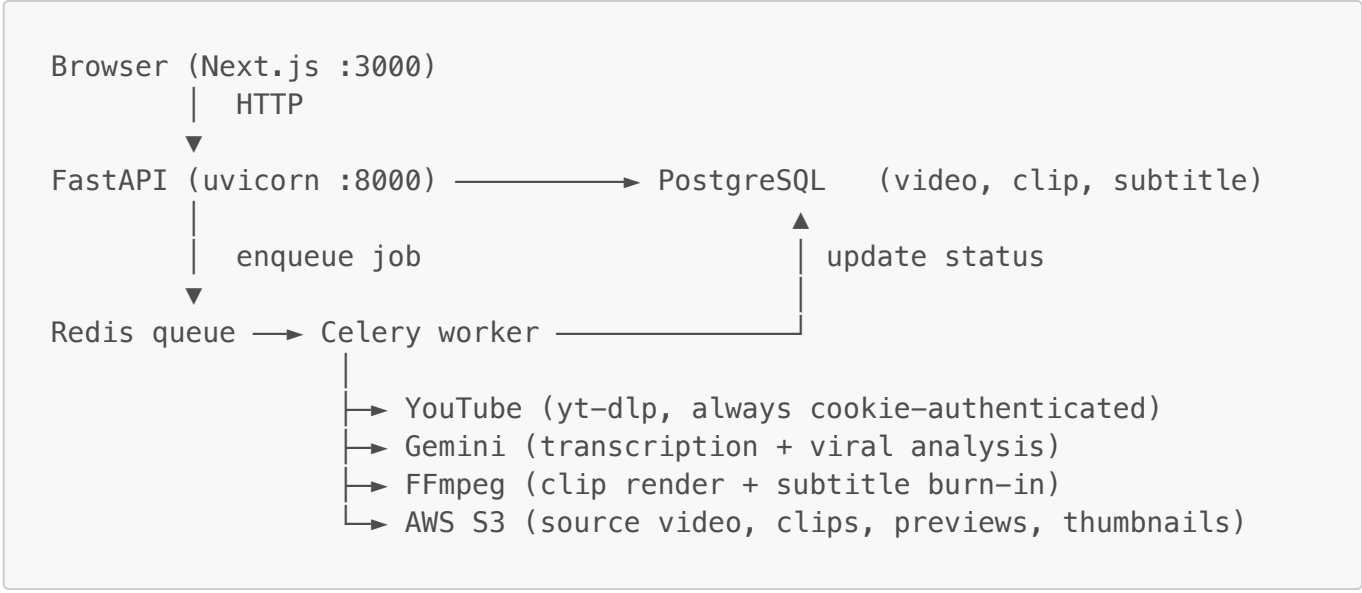
5. ZClip — AI YouTube Clipper

Role	Personal project (vibe coding)
Stack	FastAPI + Celery (Python), Next.js, PostgreSQL, Redis, AWS S3, FFmpeg (libass), yt-dlp, Gemini / Whisper, Google OAuth, Midtrans Snap

Overview. An automated YouTube → vertical-clip pipeline. A user pastes a YouTube link; the backend analyses which moments have viral potential, cuts them into clips, burns styled subtitles onto each one, and hands back editable, download-ready vertical videos for TikTok, Reels, and Shorts.

Architecture. Three concurrent processes:



The API stays responsive during long jobs because all heavy lifting — download, transcription, AI analysis, rendering — happens in the Celery worker.

How the pipeline works.

Video processing (`process_video_task`, once per video):

1. **Cache-first source retrieval** — check S3 for `source_videos/{video_id}.mp4`; if present, pull from S3 rather than touching YouTube; otherwise download audio via yt-dlp.
2. **Transcription** — audio converted to 16 kHz FLAC and sent to the speech-to-text provider, returning text *and per-word timestamps*, the foundation for subtitles and cutting.
3. **Viral analysis** — the transcript is sent to an AI provider that selects candidate moments with `start_time`/`end_time`, a score, and a rationale; out-of-range results are clipped.
4. **Persistence** — one Clip row per moment (score, style, quality), with the transcript words in that range written out as Subtitle rows.

5. **Render queueing** — a `render_clip_task` enqueued per clip, user quota deducted by video duration, video status set to `completed`.

