

From Live Concert to Pedagogical Insight: Designing an Intelligent Concert Knowledge Navigation Panel

In Son Zeng¹

¹University of Macau

Abstract

Live concerts offer rich opportunities for situated music learning, yet most educational supports, such as program notes and pre-concert talks, are separated from the musical moments they explain. This paper proposes the **Concert Knowledge Navigation Panel**, a silent, AI-assisted interface that uses live audio recognition, score/program alignment, music data mining, and just-in-time pedagogical retrieval to support concert understanding without disrupting listening. The system is designed as a six-layer architecture: live audio sensing, real-time audio-to-score recognition and alignment, pedagogical knowledge retrieval, just-in-time explanation, learner/context adaptation, and concert-safe Human-Computer Interaction (HCI) presentation. The panel aims to locate the current performance within prepared and recognized learning materials and suggests educational cues when musically meaningful events occur. We present design goals, system features, and an evaluation plan for studying whether recognition-driven, low-distraction concert knowledge navigation improves audience orientation, music understanding, and post-concert reflection.

Keywords

Concert music education, just-in-time learning, score following, music information retrieval, human-computer interaction

1. Introduction and Related Work

Concerts are powerful but under-instrumented learning environments. While live concerts expose listeners to form, harmony, thematic recurrence, orchestration, harmonic tension, expressive timing, and performance gesture in real time, prior work on program notes and pre-concert explanatory formats suggests that listeners can benefit from contextual support when encountering unfamiliar live music [1, 2]. Traditional program notes and pre-concert talks provide valuable context, but they are usually separated from the precise musical moment they explain. The **Concert Knowledge Navigation Panel** therefore addresses a specific AIED and HCI problem: how can an AI-assisted panel enhance the live concert experience by transforming real-time musical events into timely, silent, and pedagogically meaningful cues without disrupting listening immersion?

Existing real-time concert note systems and orchestra-sector implementations demonstrate the practical feasibility of synchronized, silent contextual support during live performance [3, 4, 5, 6]. EnCue, for instance, provides real-time program notes on smartphones while the orchestra is performing [3]. ConcertCue similarly describes a synchronized streaming platform in which music-aligned information is delivered to audience devices during live performance [4]. Symphony Interactive / Octava provides an implementation precedent for Wi-Fi-delivered, music-synchronized information in orchestral settings [5], while orchestra-sector discussion frames real-time program notes as support for audience understanding of musical form, context, performer commentary, and unfamiliar terms during performance [6]. Together, these sources show that synchronized concert information is both technically feasible and educationally motivated, but they also reveal the next design challenge: how can we move from pre-authored synchronized concert cues toward recognition-based design, just-in-time pedagogical retrieval, and concert-safe adaptive explanation?

HCI, score-following, and music data mining together motivate the proposed architecture. Prior HCI work shows that mobile devices in live performance should be designed as part of the audience experience rather than treated merely as distraction, implying that concert interfaces must regulate attention

AI for Education Day, SIGKDD 2026, August 12, 2026, Jeju, Korea



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

rather than maximize interaction [7]. The technical foundation comes from real-time audio-to-score alignment and score-performance alignment research. Nakamura et al. address real-time audio-to-score alignment for performances containing errors, arbitrary repeats, and skips [8], while Peter et al. show that fine-grained automatic score-performance alignment remains critical for music information retrieval (MIR) that require note-level correspondences between scores and live performances [9]. These studies support the panel's design to locate live performances within known score/program structures rather than overclaim full real-time transcription. Music data mining provides the retrieval logic: instead of relying on audio, metadata, MIDI, scores, and tags for collection-level tasks, the panel reframes musical evidence as pedagogical event retrieval [10]. Recent symbolic MIR resources such as the *When in Rome* meta-corpus of functional harmony further show how encoded music knowledge can be organized at scale for computational analysis and retrieval [11]. The panel therefore retrieves educationally meaningful event during the concert and offers pedagogical cues to help the listener learn during concerts.

The Concert Knowledge Navigation Panel responds to this gap by treating the concert as a real-time educational environment rather than as a static cultural event supplemented by program notes. It offers a silent, individualized navigation panel that helps listeners understand where they are in the music, what to notice, and why the moment matters. Building on three presentation modes, including (1) Real-time Concert Progress, (2) Real-time Music Score Interface, and (3) Real-time Music Knowledge Interface, the proposed panel develops a recognition-driven AIED architecture: live audio and event sensing estimate the current performance progress; score/program alignment demonstrates current progress in a scrollable musical score; music data mining retrieves and suggests pedagogically relevant cues; and a low-attention interface presents them without interrupting listening. In this framing, Audio Music Recognition (AMR) and score following become pedagogical infrastructure, an intelligent timing mechanism for delivering explanations when they are musically meaningful [8, 9].

The AIED contribution lies in separating AMR, pedagogical retrieval, and concert-safe presentation. Intelligent tutoring research emphasizes adaptive systems that reason about learner needs, pedagogical goals, and feedback decisions [12], while broader AIED work argues for data-informed, learner-centered support rather than merely digitized instruction [13]. For concert music education, this means the panel should decide not only whether a recognized event is theoretically accurate, but also whether it is useful, timely, and appropriate to show during live listening. Prior work on explainable AI in education and human-AI interaction further motivates audience-sensitive explanations, uncertainty handling, user control, and calibrated trust [14, 15]. Accordingly, when recognition confidence is high, the panel should provide richer just-in-time cues; when confidence is low, it should enter minimal orientation mode, or defer explanation until a safer moment.

In sum, prior work establishes four foundations: synchronized concert-note systems show that silent real-time contextual support is practically feasible [3, 4, 5, 6]; HCI research shows that audience-facing devices must be integrated into the trajectory of live performance rather than treated as ordinary screen interaction [7]; score-following and score-performance alignment show how the progress of live concert can be located within known musical structures [8, 9]; and music data mining provides methods for organizing musical evidence into retrievable information resources [10, 11]. However, these findings have rarely been integrated into a concert-safe AIED architecture. Existing concert note systems are usually pre-authored and synchronize, and AIED systems rarely address live concert listening as a low-interruption, socially situated learning context. The Concert Knowledge Navigation Panel thus addresses this gap by treating concert learning as a real-time alignment problem between music, attention, and pedagogical explanation.

2. Proposed System Design

2.1. Design Goals

The Concert Knowledge Navigation Panel is organized around three design goals that jointly balance real-time music understanding, pedagogical effectiveness, and attention-aware interaction under varying

concert conditions. These goals are aligned with the four-phase evaluation framework in Section 3.

DG1: Real-Time Music-to-Instruction Grounding under Structured Musical Alignment. The system should personalize music instruction in the live performance moment by continuously estimating the current position within a known score structure. This includes identifying instructionally meaningful events such as theme returns, harmonic cadences, orchestration transitions, tempo fluctuations, and movement boundaries. Rather than assuming that all concert learning material has been prepared in advance, the panel adopts a **score-conditioned alignment paradigm**, in which AMR-based signals and symbolic score representations jointly support real-time localization. This design supports Phase 1 technical validation and Phase 2 pedagogical evaluation by ensuring that all instructional cues are musically grounded and temporally consistent.

DG2: Concert-Safe Mode-Conditioned Attention Management. The system should preserve the audience’s listening immersion while adapting pedagogical density based on user-selected modes and performance contexts. Interaction is designed to remain silent, glanceable, and minimally intrusive, with the system determining when to show musical cues, when to provide minimal orientation, and when to defer explanation to post-concert reflection. DG2 focuses on attention management, in which the interface supports attentive listening rather than competing with the stage. This design supports Phase 3 evaluation of cognitive load, perceived distraction, and engagement across *Live Attentive*, *Live Educational*, *Learning-at-Home Attentive*, and *Learning-at-Home Educational* modes, as well as Phase 4 evaluation of robustness under incremental orchestral complexity.

DG3: Trustworthy, Explainable Human-AI Pedagogical Adaptation. The system should support transparent pedagogical reasoning through uncertainty-aware inference, lightweight user adaptation, and explicit explanation mechanisms. A dedicated Pedagogy Button enables users to inspect cue generation logic, including alignment confidence, triggering musical features, and conceptual mapping. When confidence is high, richer content may be shown; when confidence is low, the panel should reduce specificity and enter minimal orientation mode. This design supports Phase 3 evaluation of explainability and Phase 4 expert validation of cue accuracy in real concert conditions.

2.2. System Features and Prototype Mockup

The Concert Knowledge Navigation Panel is proposed as a six-layer AIED architecture that transforms live audio input into structured pedagogical knowledge while enabling mode-aware interaction, offering pedagogical reasoning and logging across all system components.

Layer 1: Live Audio, Event, and Context Sensing. This layer provides the perceptual foundation of the system by collecting real-time audio signals and contextual information required to concert understanding. It integrates venue audio feeds, device microphone inputs, and synchronized concert signals with metadata such as concert programs, piece order, movement structure, performer information, and available score materials. Instead of attempting complete musical transcription at every moment, this layer extracts robust and low-latency cues, including beat positions, note onsets, tempo variations, section-transition indicators, silence, applause, and movement boundaries. These cues establish the temporal reference required for score alignment and pedagogical cue generation. By transforming continuous performance signals into structured musical events, this layer enables the downstream modules to reason about the current state of a concert and determine appropriate educational interactions.

Layer 2: Real-Time AMR and Score/Program Alignment. This layer aligns incoming audio with prepared resources such as MusicXML, MEI, annotated PDF scores, rehearsal marks, movement order, thematic landmarks, and program notes. Its purpose is to convert live audio into a musically meaningful position estimate, such as opening theme, transition, development, recapitulation, or movement boundary. It supports DG1 by linking explanations in the current piece and performance moment. It supports DG3 by explicitly modeling uncertainty: when confidence is high, the system can support richer prompts; when confidence drops during rubato, cadenzas, applause, repeated sections, or improvisation, it enters minimal orientation mode.

Layer 3: Music Knowledge Representation and Pedagogical Retrieval. This layer retrieves explanation units indexed by score position, section type, theme, concept type, listener profile, display

length, and confidence threshold. The knowledge base includes program-note facts, composer and historical context, instrumentation cues, thematic labels, harmonic functions, listening prompts, and post-concert reflection questions. It supports DG1 by ensuring that each prompt is piece-grounded and triggered by a relevant musical landmark. It supports DG2 by retrieving short, actionable, display-ready units suitable for silent use during a live concert.

Layer 4: Pedagogical Orchestration and Just-in-Time Explanation Engine. This layer adaptively determines whether to intervene, what to show, how much to show, and when to delay. It is the core AIED decision layer of the system. Rather than presenting all available musical knowledge, it selects only prompts that are instructionally relevant, temporally appropriate, and safe for the concert context. During dense passages, it may show only a brief phrase; between movements, it may provide a richer summary. It supports DG1 through context-sensitive explanation, DG2 through low-distraction timing and brevity, and DG3 through confidence-aware decisions and adaptive prompt selection.

Layer 5: Learner and Context Adaptation. This layer maintains a lightweight and interpretable model of learner characteristics, situational context, and user-selected interaction preferences. Rather than relying on invasive profiling, the system allows users to explicitly select and adjust their preferred experience mode, such as *Attentive Listening Mode* for immersive concert engagement and *Educational Mode* for guided musical learning and analysis. In *Attentive Listening Mode*, the system prioritizes minimal and non-disruptive orientation cues such as “the main theme returns,” while in *Educational Mode*, the system provides richer analytical explanations such as “dominant seventh harmony leads to imperfect cadence” together with reflective prompts that supports post-concert discussion and deeper musical understanding. By combining explicit user control with context-aware adaptation, it supports DG2 by reducing unnecessary cognitive and visual burden, and DG3 by enabling personalization, user agency, and feedback-driven interaction.

Layer 6: Concert-Safe HCI Presentation and Reflection. This layer presents knowledge through low-brightness visual prompts, optional haptic cues, progress arcs, icons, short text, score fragments, and post-concert review cards. The default view is a glanceable Concert Progress tab, supported by Knowledge and Score tabs when users want more detail. It supports DG2 by preserving the audience’s attention to the live performance through silent and peripheral interaction, and DG3 by enabling review cards, and post-concert reflection, turning temporary concert prompts into reusable learning materials.

Prototype Mockup. The prototype instantiates the proposed architecture through a mode-driven interface designed to support different listening goals and learning contexts (see <https://curved-tacky-86410009.figma.site/>). The system provides four adaptive experience modes: *Live Attentive Mode*, *Live Educational Mode*, *Learning-at-Home Attentive Mode*, and *Learning-at-Home Educational Mode*. These modes represent different configurations of information density, pedagogical support, and user engagement. After initialization, the audience-facing interface visualizes the detected performance progress and structural landmarks generated from the sensing and alignment layers. During live concert conditions, the interface foregrounds a minimal Live Cue Panel that displays AMR-aligned musical position estimates and highlights structural events. In *Learning-at-Home* modes, particularly, the system transitions to a structured exploration interface where pedagogical events are reorganized into a navigable timeline synchronized with score representations. Through these adaptive modes, the prototype demonstrates how real-time sensing, musical alignment, and pedagogical orchestration can jointly support both audience experience and music learning.

3. Research Design and Evaluation Plan

The evaluation of the Concert Knowledge Navigation Panel follows a four-phase progressive validation framework that incrementally tests system functionality, pedagogical effectiveness, interaction behavior, and system stability under increasingly realistic conditions. This staged design ensures that technical robustness, learning impact, and real-world usability are systematically evaluated from controlled simulation to ecological deployment.

In Phase 1, the system is evaluated under simulated concert conditions using recorded musical

excerpts and annotated scores. This phase focuses on technical validation, including AMR alignment accuracy, score synchronization stability, cue timing precision, and system latency. The objective is to establish the reliability of real-time musical position tracking and pedagogical event triggering under controlled and repeatable conditions.

Phase 2 evaluates pedagogical effectiveness through a between-subject experiment comparing three conditions: a baseline listening condition without system support, a traditional program note condition, and the full Concert Knowledge Navigation Panel condition. Participants engage with short concert stimuli and are evaluated on listening awareness, structural comprehension, perceived listening satisfaction, and post-concert retention. This phase examines whether real-time concert pedagogical augmentation improves music understanding while preserving the listening experience.

Phase 3 adopts a within-subjects experimental design to evaluate user interaction behavior across four conditions: Live Attentive, Live Educational, Learning-at-Home Attentive, and Learning-at-Home Educational modes. This phase measures cognitive load using NASA-TLX, perceived distraction, learning gains, and engagement depth. We assess learning outcomes through pre- and post-tests of conceptual understanding of music theory and structure within the concert repertoire. In addition, we analyze interaction behavior from system logs, including cue-click frequency, pedagogy button usage, mode-switching patterns, score-interaction trajectories, and review-card engagement. Semi-structured interviews further examine perceived immersion during the concert, trust in pedagogical cues, and perceived naturalness of each mode.

Finally, Phase 4 evaluates system performance in real orchestral concert environments. This phase examines system robustness under real-world conditions, including multiple movements, applause transitions, conductor-dependent tempo variation, and complex orchestration. Technical measures include AMR alignment stability, cue latency, and mode switching reliability under live conditions, while behavioral and qualitative data include interaction logs, dwell time patterns, post-concert questionnaires, expert evaluation of cue correctness, and semi-structured audience interviews. This phase provides ecological validation of the system’s capability of functioning as a real-time pedagogical augmentation tool in authentic concert settings.

Across four phases, the evaluation operationalizes the system’s core design assumptions. Phase 1 and Phase 2 validate audio-score alignment reliability and pedagogical effectiveness under controlled conditions, Phase 3 examines mode-dependent user learning and interaction behavior, and Phase 4 tests ecological robustness in live orchestral environments with temporal and acoustic complexity. Together, the results support the central claim that reliable concert understanding does not require full Audio2Music transcription. Instead, stable music score alignment combined with pedagogical retrieval is sufficient to deliver timely, meaningful, and explainable music learning experiences in live concert settings.

4. Conclusion and Future Work

This paper introduced the Concert Knowledge Navigation Panel, a recognition-driven AIED architecture for transforming live concerts into silent, situated learning environments. Its novelty lies not simply in using AI during performance, but in engineering situated explanation, uncertainty-aware alignment, and human-centered educational value rather than universal transcription. By separating recognition, pedagogical retrieval, explanation orchestration, and concert-safe presentation, the panel reframes concert education as a real-time alignment problem between music, attention, and pedagogical explanation.

