

Local-First Dubbing Workflow Checklist

A stage-by-stage checklist for shipping AI-dubbed video without giving up control. Use it to audit an existing pipeline or spec a new one. Print it, share it, adapt it — attribution appreciated.

1. Ingest & project setup

- ☐ Source media stored on a local disk you control (not a synced cloud folder).
- ☐ Working project folder is excluded from cloud backup services that scan file contents.
- ☐ Source clip demuxed; original picture stream preserved untouched.
- ☐ Audio normalized to a single sample rate before ASR.
- ☐ Project has a target language, target loudness, and delivery spec written down.

2. Transcription (ASR)

- ☐ ASR runs locally (DirectML / TensorRT RTX / CPU fallback), not against a remote API.
- ☐ Transcript is editable per line, with timestamps preserved.
- ☐ Names, jargon, and product terms reviewed and corrected before translation.
- ☐ Overlapping speech flagged for manual cleanup.

3. Translation

- ☐ Translation is editable line-by-line without re-running ASR.
- ☐ Register, formality, and idioms reviewed against target-market norms.
- ☐ On-screen text and burned-in captions handled separately from dialogue.
- ☐ Glossary of protected terms (brands, product names) applied consistently.

4. Diarization & speaker mapping

- ☐ Each speaker has a stable ID across the whole project.
- ☐ Reassigning a speaker on one line doesn't invalidate other stages.
- ☐ Voice prints stored locally, never uploaded without explicit opt-in.

5. Voice generation (TTS)

- ☐ TTS runs locally with acceleration appropriate to the machine.
- ☐ Per-line regeneration is possible without rerunning the pipeline.

- ☐ Timing budget per line respects the shot; overruns flagged, not silently trimmed.
- ☐ Prosody / pacing controls exposed for problem lines.

6. Mix & preview

- ☐ Original music and effects preserved; only dialogue is ducked/replaced.
- ☐ Optional stem separation available for tricky source mixes.
- ☐ Preview render is fast enough to iterate — not a queued cloud job.
- ☐ Final mix meets the delivery loudness spec (e.g. -23 LUFS EBU R128, -16 LUFS streaming).

7. Reliability & recovery

- ☐ Jobs are resumable after crash, cancel, OOM, or power loss.
- ☐ Failure classes are typed (download, lock, OOM, cancel) — not string-matched.
- ☐ Errors name the stage AND the cause; no generic "failed".
- ☐ Completed stages are not silently re-run when one line changes.

8. Privacy & sovereignty

- ☐ Media, transcripts, and voice prints stay on the workstation by default.
- ☐ Any cloud stage is opt-in per project AND per stage — never implicit.
- ☐ NDA / client footage never leaves the machine without written approval.
- ☐ Model artifacts are checksum-verified before load.

Companion guide: trackdub.com/guides/ai-dubbing-guide · © Trackdub. Free to share with attribution.