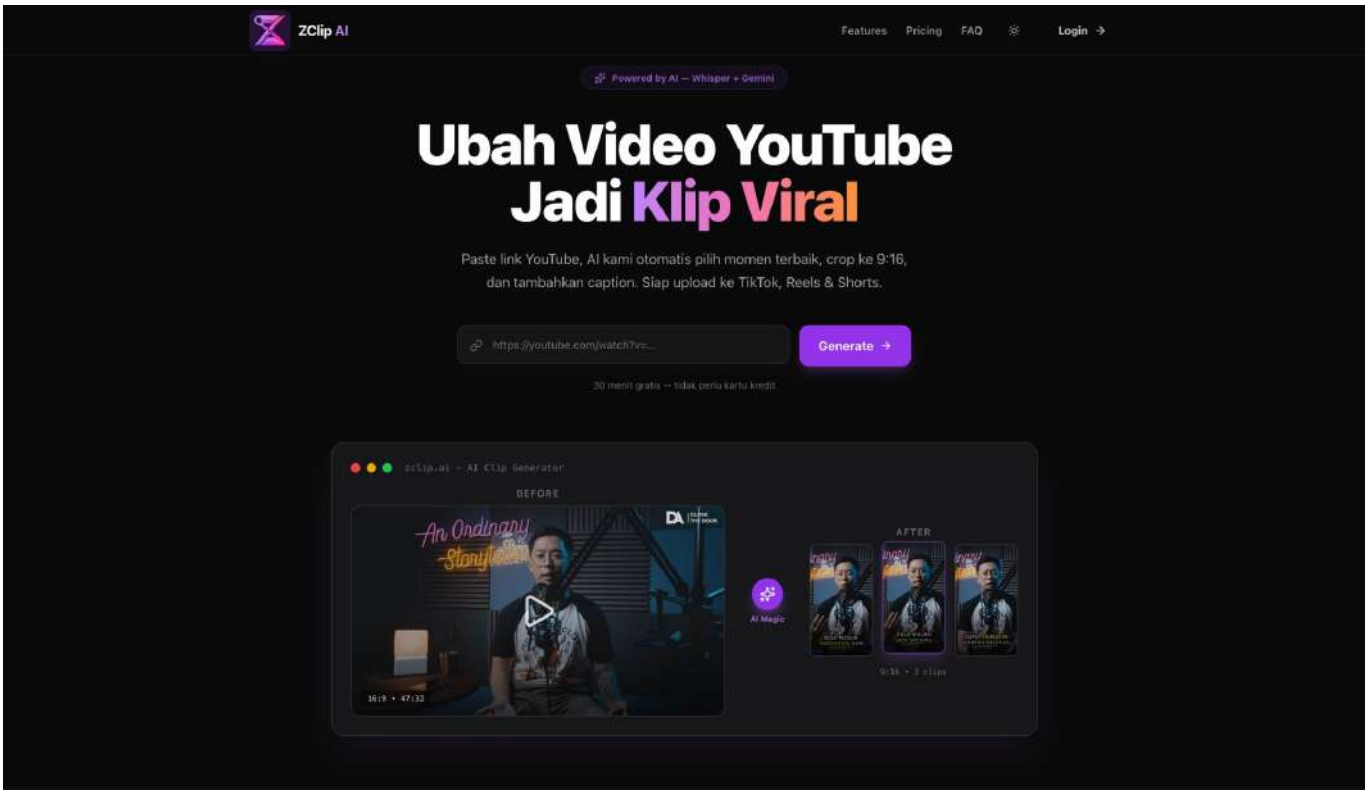
Clip rendering (`render_clip_task`, per clip):

