

FAL vs Replicate

The two big AI media APIs, compared honestly. I pay for both and use both daily. Here is exactly what each one wins, what they cost, and how to set up the same stack I run.

01
Replicate wins on catalog16,000+ models

Replicate is the giant library: official models plus a huge **community hub**, and it covers LLMs and embeddings too, not just media. If a model exists, it is probably on Replicate. Start at **replicate.com/explore**.

02
FAL wins on speed and pricingPer video
FAL is built only for media, so it is fast: **near-zero cold starts** on warm models. And you pay **per output** (per image, per second of video), a fixed price. Replicate bills many models per GPU-second, so a slow run costs more.

03
FAL gets the exclusivesOmni = FAL only
Gemini Omni Flash (Google's video model with synced audio) is on FAL and not on Replicate. FAL also tends to get hot media models day one. Replicate's Google lineup is the **Veo** family.

04
The big models cost the same~\$3 / 10s
For official flagships like **Seedance 2.0**, both platforms charge the vendor's price: about **\$0.30 per second** at 720p, so a 10 second clip is ~\$3 either way. Pick on speed and availability, not price.
My setup: one key each, FAL first, Replicate as fallback

Copy

```
$ echo 'FAL_KEY=your-key-here' >> .env
$ echo 'REPLICATE_API_TOKEN=your-token' >> .env
# both SDKs pick these up automatically
```

The Stack · everything you need

 [Claude Code](#)
[Claude Code](#)
[The Agent](#)
[The AI agent that writes, designs, cuts and ships everything in this guide. Free to install.](#)

[Install →](#)
 [Zernio](#)
[Zernio](#)
[Posting + DMs](#)
[Posts to IG, TikTok and story, and runs the comment-to-DM automation that delivered this page.](#)

[Start posting →](#)
 [ElevenLabs](#)
[ElevenLabs](#)
[Your Voice](#)
[Clones your voice from 60s of audio and returns word timings your edits can sync to.](#)

[Get your voice →](#)
 [HeyGen](#)
[HeyGen](#)
[Your Avatar](#)
[Your talking avatar for hook shots and full videos, driven by script or API.](#)

[Try it →](#)
Want the full comparison + my agent stack?
The exact pipelines, reels and systems, inside the community.
[Try It Free →](#)

 [Zdeno Mucina](#)
Not ready to join yet?
Take one system like this per week by email instead.

Send it weekly →

[Download as PDF ↓](#)
[Start My Free Trial →](#)