Generative B-Roll Service - Build Blueprint

Fora Soft Learn · AI for Video Engineering · Capstone 10.4

1 - The reference architecture: router-first, two gates

Front door	Editor brief (text or reference still) -> structured generation request.
Router	Pick a swappable model by cost, look, and commercial-safety. The model is a part, not a partner.
Two gates	Quality gate (auto checks + human) THEN rights+provenance gate (C2PA + rights ledger).
Handoff	Conform -> encode -> package into the platform's OTT pipeline. Audit every step.

2 - Build vs buy (2026): rent the models, integrate the encoder, build the rest

```
BUY / HOST  video models (Veo 3.1 / Runway / Kling / Marey / open Wan, LTX) - fastest-moving part
BUILD       model router + brief builder + quality gate + rights/provenance ledger = your product
INTEGRATE   conform / encode / package on the platform's existing OTT encoder - never rebuild it
NOTE        OpenAI switched Sora off in 2026 (API end 24 Sep) - that is why you abstract the model.
```

3 - The 2026 model field (price per second of output; verify before you ship)

```
Google Veo 3.1  ~$0.15-0.40/s  native audio   | Runway Gen-4 Turbo  ~$0.05/s   | Kling  ~$0.10/s
COMMERCIAL-SAFE: Adobe Firefly Video (licensed + IP indemnity) | Moonvalley Marey (licensed)
Route paid-title / photoreal clips to a licensed, indemnified model; cheap general models for low-risk.
Open weights (Wan, LTX): your GPU cost only, your data stays in - but you own the legal review.
```

4 - Per-clip cost arithmetic (one 6-second clip)

```
Generation   6 s x $0.15/s x 4 candidates  =  ~$3.60   (+ brief & gates ~$0.10-0.40)  => ~$4 / clip.
Cheap route (general model, fewer tries) ~ $1.50. Premium stock $50-200. Custom shoot $1,000s.
Even the safe generative path is 1-2 orders cheaper than licensing. Pays off above ~20-30 clips/month.
Keep the always-on service infrastructure on a SEPARATE fixed line from the per-clip compute.
```

5 - Build order: each milestone ships a working tool

```
1 Gated single-model tool (thin router + auto quality check + human review).
2 Provenance + rights ledger (C2PA credential on every clip) - second, NOT last.
3 Multi-model router (add a licensed, indemnified model + routing policy).
4 OTT encode-and-package integration.    5 Scale, cost controls, Article 50 disclosure.
```

6 - Go-live + rights checklist (engineering context, not legal advice)

- [ ] Model router in place before anything calls a model - any vendor swappable by config.
- [ ] Paid-title / photoreal clips routed to a licensed, indemnified model (Firefly / Marey).
- [ ] C2PA content credential on every generated clip AND recorded in your rights ledger (transcode strips it).
- [ ] Scope line held: no real identifiable person, real brand mark, or real event shown as factual.
- [ ] EU AI Act Article 50: AI clips disclosed to EU viewers; machine-readable mark from 2 Aug 2026.
- [ ] Cost instrumented: log every model call, tune candidates-per-clip, set per-show budgets.