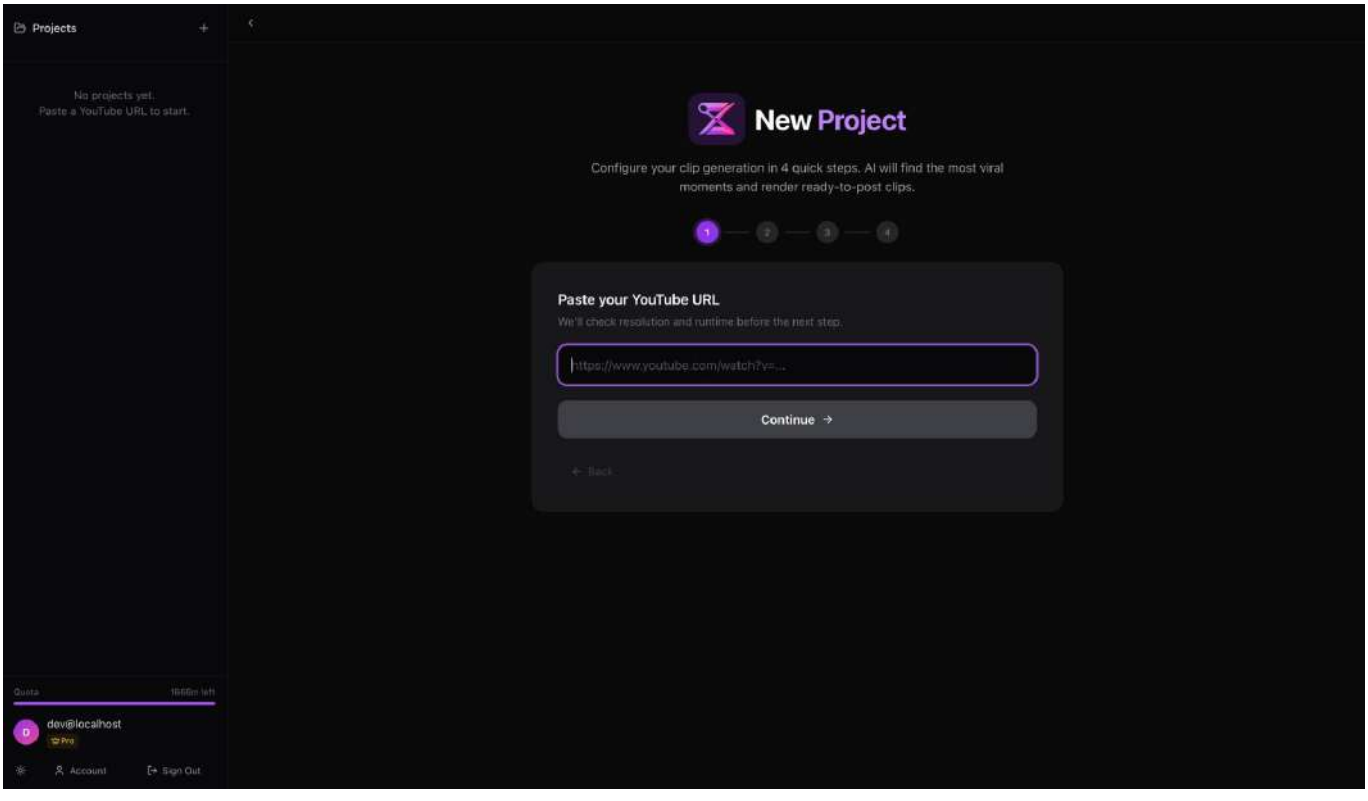
1. Fetch the source (S3 cache, else segment download from YouTube).
2. Cut to `start_time`–`end_time`, crop to 9:16 (or leave 16:9), scale to target resolution.
3. Generate an `.ass` subtitle file from the per-word timestamps and burn it in with FFmpeg's `ass` filter using the selected style.
4. Upload to S3, generate a lightweight preview and thumbnail, store the URLs on the clip row, and mark it `completed`.

