Longitudinal media review for audition-oriented violin practice

Alan N. Pham[1]

[1]AO Labs, Worcester, MA, USA

High-level music training produces dense evidence: daily practice videos, repeated repertoire, changing technical constraints, and periodic performance goals. Most of that evidence is difficult to use because it is stored as unstructured media rather than as a longitudinal record. We report Curtis Media Review, a live AO Labs system that indexes public practice media, identifies practice candidates, records model-assisted section reviews, and converts the current state into a compact practice record. The checked live state on 7 May 2026 used the YouTube channel `@nalalan` as the primary source, indexed 207 public videos, marked 102 as practice candidates, marked 82 as long-form candidates, stored 6 media samples, retained 11 reviewed sections, and recorded 51 Curtis-focused findings. No repertoire work is currently named at clear confidence. The first screen therefore reports *Piece being identified*, an identifying state instead of a piece-completion percentage, and one major recording/practice constraint. Possible labels and user-rejected false labels are not allowed to become current repertoire titles, and new piece-title output requires corroboration from a second audio verifier before it can update the piece list. The system uses GPT-5.5 for vision and text review and separate audio models for identification and verification, while preserving explicit limits: Curtis admission cannot be predicted from current samples, weak audio/video evidence is marked unjudged, and the system reports practice signals rather than admissions probability. The contribution is an experimental method for turning accumulated practice media into auditable feedback loops, progress history, and source-bounded coaching suggestions.

1 Introduction

Elite music preparation depends on repeated, deliberate practice and feedback over long time horizons (*1*, *2*). The useful data are not only polished performances. They include ordinary practice sessions, failed takes, camera angle limitations, recurring technical problems, and changes in focus across days. A violinist recording daily practice can therefore produce a large longitudinal dataset, but the raw dataset does not automatically become an operating record.

Curtis Media Review addresses this problem as a live software system rather than as a portfolio page (*3*). The system indexes public practice media, classifies candidate practice logs, records technical observations, and keeps the current practice constraint visible. It is designed for the Curtis Institute goal, but it does not claim to estimate an acceptance probability. The current role of the system is narrower: observe practice evidence, retain history, separate judged from unjudged evidence, and return a small number of practice constraints that can be tested in subsequent recordings.

The system is also part of AO Labs' broader progress architecture. In May 2026, AO Progress was introduced as a shared ledger for cross-application state and long-term changes across public apps, papers, CV artifacts, Curtis practice state, Imagineer state, and Relay state (*4*). Curtis supplies one stream in that ledger: timestamped media and review evidence for musical development.

2 System state

The first production state used the live backend at `https://curtis.aolabs.io/api/curtis/media-status`. The checked state on 7 May 2026 reported the values in Table 1. These are operating facts, not outcome claims.

2.1 Media indexing

The source inventory stores platform, title, publication time, duration, view count, candidate reasons, media kind, and blocker state. Dated long-form practice logs are treated as high-value longitudinal records because the title contains a practice date and the duration preserves session-level evidence. Other videos are retained but not forced into the practice lane.

Table 1. Curtis production state on 7 May 2026.

Signal	Current value
Primary source	YouTube channel `@nalalan`
Inventory	207 indexed videos
Practice candidates	102 videos
Long-form candidates	82 videos
Media samples	6 captured samples
Reviewed sections	11 sections
Curtis-focused findings	51 findings
Latest dated practice log	5-1-26
Review model	GPT-5.5 for vision/text, separate audio models for audio evidence and piece verification
Current piece	Piece being identified; no clear repertoire title
Curtis-level completion	Withheld until the piece is identified at clear confidence
Current focus	Audition-style setup consistency
Current constraint	Open strings only, mirror-check contact point
Boundary	Curtis admission cannot be predicted from current samples

2.2 Review evidence

The review layer stores section-level findings with a dimension, evidence source, judgment, and practice constraint. The current dimensions include tone, intonation, shifts, time, articulation, and audition delivery. Findings can be strong, needing work, or unjudged. The unjudged state is part of the system design: if the sound is not available, the hand is obscured, the camera angle hides a transition, or still frames do not support a timing claim, the system should not convert weak evidence into a confident critique.

3 First readout

The first live readout is experimental but usable. The strongest immediate value is that the system has already converted a large, timestamped practice archive into a queryable corpus with captured media samples and stored section reviews. The current piece state is not a named repertoire work. The home screen reduces the current state to the active piece-identification state, a withheld completion percentage, and one major practice or recording constraint. Generic etude/caprice evidence remains outside the piece-title field, and possible or uncorroborated repertoire labels are suppressed. The

current active focus is audition-style setup consistency, with open strings only and mirror-check contact point as the practice condition. These statements are not generic pedagogy; they are extracted from the stored review findings, piece record, and current progress plan.

Table 2. Current evidence classes and limits.

Class	Stored signal	Limit
Audio	Tone, intonation, shifts, time, articulation	Some audio reviews fail or cannot inspect the excerpt cleanly
Video	Setup, posture, bow lane, visible hand/bow evidence	Still frames can hide motion, sound, and transition accuracy
Metadata	Dates, duration, channel, view count, practice-candidate labels	Metadata cannot judge playing quality by itself
Progress plan	One focus, practice constraint, short session plan	The plan is a next experiment, not an outcome prediction

4 Discussion

Curtis Media Review is intentionally conservative. It does not rank the player against applicants, infer admissions probability, or convert every video into a claim. Instead, it keeps a running technical record. This makes the system useful even before it becomes musically sophisticated: it can preserve practice volume, surface recurring constraints, identify missing evidence, and make subsequent recordings easier to compare against earlier ones.

The main open problem is repertoire identification and evidence quality. YouTube metadata can identify candidate videos, and the current owner-sync path can provide direct audio/video samples, but unclear camera framing, missing score/title context, and short excerpts still prevent reliable piece naming. The backend therefore keeps generic labels out of the piece-title field, stores them as candidate evidence, and marks weak evidence explicitly. A better system would combine reliable excerpt extraction, audio fingerprinting, score/title context, visual posture analysis, and longitudinal trend scoring.

The broader experiment is whether a practice system can make musical development legible over time. A useful Curtis system should show whether tone, intonation, rhythm, articulation, shifts, memory, endurance, and audition delivery are improving across dated recordings. It should also preserve the source evidence behind each claim so that the dashboard does not become motivational text detached from the actual playing.

5 Materials and methods

Backend

The backend is a FastAPI service deployed on Railway. It exposes `https://curtis.aolabs.io/api/curtis/media-status` for compact state, `https://curtis.aolabs.io/api/curtis/ops-check` for fuller operational state, sample indexes for duplicate-window avoidance, and run endpoints for scans, media probing, analysis, and coaching. Persistent runtime state stores sources, inventory, review findings, media samples, analysis runs, and scan history.

Source scan

The default public source is `https://www.youtube.com/@nalalan`. The scanner queries public video metadata, classifies videos by title, duration, and candidate reasons, and stores inventory records. Dated practice logs and long-form videos are treated as practice candidates. The scan does not itself inspect performance quality.

Model review

The model layer reviews sampled media sections where usable excerpts exist. GPT-5.5 is used for text and visual review, and separate audio models are configured for audio review and piece-title verification. Findings are sanitized so weak evidence terms such as obscured, not audible, or no clear evidence produce an unjudged state rather than an unsupported critique. Piece labels are also sanitized: only clear-confidence composer, work, movement, etude/caprice number, opus, key, or catalog evidence can become a piece title, and the proposed title must be corroborated by an independent verifier before it updates the current piece list. Possible labels are treated as unidentified. The owner-browser sync uses the sample index to prioritize uncaptured public YouTube windows rather than re-uploading already reviewed excerpts, and the piece-identification pass can test multiple active moments inside each captured sample before withholding the title. Piece records retain day-level progress entries so the home screen can show today's percent separately from the longer-running piece record when a piece is actually identified.

Progress ledger

AO Progress tracks Curtis as part of a larger multi-application state record. The first progress snapshot included Curtis home, Curtis media state, other AO Labs apps, Imagineer state, Imagineer paper, Relay state, and public project surfaces. Curtis contributes dated practice inventory, current focus, review status, and future trend signals.

6 Limitations

The current system cannot predict acceptance at Curtis. It also cannot replace a musician, teacher, jury, studio lesson, or audition process. It can be wrong when the source media are incomplete, when audio is unavailable, when video frames hide the physical action, or when model review fails. The useful claim is therefore limited: Curtis Media Review turns practice media into a persistent, source-bounded record that can support iteration over time.

References

1. K. A. Ericsson, R. T. Krampe, C. Tesch-Romer, The role of deliberate practice in the acquisition of expert performance. *Psychological Review* **100**, 363–406 (1993).

2. A. Williamon, Ed., *Musical Excellence: Strategies and Techniques to Enhance Performance* (Oxford University Press, Oxford, 2004).

3. A. N. Pham, *Curtis Media Review*, Live software system, curtis.aolabs.io, 2026.

4. A. N. Pham, *AO Progress Ledger*, Live longitudinal state system, progress.aolabs.io, 2026.