PART 2

FreeLLMAPI ka poora setup, step by step

Ek hi API key se kai free AI providers (Groq, Google AI Studio, Cerebras aur bhi) use karo. Apne laptop par, 10 minute mein.

Guide by **@sikhodigitally_** (Instagram)

FreeLLMAPI kya hai?

FreeLLMAPI ek open-source (MIT) **proxy** hai jo tumhare laptop par chalta hai. Ye alag-alag providers ke free tiers ko ek hi **OpenAI-compatible endpoint** ke peeche jod deta hai. Ek provider ki limit khatam ho jaye to ye khud agle provider par switch kar leta hai (failover). Tumhe sirf ek **unified API key** yaad rakhni hai.

> **Zaroori baat:** ye project **personal experimentation** ke liye hai aur single-user, local setup ke liye bana hai. Proxy ko public internet par expose mat karo. Har provider ke free tier ki limits aur rules alag hote hain, aur wo kabhi bhi badal sakte hain. Kisi bhi provider ka free tier abuse mat karo.

Shuru karne se pehle (Requirements)

Node.js 20+	openssl	Free provider keys
npm ke saath aata hai. Check: `node -v`	encryption key banane ke liye (Mac/Linux mein pehle se hota hai)	kam se kam 2 ya 3 provider ke account

Free keys kahan se milengi (official sites): Groq (console.groq.com), Google AI Studio (aistudio.google.com), Cerebras (cloud.cerebras.ai). Signup karke "API keys" section mein key generate karo. Wahi key Step 4 mein paste hogi.

Windows user ho? `openssl` wali command ke liye Git Bash ya WSL use karo. Ya key manually bhi bana sakte ho: koi bhi 64 characters ka random hex string (0-9, a-f) chalega.

1 Repo clone karo aur npm install chalao

Terminal kholo (Mac: Terminal, Windows: PowerShell ya Git Bash) aur ye teen commands ek ek karke chalao:

```
git clone https://github.com/tashfeenahmed/freellmapi.git
cd freellmapi
npm install
```

Kya hoga: project tumhare folder mein download hoga aur saari dependencies install ho jayengi. Isme 1 se 3 minute lag sakte hain.

2 .env file banao (encryption key + port)

Pehle example file ko copy karke `.env` banao. Phir ek random encryption key aur port usme daalo:

```
cp .env.example .env
printf "ENCRYPTION_KEY=$(openssl rand -hex 32)\nPORT=3001" >> .env
```

Ab `.env` file kuch aisi dikhni chahiye:

```
ENCRYPTION_KEY=<64 characters ka random hex>
PORT=3001
```

Dhyan rakho: ENCRYPTION_KEY tumhari provider keys ko encrypt karti hai (AES-256-GCM). Isse kisi ko share mat karo aur `.env` ko GitHub par push mat karo. Agar ye key kho gayi to saved provider keys wapas padhi nahi jayengi, dobara add karni padengi.

3 Server chalao aur dashboard kholo

```
npm run dev
```

- Terminal ko **band mat karna**. Server isi mein chal raha hai.
- Browser mein kholo: `http://localhost:3001`
- Dashboard dikh gaya to server sahi chal raha hai.

Agar localhost:3001 par kuch nahi khulta: dev mode mein kuch setups mein dashboard alag port par chalta hai (jaise `localhost:5173`). Terminal ke output mein jo URL likha ho, wahi kholo.

4 Provider keys add karo (Keys tab)

- Dashboard mein **Keys** tab kholo.
- **+ Add key** dabao.
- Provider chuno (Groq, Google AI Studio, Cerebras ya koi aur) aur uski API key paste karo.
- Save karo. Jitne providers ki key daloge, utna zyada free quota milega aur failover utna behtar chalega.

Tip: kam se kam 3 providers add karo. Ek ki daily limit khatam hone par proxy automatic agle par chala jata hai, to kaam rukta nahi.

5 Unified API key copy karo

Usi dashboard mein tumhe ek **Unified API key** milegi. Ye `freellmapi-` se shuru hoti hai. Ab se baaki saari jagah (apps, scripts, Claude Code) **sirf yahi key** use hogi. Provider ki asli keys kahin aur paste karne ki zaroorat nahi.

Ek key, sab providers ke liye. Ise apne paas safe rakho.

6 Claude Code (ya apne app) mein connect karo

Tarika A: environment variables (jaisa reel mein dikhaya)

```
export ANTHROPIC_BASE_URL=http://localhost:3001
export ANTHROPIC_AUTH_TOKEN=<tumhari-unified-key>
```

Phir usi terminal mein `claude` chalao. Windows PowerShell mein `export` ki jagah `$env:ANTHROPIC_BASE_URL="http://localhost:3001"` likho.

Tarika B: auto-setup command (project ka apna helper, config ka backup bhi leta hai)

```
npx freellmapi setup-claude --url http://localhost:3001 --api-key <tumhari-unified-key>
```

Tarika C: apne Python / JS code mein (OpenAI SDK, LangChain, LlamaIndex sab chalte hain)

```
from openai import OpenAI

client = OpenAI(
    base_url="http://localhost:3001/v1",
    api_key="freellmapi-xxxxxxxx",   # unified key
)
resp = client.chat.completions.create(
    model="auto",   # router khud best free model chunega
    messages=[{"role": "user", "content": "Namaste!"}],
)
print(resp.choices[0].message.content)
```

7 Routing "Balanced" rakho. Setup ho gaya!

Dashboard mein routing ko **Balanced** par rakho. Isse proxy speed aur limits ke beech balance banakar request ko healthy provider par bhejta hai. Ab test karo:

```
curl http://localhost:3001/v1/chat/completions \
  -H "Authorization: Bearer <tumhari-unified-key>" \
  -H "Content-Type: application/json" \
  -d '{"model":"auto","messages":[{"role":"user","content":"Hi"}]}'
```

Jawab aa gaya to setup 100% sahi hai.

Problem aaye to (Troubleshooting)

Problem	Kya karo
`npm install` fail ho raha hai	`node -v` check karo. 20 se kam hai to Node.js update karo.
Dashboard nahi khul raha	Terminal mein `npm run dev` chal raha hai ya nahi dekho. Output mein diya hua URL/port kholo.
Port 3001 busy hai	`.env` mein PORT badal do (jaise 3002) aur URLs bhi usi hisaab se badlo.
401 / unauthorized error	Provider key nahi, **unified key** use karo. Key mein extra space ya line break to nahi?
Rate limit / 429 error	Us provider ka free quota khatam hua. Aur providers ki keys add karo.
Server restart ke baad keys gayab	`.env` ka ENCRYPTION_KEY wahi rakho jo pehle tha. Badla to keys dobara add karni padengi.

Seedhi baat (limits)

- Free tier models paid wale (Claude, GPT) jitne strong nahi hote. Chhote kaam, learning aur testing ke liye achhe hain, bade production kaam ke liye nahi.
- Har provider ki rate limit aur terms alag hain aur badalte rehte hain. Uski official site check karte raho.
- Ye tool tumhare machine par chalta hai. Laptop band to proxy band.
- Sensitive ya client data free providers ko mat bhejo. Unki privacy policy padh lo.

Ye guide kaam aayi?
Roz aise free AI tools aur vibe coding ke setups ke liye follow karo **@sikhodigitally_** Instagram par. Reel par comment karo aur agli guide ka link DM mein aa jayega.

Source: project README aur public guides (github.com/tashfeenahmed/freellmapi). Commands aur ports project ke version ke hisaab se badal sakte hain. Ye guide sirf educational hai.