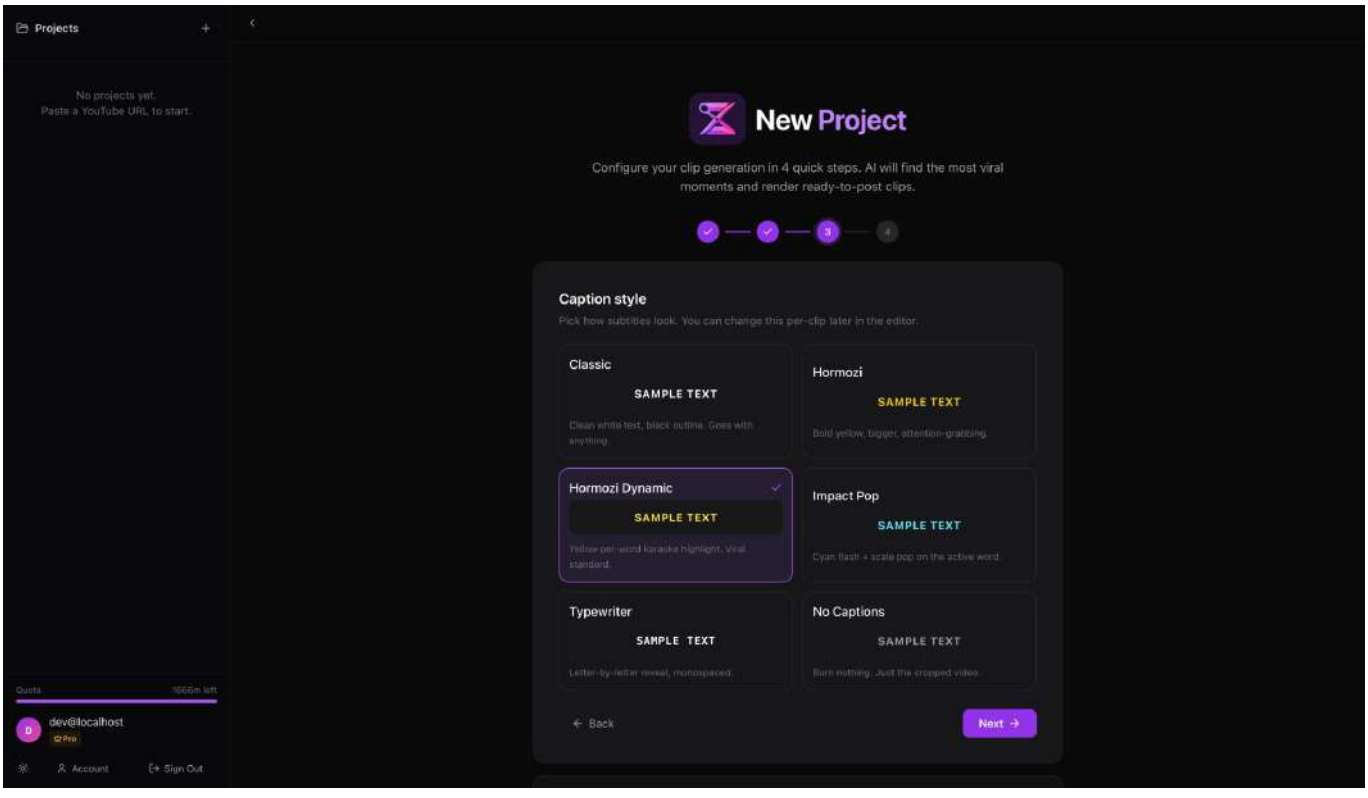
Highlights:

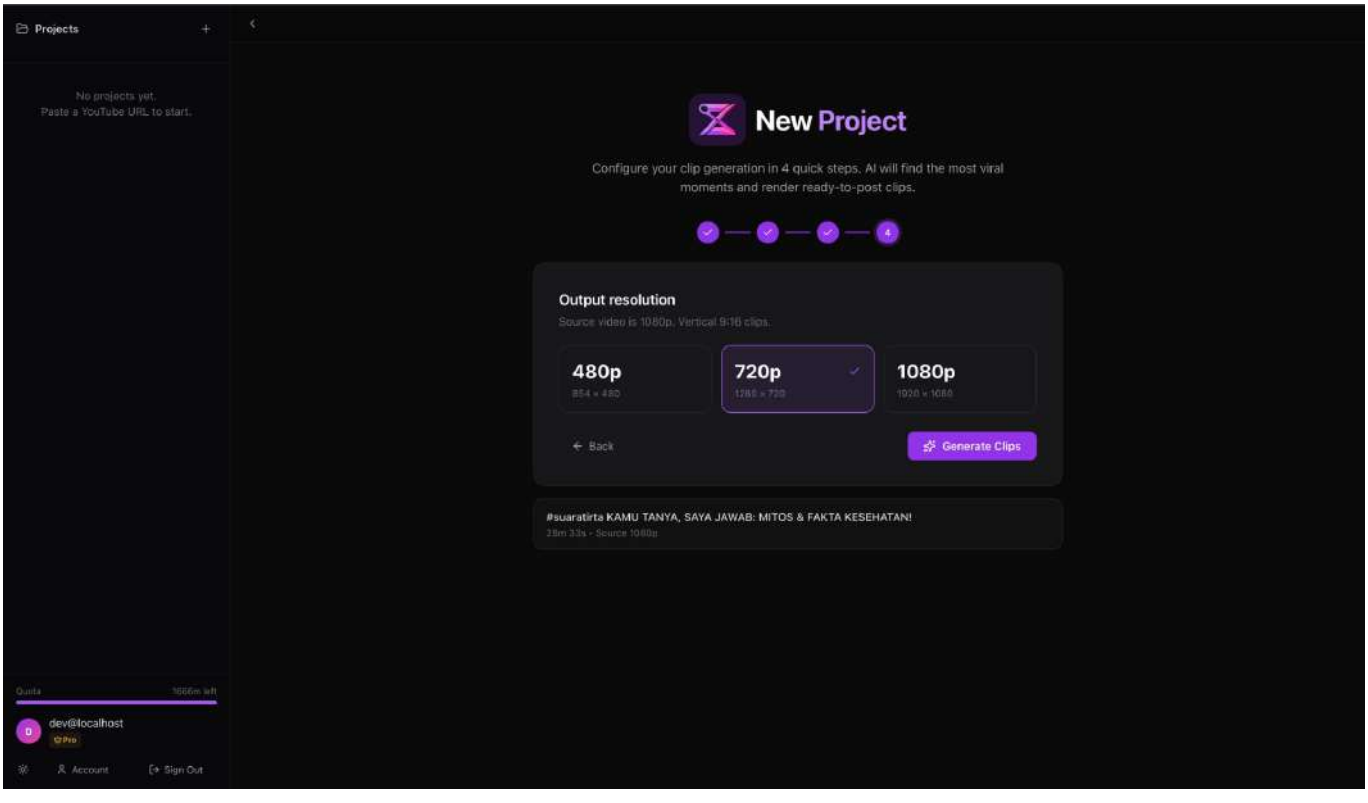
- Built an event-driven background processing system with explicit status lifecycle: `downloading → transcribing → analyzing → rendering → completed / failed`, polled by the UI for progress.
- Implemented **S3 caching of source videos** to avoid redundant YouTube downloads and cut downstream latency and quota cost.
- Designed the subtitle model around per-word timestamps, enabling word-level subtitle editing, timing corrections (`POST | PATCH | DELETE /subtitles`), and per-word karaoke-style caption animations.
- Delivered a 4-step project wizard (URL → clip length → caption style → output resolution) with a clip grid, viral-score badges, and a modal editor supporting preview, crop adjustment, caption-style switching (which triggers a background re-render), and a lightweight video proxy endpoint for smooth mobile playback.
- Configurable AI stack: transcription via Gemini `gemini-3.5-transcribe` or Groq Whisper; viral analysis via Gemini, Claude, or DeepSeek.
- Auth in dev bypass mode locally, Google OAuth in production; optional Midtrans Snap tier upgrades with quota enforcement.
- Documented operational pitfalls end to end — FFmpeg builds requiring `libass` for the `ass` filter, cookie-only YouTube access, and exact S3 region/endpoint matching to avoid `SignatureDoesNotMatch`.

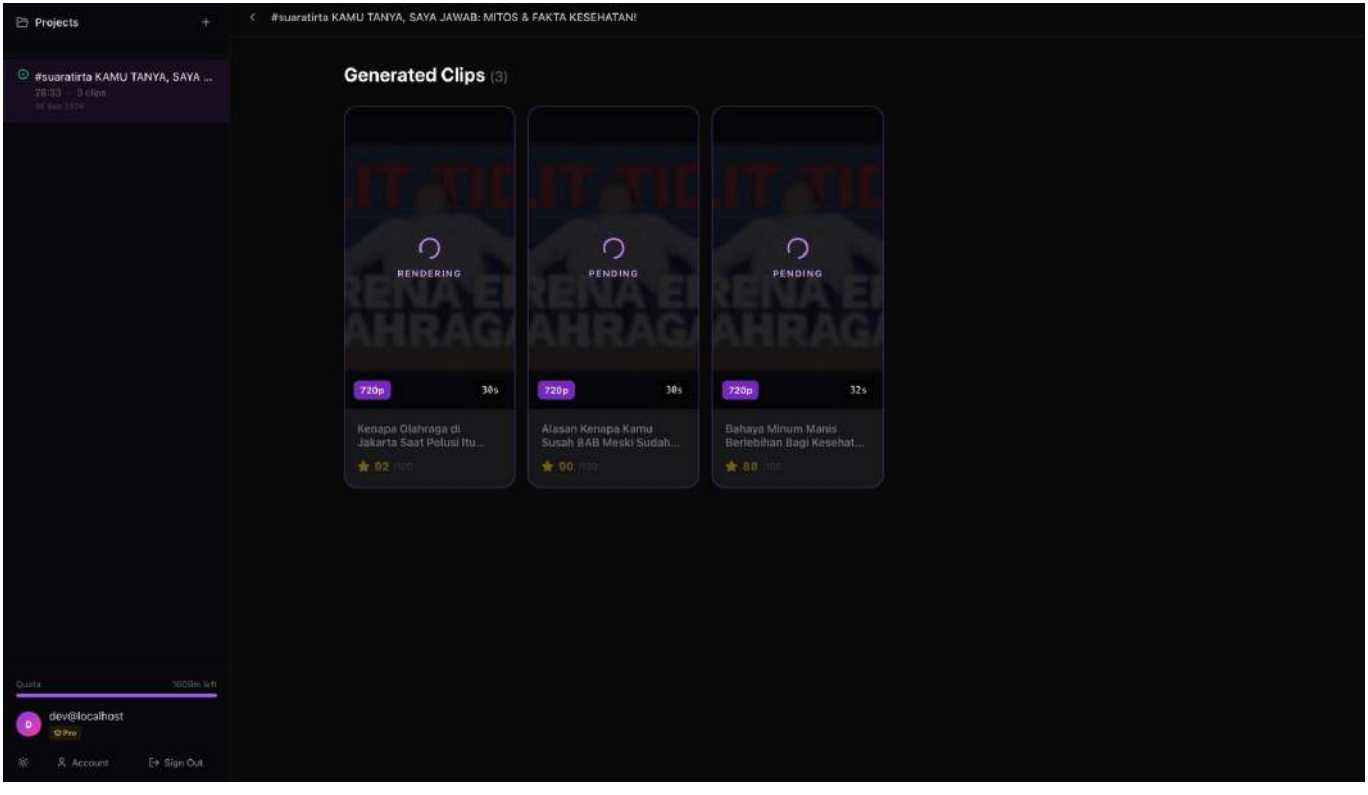
Screenshots:


