UPDATED EDITION — 18 OCTOBER 2026

AI Video Production Blueprint

স্বয়ংক্রিয় মোশন ভিডিও তৈরির পূর্ণাঙ্গ পাইপলাইন — স্ক্রিপ্ট থেকে MP4 পর্যন্ত

এই গাইডের মূল ধারণা আগের সংস্করণের মতোই: স্ক্রিপ্ট, ভয়েসওভার, সাবটাইটেল ও কোড-ভিত্তিক মোশন গ্রাফিক্স এক পাইপলাইনে জোড়া দেওয়া। যা বদলেছে তা হলো টুলগুলোর সর্বশেষ সংস্করণ, খরচের বাস্তব হিসাব, আর মাঠে কাজ করতে গিয়ে শেখা ভুল ও তার সমাধান।

1 **Script Drafting**
ভয়েসওভার স্ক্রিপ্ট ও প্রতি বাক্যে Visual Cue

2 **TTS Voiceover**
Gemini TTS দিয়ে প্রাকৃতিক ভয়েসওভার অডিও

3 **Precise SRT**
অডিও থেকে টাইমকোডসহ সাবটাইটেল ফাইল

4 **Master Prompt**
৫ ধাপের ফর্মুলায় সাজানো নির্দেশনা

5 **Claude Code**
Claude Remotion কোড বানায়; render হয় লোকালে

সৎ ব্যাখ্যা: Claude ভিডিও render করে না

Claude কোনো rendering engine নয়। Claude Remotion (React-ভিত্তিক) কোড তৈরি করে দেয়। সেই কোড নিজের কম্পিউটারে Remotion CLI দিয়ে চালালেই MP4 তৈরি হয়। এই পার্থক্যটা বুঝলে প্রত্যাশা ঠিক থাকে এবং error এলে কোথায় খুঁজতে হবে তাও পরিষ্কার হয় — কোডে, Claude-এর “render”-এ নয়।

10-15 মিনিট
অডিও ও SRT তৈরি থাকলে কোড থেকে MP4 পর্যন্ত আনুমানিক সময়

9:16 / 16:9
Shorts/Reels-এর জন্য vertical, YouTube-এর জন্য widescreen

30 fps লক
SRT টাইমকোডের সঙ্গে মিলিয়ে frame rate আগেই নির্দিষ্ট করুন

এই সংস্করণে নতুন যা আছে

- সর্বশেষ মডেল: Claude 5.5 family এবং Gemini 3.8 Flash TTS — পুরনো মডেলের নাম সরিয়ে দেওয়া হয়েছে।
- প্রতি ধাপে কোন মডেল ব্যবহার করবেন তার আলাদা টেবিল (পৃষ্ঠা ৩)।
- খরচের বাস্তব হিসাব ও free-tier দিয়ে শুরু করার পথ (পৃষ্ঠা ৪)।
- সাধারণ ভুল ও সমাধান: SRT drift, বাংলা ফন্ট, safe zone, word-highlight sync (পৃষ্ঠা ৪)।
- কখন এই পাইপলাইন ব্যবহার করবেন না — সেই সীমানাটাও স্পষ্ট করা হয়েছে।

Bornil Sohel — Visual & Tech Strategist
bornilsohel.pro.bd
© 2026 Bornil Sohel. All rights reserved.

৫ ধাপ বিস্তারিত: প্রতিটি ধাপে ঠিক কী ঘটে

ধাপ ০১ — স্ক্রিপ্ট তৈরি

ChatGPT বা Claude দিয়ে ভয়েসওভার স্ক্রিপ্ট লিখুন। প্রতি লাইনে visual cue রাখুন — কোন কথার সঙ্গে স্ক্রিনে কী দেখা যাবে। কাঠামো রাখুন: প্রথম ৩ সেকেন্ডে Hook, মাঝখানে মূল কথা, শেষে Call-to-Action। ৬০ সেকেন্ডের ভিডিওর জন্য বাংলায় মোটামুটি ১৩০-১৫০ শব্দ যথেষ্ট; এর বেশি হলে TTS দ্রুত বলতে বাধ্য হয় এবং sync কঠিন হয়ে যায়।

