanding.

REPORTS

- ML practitioners can sometimes have trouble sharing their work with their team and internal stakeholders. Reports lets them do just that.
- Reports are interactive documents where practitioners can take notes, share findings, plan future research, and, above all else, collaborate across themes.

STITCH FIX
Your partner in personal style

“We are excited to use W&B capabilities to monitor the fine-tuning of our OpenAI models that help improve how we understand client feedback at Stitch Fix.”

REZA SOHRABI
Staff ML Engineer, Stitch Fix Algorithms



LEARN MORE

ABOUT WEIGHTS & BIASES

Weights & Biases is the premier, developer-first MLOps platform. We help ML teams unlock their productivity by optimizing, visualizing, collaborating on, and standardizing their model and data pipelines – regardless of framework, environment, or workflow.