DAI XUDONG

From Musical Structure to Machine Creativity

COMPOSITION / SYSTEMS / HUMAN INTENTION

Selected works and ongoing research

Composer
Music AI Researcher
Creative Technologist

I use musical knowledge to build interpretable and interactive creative systems.

My work explores how musical structure and human intention can become computational representations for creative systems.

COMPOSITION

Structure, tension and musical authorship

↓

SYMBOLIC MUSIC SYSTEMS

Scores and rules that a musician can inspect

↓

MACHINE CREATIVITY

Testable attention and responsive interaction

Composition studies at Xi’an Conservatory of Music ground my work in harmony, orchestration and form.

Xunpu Studio

Inspectable music intelligence for composers and students

READ
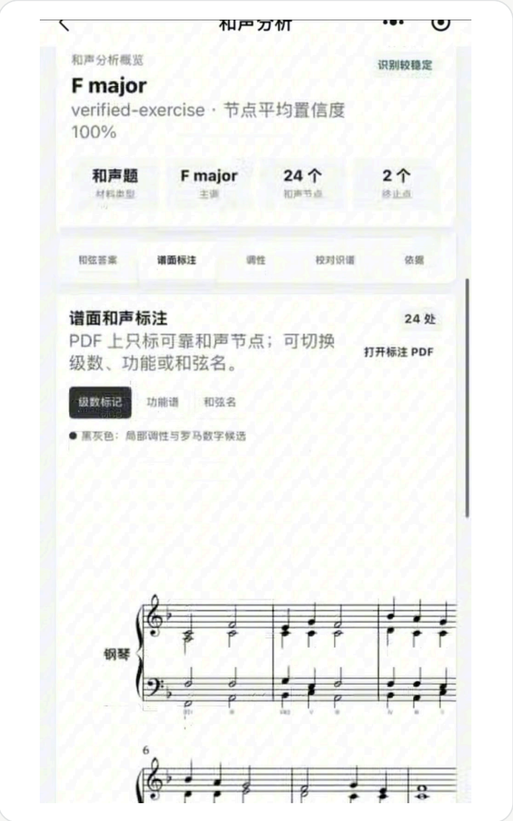
Score input becomes a symbolic representation.

REASON
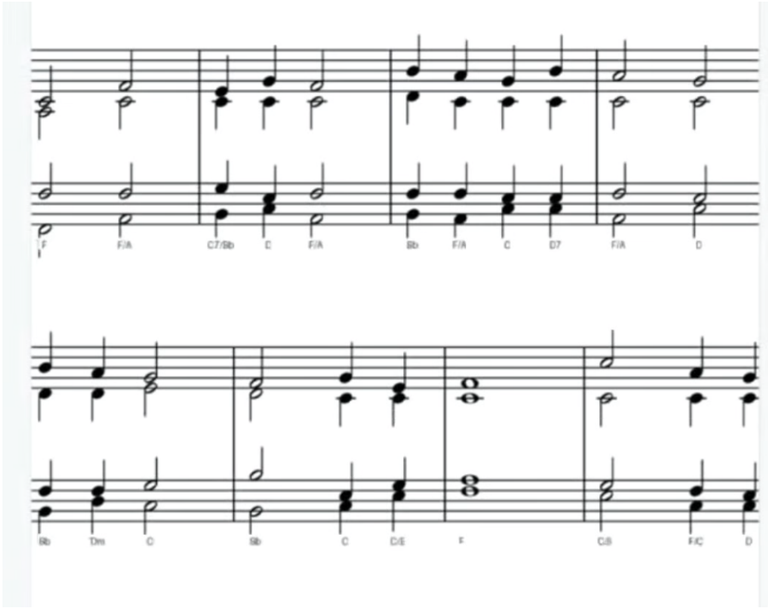
Harmony, counterpoint and form analysis produce reviewable outputs.

REVIEW

Musicians inspect and revise the analysis.



SCORE ANNOTATION / RECORDED OUTPUT



Candidate chord labels remain open to musical review.

Xunpu mobile + Xunpu Studio desktop. Actual recording. Displayed confidence is a system output.

A shared symbolic representation

INPUT

PDF / image
MusicXML / MIDI

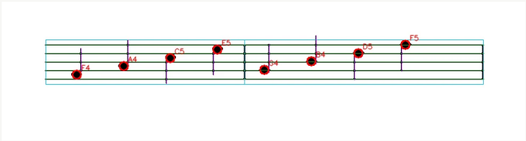
SYMBOLIC REPRESENTATION

PRODUCT ROUTE

Harmony / Counterpoint / Form

Reviewable annotations
AI-assisted explanation
Human review

SUPPORTING PERCEPTION RESEARCH



GraphOMR

Printed single-staff baseline.
Two fixtures, 16 synthetic variants.

INDEPENDENT RESEARCH / MGSA
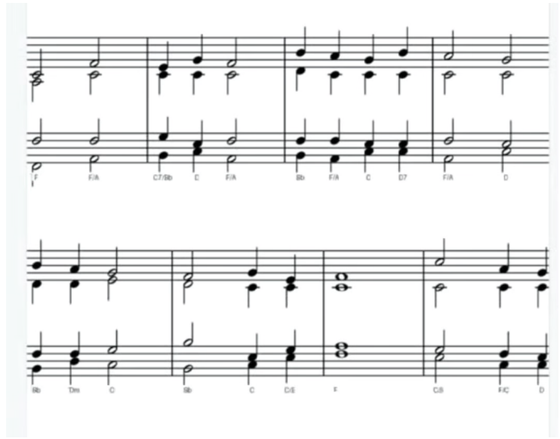
Musicological priors in attention

Measure-level boundary-detection pilot.
Separate from deployed product inference.

Solid: implemented product route. Dashed: independent research. OMR is outside the MGSA experiment.

Musical reasoning, made inspectable

HARMONY



Candidate chord names on the score

Inspect harmonic labels against the original notes.

COUNTERPOINT

```
student_homework = [
    [48, 55],  # 第一拍
    [50, 57]   # 第二拍（错误发生）
]

feedback = grader.check_parallel_fifth(student_homework)
for issue in feedback:
    print(issue)
```

Rule-based checks in code

Parallel-fifth checking route shown in the source presentation.

FORM



Measure ranges and section reports

Inspect the report before accepting a structural label.

Recorded outputs and code evidence. Candidate analysis requires human review. This page is not an accuracy benchmark.
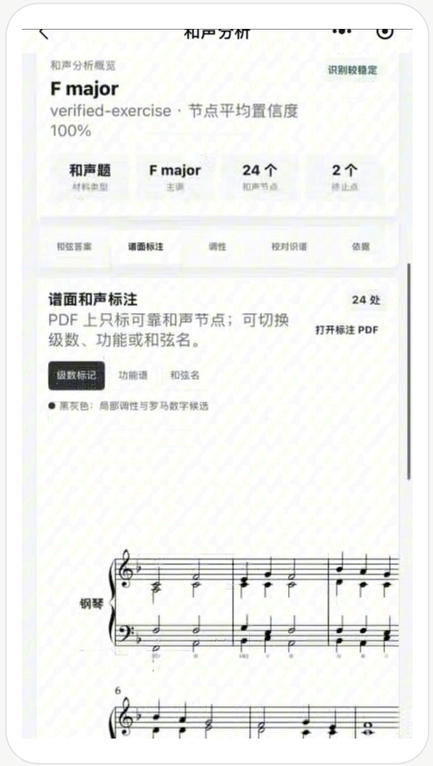
Xunpu Studio: from score to report

Actual mini-program recording / upload, annotations and structural reports

01 / UPLOAD



02 / HARMONY



03 / FORM



Read. Analyze. Review.

A score becomes a report and a reviewable musical object. Analysis remains open to human judgment.

50 participants

1,247 core-function calls

8 weeks

Frames from applicant-supplied demos. Displayed confidence is a system output. Closed-pilot figures are CV-reported.

Musicology-Guided Structural Attention

MEASURE-LEVEL BOUNDARY DETECTION / ONGOING RESEARCH

Can musicological priors improve structural boundary detection in symbolic music?

MUSICOLOGICAL PRIORS

SYMBOLIC SCORE	REPRESENTATION	MGSA ENCODER	OUTPUT
MusicXML	→ Note / measure embedding	→ Self-attention + priors	→ Boundary probability

$$S_{ij} = \frac{Q_i K_j^{\mathsf{T}}}{\sqrt{d_k}} + \lambda_m M_{ij} + \lambda_h H_{ij} + \lambda_c C_{ij} + \lambda_r R_{ij} + \lambda_t T_{ij} - \lambda_b B_{ij}$$

Supervision: phrase structural-ending annotations projected onto measure gaps. OMR is outside this experiment.

Encoding musical knowledge into attention

M MOTIF

Interval and duration similarity

H HARMONY

Coarse harmonic-function affinity

C CADENCE

Predicted cadential salience

R REPETITION

Transposition-tolerant recurrence

T METRIC

Similarity of metrical emphasis

B BOUNDARY

Penalty across score-derived closure cues

Content-based attention **+ musicological relations** **– boundary penalty**

Learned, nonnegative prior weights. These are operational proxies for musical relationships.

A matched-control pilot

Fold 0 / five seeds / five epochs / Mozart symbolic corpus

MODEL	EXACT BOUNDARY F1	±1 MEASURE F1
Mean-pooling Transformer	0.3133 ± 0.0164	0.4324 ± 0.0109
Matched pooling / no priors	0.3297 ± 0.0135	0.4588 ± 0.0136
MGSA / matched pooling + priors	0.3332 ± 0.0156	0.4569 ± 0.0317

PRELIMINARY / MIXED RESULT

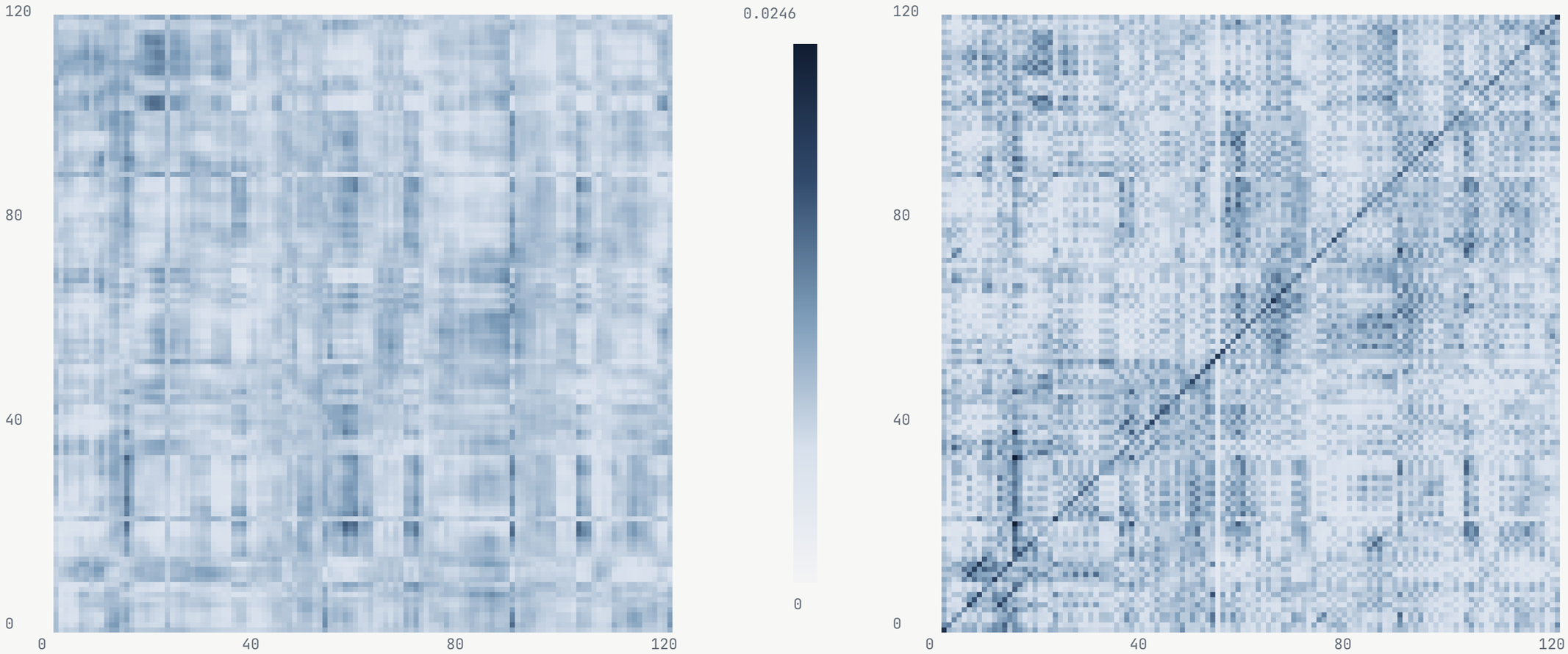
A small exact-boundary improvement. A slightly lower tolerant score.

No claim of consistent or significant gains. Mean ± seed SD. Next: grouped evaluation and error analysis.

Attention in a held-out score

MATCHED CONTROL / NO PRIORS

MGSA / MUSICOLOGICAL PRIORS



K.283-I, seed 42. Last layer, mean over heads. Axes: query/key measure (0-based). Attention is not proof of musical reasoning.

Algorithmic Awakening

AI-assisted audiovisual music / composition and sound design





Official awardees ↗

HKUST AI FILM FESTIVAL

Best AI Music Production Award

The 2nd HKUST AI Film Festival × MiniMax / 2026

Algorithmic Awakening

04:20 / external audio

Listen ↗

GENERATE

Custom serial / random-sequence code supplies material.

CURATE

I select fragments, register and density.

ARRANGE

I shape form, pacing and sound design.
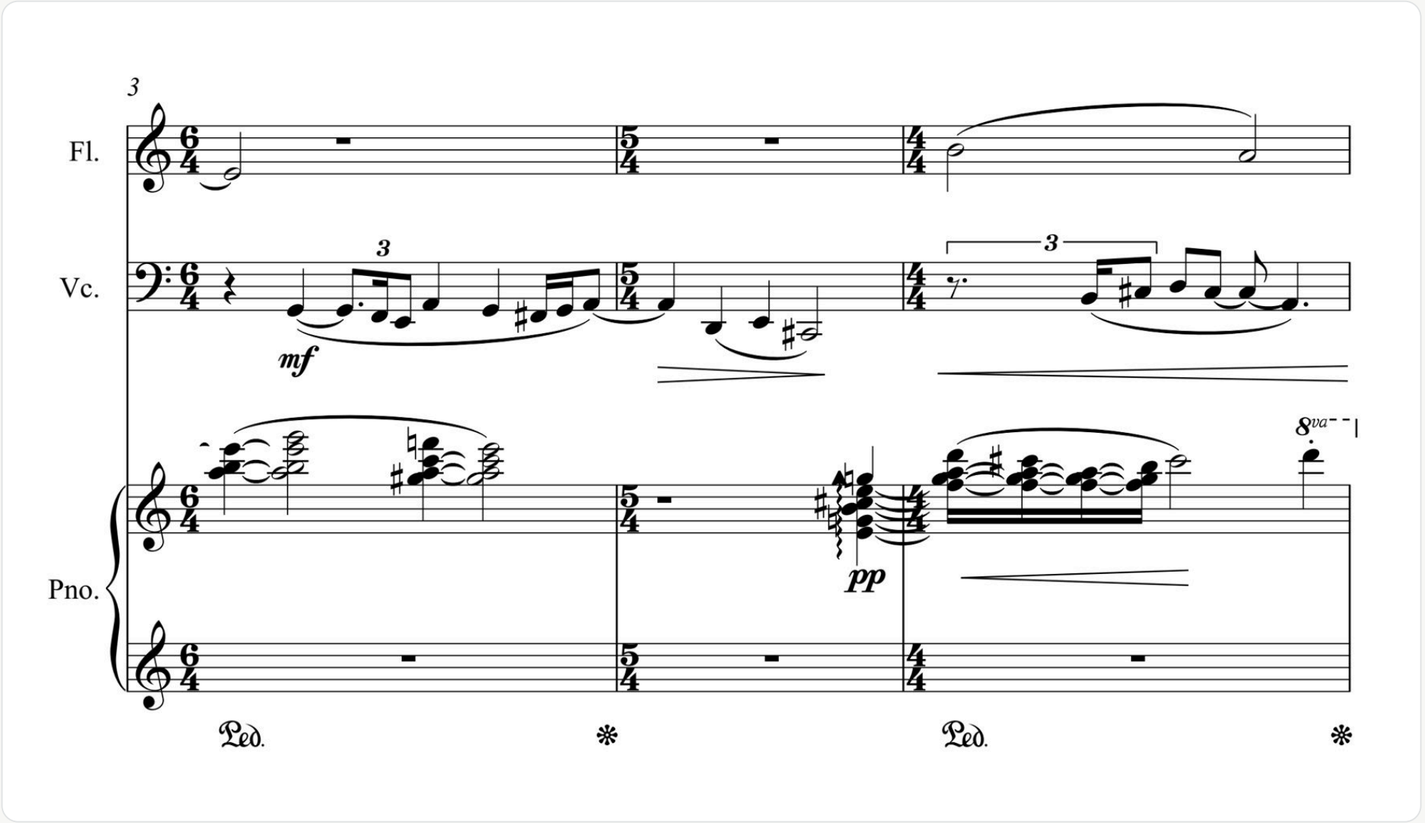
FINISH

I mix and master the final sound in Studio One.

Code proposes material. I remain responsible for musical form and final sound.

LING

Flute, cello and piano / 2024



Original score, measures 3–5. Estimated score duration: 6′22″. The award supports the work.

MEASURES 3–5

At measure 3, the cello enters beneath a sustained flute line. Metre changes from 6/4 to 5/4 and 4/4, while dense piano sonorities reshape the texture.

Full score ↗
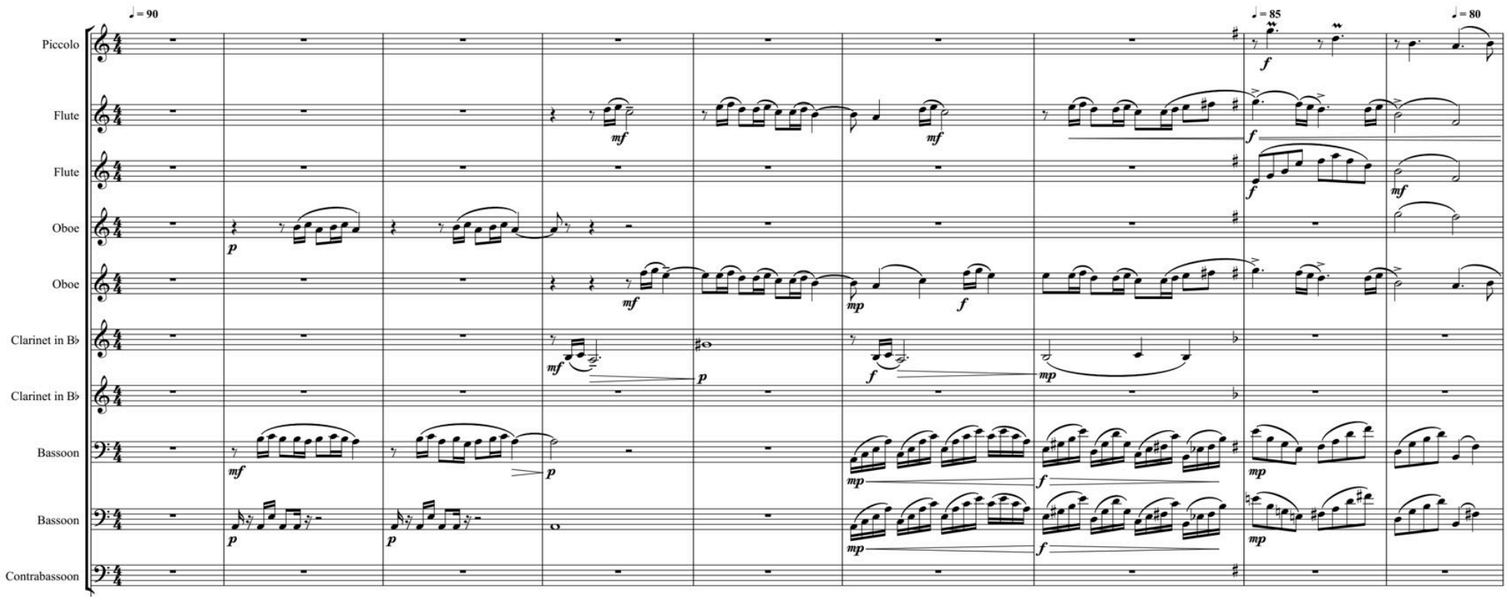


Gold Prize

Erik Satie International Music Competition
Innovative Writing / 2024

The Vanished Kingdom

Orchestral work / register, instrumental colour and articulation



OPENING WOODWINDS / ORIGINAL SCORE

The oboe figure expands into contrasting woodwind layers.

The Vanished Kingdom

08:25 / external audio

Listen ↗

Full score ↗

Gesture as an intention signal

MediaPipe Hands + Tone.js / separate chord-control prototype



CURRENT / RESEARCH

SYMBOLIC UNDERSTANDING

Xunpu system and MGSA pilot

PROTOTYPE

GESTURE CONTROL

Hand features and chord parameters

FUTURE

SCORE FOLLOWING

Timing, phrasing and performer intent

FUTURE

ADAPTIVE ACCOMPANIMENT

Responsive co-performance

Try prototype ↗

Idle interface. Score following and accompaniment remain future research.

Selected Engineering / Outputs

Xunpu / GraphOMR

Music intelligence and explainable OMR

Printed single-staff research baseline.

Repository ↗

ReelMesh Studio

Creative production systems

A local workbench with creator review.

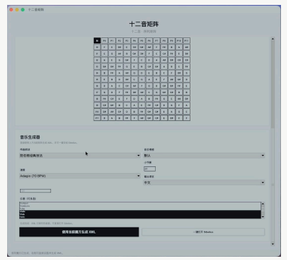
Repository ↗

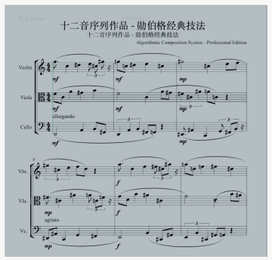
ShortDramaAgent

Structured audiovisual planning

Narrative events become editing and sound cues.

Repository ↗





TWELVE-TONE TOOL / ACTUAL DEMO

Matrix + generation controls MusicXML notation

Software copyright: 2026SR0486941

ONGOING RESEARCH

MGSA Manuscript in preparation

github.com/xudongdai-ai

Generated serial material supports compositional practice. ReelMesh–ShortDramaAgent integration remains planned.

Composition is my foundation.
Human intention is my next question.

PRACTICE
COMPOSITION

→

CURRENT
SYMBOLIC UNDERSTANDING

→

ONGOING
STRUCTURAL ATTENTION

-->

FUTURE
PERFORMER INTENTION

-->

FUTURE
HUMAN–AI CO-CREATION

Dai Xudong

Listening room ↗

daixudong.art

github.com/xudongdai-ai

shuaishuaidai2@gmail.com