ধাপ ০২ — ভয়েসওভার অডিও (Gemini TTS)

Google AI Studio-তে **gemini-3.8-flash-tts** বা **gemini-3.8-flash-lite-tts** ব্যবহার করুন (২৩ সেপ্টেম্বর ২০২৬-এ প্রকাশিত)। এতে 2,000+ voice, 100+ ভাষা এবং 30-second voice cloning আছে; AI Studio-র free tier-এও পাওয়া যায়। স্ক্রিপ্ট পেস্ট করে tone ও pacing ঠিক করে .mp3 বা .wav ডাউনলোড করুন। ফাইনাল অডিওটাই master — এর পর স্ক্রিপ্ট বদলালে SRT আবার বানাতে হবে।

ধাপ ০৩ — মিলিসেকেন্ড-নির্ভুল SRT

ফাইনাল অডিও Gemini-তে আপলোড করে নির্দেশ দিন: “Generate a standard SRT file with millisecond-accurate timestamps for this audio.” পাওয়া SRT যাচাই করুন — প্রথম ও শেষ cue অডিওর শুরু-শেষের সঙ্গে মেলে কিনা। এখানে ভুল হলে পুরো ভিডিওর sync ভুল হবে, তাই এই ধাপে তাড়াহুড়ো নয়।

ধাপ ০৪ — Master Prompt তৈরি

পরের পৃষ্ঠার ৫ অংশের ফর্মুলা অনুসরণ করুন: Context, Task, Constraints, Output Format, Example। রেফারেন্স ভিডিও থাকলে তার transition, color scheme ও composition আলাদা করে লিখে দিন — “ওই ভিডিওর মতো” বললেই Claude বুঝবে না।

ধাপ ০৫ — Claude Code ও লোকাল Render

Claude Sonnet 5.5-কে Master Prompt, অডিও ফাইলের নাম ও SRT-এর পুরো লেখা দিন। Claude একটি single-file Remotion component দেবে। কোডটি লোকাল Remotion setup-এ রেখে CLI দিয়ে render করুন — এখানেই আসল MP4 তৈরি হয়।

© 2026 Bornil Sohel. All rights reserved.

Prompt ফর্মুলা ও Master Prompt Template

ভালো prompt মানে লম্বা prompt নয় — সাজানো prompt। নিচের ক্রমটি বজায় রাখুন:

অংশ	কাজ
1. Context	আপনি কে, ভিডিওটা কার জন্য, কোন platform-এ যাবে
2. Task	ঠিক কী বানাতে হবে — এক prompt-এ একটাই task
3. Constraints	সীমা ও নিষেধ: সময়, safe zone, frame rate, কী করা যাবে না
4. Output Format	আউটপুটের চেহারা: single-file Remotion component, ratio, palette
5. Example	নমুনা timecode sync বা style value দিয়ে "এরকম চাই" দেখানো

কপি-রেডি Master Prompt Template

```
Role: Expert Remotion / Motion Graphics Video Developer.
Context: [আপনার পরিচয়, দর্শক, platform — যেমন Facebook Reels, বাংলা দর্শক]
Task: অডিও ও SRT অনুযায়ী kinetic typography ভিডিওর
single-file Remotion (React/TSX) component তৈরি করুন।
Input: Audio: [audio.mp3] | SRT: [পুরো SRT লেখা পেস্ট করুন]
Spec: Ratio: 9:16 (1080x1920) | FPS: 30 | Duration: audio-locked
Palette: Background #0d1117 | Accent #38bdf8 | Highlight #f59e0b
Constraints:
1. শব্দ ধরে ধরে highlight (word-by-word) — SRT timecode হুবহু মিলবে
2. সাবটাইটেল safe zone-এ থাকবে, platform UI দিয়ে ঢাকা যাবে না
3. বাংলা লেখায় Hind Siliguri ফন্ট embed করবেন
Example: 00:00:02,000 এ "প্রযুক্তি" শব্দটি highlight রঙে বড় হবে।
```

কোন ধাপে কোন মডেল (অক্টোবর ২০২৬ অনুযায়ী সর্বশেষ)