ZClip AI
Features
Pricing
FAQ
Login
Powered by AI — Whisper + Gemini
Ubah Video YouTube
Jadi Klip Viral
Paste link YouTube, AI kami otomatis pilih momen terbaik, crop ke 9:16,
dan tambahkan caption. Siap upload ke TikTok, Reels & Shorts.
https://youtube.com/watch?v=...
Generate
30 menit gratis — tidak perlu kartu kredit
zclip.ai — AI Clip Generator
BEFORE
AFTER
AI Magic


Projects
No projects yet.
Paste a YouTube URL to start.
Quota
dev@localhost
Pro
Account
Sign Out
New Project
Configure your clip generation in 4 quick steps. AI will find the most viral
moments and render ready-to-post clips.
Paste your YouTube URL
We'll check resolution and runtime before the next step.
https://www.youtube.com/watch?v=...
Continue
Back


Projects
No projects yet.
Paste a YouTube URL to start.
New Project
Configure your clip generation in 4 quick steps. AI will find the most viral moments and render ready-to-post clips.
Caption style
Pick how subtitles look. You can change this per-clip later in the editor.
Classic
SAMPLE TEXT
Clean white text, black outline. Goes with anything.
Hormozi
SAMPLE TEXT
Bold yellow, bigger, attention-grabbing.
Hormozi Dynamic
SAMPLE TEXT
Yellow per-word karaoke highlight. Viral standard.
Impact Pop
SAMPLE TEXT
Cyan flash + scale pop on the active word.
Typewriter
SAMPLE TEXT
Letter-by-letter reveal, monospaced.
No Captions
SAMPLE TEXT
Burn nothing. Just the cropped video.
Back
Next
Quota
dev@localhost
Pro
Account
Sign Out


Projects
No projects yet.
Paste a YouTube URL to start.
New Project
Configure your clip generation in 4 quick steps. AI will find the most viral moments and render ready-to-post clips.
Output resolution
Source video is 1080p. Vertical 9:16 clips.
480p
720p
1080p
Back
Generate Clips
#suaratirta KAMU TANYA, SAYA JAWAB: MITOS & FAKTA KESEHATAN!
Quota
dev@localhost
Pro
Account
Sign Out


Projects
#suaratirta KAMU TANYA, SAYA ...
#suaratirta KAMU TANYA, SAYA JAWAB: MITOS & FAKTA KESEHATAN!
Generated Clips (3)
RENDERING
PENDING
PENDING
720p
30s
720p
30s
720p
32s
Kenapa Olahraga di Jakarta Saat Polusi Itu...
Alasan Kenapa Kamu Susah BAB Meski Sudah...
Bahaya Minum Manis Berlebihan Bagi Kesehat...
82
90
88
Quota
dev@localhost
Pro
Account
Sign Out