Fireworks × Microsoft Foundry

The right model for every software and platform workload, all on Azure.

Open weight models on Fireworks on Foundry. Billed and governed through Azure.

SOFTWARE AND PLATFORM TEAMS BUILDING ON FIREWORKS

Adobe · Replit · Genspark · Harvey · Lovable

Internal use only. Customer logos must be removed or replaced before sharing externally.

01 / YOUR AI JOURNEY ON FOUNDRY

STAGE 1 · SHIP

Run your highest-volume workloads on open models, at frontier quality

Code generation and review

Give every developer frontier-level coding help, at the volume they actually use it. Completion, review and agentic coding across your codebase.

PROOF

Trilogy runs its own harnesses at 1/5 the cost.

Knowledge work and support agents

Answer every customer and analyst question in real time, on every call. Research, policy Q&A, summaries and support agents.

PROOF

Kimi K3 tops the Harvey and Sierra agent benchmarks.

STAGE 2 · TRAIN

Teach a model your business

Fine-tune an open model on your evals and production data so it learns your products, policies and language. A Fireworks engineer helps you train it.

PROOF

Genspark's tuned model beat frontier models by **12%**.

STAGE 3 · OWN

Run it on Azure as your own

Your competitive edge, in your tenant. Bring your model's weights into Foundry and run them under Azure governance and billing. The IP stays yours.

FULLY MANAGED

No GPUs to manage. Fireworks runs them for you.

02 / SECURITY AND GOVERNANCE

Microsoft terms apply
Covered by the Microsoft Product Terms and DPA.

Model labs never see your data
Model providers have no access to customer data.

Your region, your Azure controls
Global Standard or Global Provisioned, with data at rest in the region you choose.

No separate BAA needed
Fireworks operates as a Microsoft subprocessor.

Microsoft security tools built in
Entra, Purview, and Defender controls through Foundry.

Full details
Privacy and compliance FAQ · trust.fireworks.ai

GET STARTED IN DAYS

azure@fireworks.ai

01 Turn it on
Enable it in the Foundry portal, pay as you go.

02 Prove it
Run your own evals with a Fireworks engineer.

03 Scale
Grow into provisioned throughput or BYOW.