ধাপ	সুপারিশ	কেন
স্ক্রিপ্ট	Claude Sonnet 5.5 / ChatGPT	দ্রুত draft ও visual cue সাজানো
ভয়েসওভার	gemini-3.8-flash-tts	2,000+ voice, voice cloning, free tier
SRT	Gemini (অডিও আপলোড)	timestamp-সহ SRT আউটপুট
Remotion কোড	**Claude Sonnet 5.5** (২৮ সেপ্টেম্বর ২০২৬)	কোডের জন্য সেরা balance; সবচেয়ে কঠিন কাজে Opus 5.5
ছোট sub-task	Claude Haiku 5.5 (৭ অক্টোবর ২০২৬)	সস্তা ও দ্রুত — caption ভাগ করা, নাম বদলানো

© 2026 Bornil Sohel. All rights reserved.

খরচ, সাধারণ ভুল ও Quality Control

খরচের বাস্তব হিসাব

- **TTS:** Gemini TTS প্রায় $0.81 প্রতি ঘণ্টা অডিও (২০২৬ সালের শেষ পর্যন্ত এই হার) — ৬০ সেকেন্ডের ভিডিওর খরচ প্রায় নগণ্য।
- **Claude:** API ব্যবহারে token অনুযায়ী খরচ; chat-এ draft করলে প্রায় শূন্য।
- **Free-first পথ:** আগে AI Studio free tier ও Claude free chat দিয়ে পুরো workflow পরীক্ষা করুন, তারপরই paid-এ যান।

Iteration Loop — একবারে perfect আশা করবেন না

Generate → Preview → কোন জায়গায় সমস্যা তা নির্দিষ্ট করে prompt-এ fix লিখুন → আবার generate। “ভালো হয়নি, আবার করুন” বললে একই ভুল ফিরে আসে; “দ্বিতীয় cue ০.৪ সেকেন্ড দেরিতে আসছে” বললে ঠিক হয়। প্রতি round-এ একটাই বড় fix নিন।

সাধারণ ভুল ও সমাধান

সমস্যা	সমাধান
SRT drift — শেষে গিয়ে sync মিলছে না	SRT ফাইনাল অডিও থেকেই বানান; মাঝের ২-৩টি cue ম্যানুয়ালি মিলিয়ে দেখুন; প্রয়োজনে cue ভাগ করুন
বাংলা অক্ষর ভাঙা বা বাক্স দেখায়	Remotion-এ Hind Siliguri-র মতো সঠিক বাংলা ফন্ট embed করুন; system font-এর উপর ভরসা নয়
সাবটাইটেল platform UI দিয়ে ঢাকা পড়ে	9:16-এ লেখা মাঝের safe zone-এ রাখুন; নিচের প্রায় ২০% ও ডান পাশ button এলাকা এড়িয়ে চলুন
Word highlight কথার সঙ্গে মেলে না	শব্দভিত্তিক timestamp নিন; শুধু line cue দিয়ে word animation হুবহু হবে না
অডিও clipping / hiss	ফাইনাল loudness প্রায় -14 LUFS রাখুন; render-এর আগে অডিও normalize করুন

QC Checklist — render-এর আগে

- অডিও -14 LUFS, কোনো clipping নেই
- SRT শুরু, মাঝ ও শেষ — তিন জায়গায় sync মিলেছে
- Frame rate কোডে 30 fps লক করা
- বাংলা ফন্ট ঠিক render হচ্ছে
- Background ও লেখার contrast মোবাইলে পড়া যায়
- Safe zone-এর বাইরে কোনো জরুরি লেখা নেই

কখন এই পাইপলাইন ব্যবহার করবেন না

Character animation বা cinematic shot এই পাইপলাইনের কাজ নয় — সেগুলোর জন্য Google Flow বা Kling-ধরনের video generator ব্যবহার করুন। এই পাইপলাইন সবচেয়ে ভালো কাজ করে: kinetic typography, explainer-ধাঁচের ভিডিও, quote/template ভিডিও এবং তথ্যভিত্তিক Shorts-এ — যেখানে লেখা, সময় ও layout-ই মূল বিষয়।

Bornil Sohel — Visual & Tech Strategist
bornilsohel.pro.bd
© 2026 Bornil Sohel. All rights reserved.