

# Designing Persistent Musical Metaverses: A Multi-Dimensional Framework for Interaction, Collaboration and Creative Space Organization

Michel Buffa, Adam Mir-Sadjadi, Samuel Demont, Yassine Qat, Pierre Constantin, Florent Robert, Marco Winckler  
Université Côte d'Azur  
France  
firstname.lastname@univ-cotedazur.fr

**Abstract**—Persistent Musical Metaverses are immersive virtual environments where musicians create, collaborate, and develop musical projects over extended periods. While existing Musical XR systems have primarily focused on immersive interaction and virtual instruments, they provide limited support for organizing collaborative activities and ensuring the long-term continuity of creative work. This paper proposes a multi-dimensional framework for designing Persistent Musical Metaverses, structured around three complementary design dimensions: Interaction, Collaboration, and Creative Space Organization. