

MEHEARSAL

AI-Powered Virtual Band

PROJECT BACKGROUND

Digital music practice tools are increasingly important, but many existing solutions still require a trade-off between accessibility and technical capability. Professional-grade audio systems often depend on installation, complex setup, or desktop software, while browser-based tools are easier to access but usually limited in low-latency performance and real-time interaction. This creates a gap for musicians who need flexible, responsive, and low-barrier rehearsal tools.

The MEHEARSAL project, developed with Muse Scene Lab in Spain, addresses this gap by creating a browser-based rehearsal and practice environment for musicians who want to rehearse solo with a virtual band. The system is designed to support real-time control of tempo, dynamics, volume, and instrument isolation through gesture and voice, providing a more adaptive alternative to static backing tracks and conventional practice software.

KEY FEATURES

WEB AUDIO ENGINE

- **Multi-Instrument Synthesis:** Rhodes MK-80 (modal synthesis), Electric Bass and Acoustic Guitar (digital waveguide/Karplus-Strong modelling), alongside existing sample-based instruments
- **Real-Time WASM Performance:** JUCE/C++ audio engine compiled via Emscripten, running at near-native speed with no garbage-collection audio dropouts
- **Low-Latency Pipeline:** AudioWorklet renders audio on a dedicated thread, targeting sub-200ms round-trip latency
- **Tempo-Safe Playback:** Real-time BPM time-stretching with no audible artifacts, even down to 75% speed
- **Stable Polyphony:** Supports simultaneous voices per instrument (Rhodes: 16, Bass: 4, Guitar: 8) without degradation

MOTIVATION

The increasing demand for accessible, collaborative music tools highlights a critical gap: modern web browsers lack the infrastructure to support professional-grade, real-time audio performance coupled with intuitive gesture-based interaction. MEHEARSAL addresses this by developing an integrated, immersive music practice platform entirely within the browser, combining a high-fidelity WebAudio engine with real-time hand gesture recognition.

The Mehearsal Core project comprises two coordinated components: (1) a WebAssembly compilation pipeline using Emscripten to extend an existing proprietary C++ audio architecture, delivering a thread-safe AudioWorklet processor capable of synthesizing multiple instruments with seamless tempo manipulation, and (2) a client-side gesture recognition system built entirely in WASM using machine learning inference, enabling musicians to control playback and performance through intuitive hand gestures. Both systems communicate via a unified MIDI protocol, creating a cohesive human-computer musical interaction experience.

HAND TRACKING

- **Touch-Free Control:** Musicians control tempo, muting, and session state without looking away from their instrument
- **Client-Side Gesture Recognition:** MediaPipe Hand Landmarker + WASM-compiled inference, fully on-device with no cloud dependency
- **Robust Gesture Set:** Static gestures (fist, open palm, index point) and temporal gestures (wave), with confidence/debounce filtering against false triggers
- **Lighting & Device Invariance:** Landmark normalization (translation, scale, rotation) keeps recognition consistent across setups
- **Direct Audio Integration:** Recognized gestures dispatch straight into the audio engine via a shared MIDI command interface