Future work will develop and evaluate a prototype using public-domain solo, chamber, or orchestral repertoire with prepared landmarks and educational cues. Usability testing with concert planners and music educators will use think-aloud protocols to assess feasibility, concert preparation workflows, and user experience. Beta testing with experienced audiences will examine system flow, visibility, and perceived non-intrusiveness. Audience questionnaires and interviews will evaluate perceived helpfulness, orientation, engagement, and post-concert reflection. Overall, the Concert Knowledge

Navigation Panel supports concert learning by enabling users to navigate concert progress, view real-time scoring cues, and gain conceptual knowledge about the ongoing concert, contributing to AIED, HCI, music education, and MIR communities.

Declaration on Generative AI

During the preparation of this work, the author used ChatGPT by OpenAI for structural refinement, citation formatting, grammatical correction and LaTeX formatting support. The author reviewed and edited all generated content and takes full responsibility for the final manuscript.

References

- [1] D. Bennett and J. Ginsborg, "Audience reactions to the program notes of unfamiliar music," *Psychology of Music*, vol. 46, no. 4, pp. 588–605, 2018, doi: 10.1177/0305735617721339.
- [2] M. Wald-Fuhrmann, L. Egermann, L. Czepiel, R. O'Neill, and C. Weining, "Concert formats influence how audiences experience live classical music," *Humanities and Social Sciences Communications*, vol. 12, Art. 110, 2025, doi: 10.1057/s41599-025-06461-9.
- [3] Des Moines Symphony, "Virtual Program Notes from EnCue," 2020. [Online]. Available: <https://dmsymphony.org/about/news/virtual-program-notes-from-encue/>
- [4] MIT Music Technology Lab, "ConcertCue," n.d. [Online]. Available: <https://musictech.mit.edu/concertcue/>
- [5] UMBC Imaging Research Center, "Symphony Interactive," n.d. [Online]. Available: <https://irc.umbc.edu/symphony-interactive/>
- [6] S. Brown, "Notes on the Go," *Symphony*, Spring 2018. [Online]. Available: <https://symphony.org/wp-content/uploads/2024/08/Notes-on-the-Go.pdf>
- [7] C. Greenhalgh, A. Hazzard, S. Benford, and L. Cliffe, "Crafting trajectories of smart phone use at the opera," *ACM Transactions on Computer-Human Interaction*, vol. 29, no. 6, pp. 1–37, 2022, doi: 10.1145/3531007.
- [8] T. Nakamura, E. Nakamura, and S. Sagayama, "Real-time audio-to-score alignment of music performances containing errors and arbitrary repeats and skips," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 2, pp. 329–339, 2015, doi: 10.1109/TASLP.2014.2381276.
- [9] S. D. Peter, S. Chowdhury, A. Maezawa, and G. Xia, "Automatic note-level score-to-performance alignments in the ASAP dataset," *Transactions of the International Society for Music Information Retrieval*, vol. 6, no. 1, pp. 99–108, 2023, doi: 10.5334/tismir.149.
- [10] T. Li, M. Ogihara, and G. Tzanetakis, Eds., *Music Data Mining*. Boca Raton, FL, USA: Chapman & Hall/CRC, 2012.
- [11] M. Gotham, G. Micchi, N. Napoles Lopez, and M. Sailor, "When in Rome: A meta-corpus of functional harmony," *Transactions of the International Society for Music Information Retrieval*, vol. 6, no. 1, pp. 150–166, 2023, doi: 10.5334/tismir.165.
- [12] B. P. Woolf, *Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing E-Learning*. Burlington, MA, USA: Morgan Kaufmann, 2009.
- [13] R. Luckin, W. Holmes, M. Griffiths, and L. B. Forcier, *Intelligence Unleashed: An Argument for AI in Education*. London, UK: Pearson, 2016.
- [14] H. Khosravi, S. B. Shum, G. Chen, C. Conati, Y.-S. Tsai, J. Kay, S. Knight, R. Martinez-Maldonado, S. Sadiq, and D. Gasevic, "Explainable artificial intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 3, Art. 100074, 2022, doi: 10.1016/j.caeai.2022.100074.
- [15] S. Amershi et al., "Guidelines for human-AI interaction," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, doi: 10.1145/3290605.3300233.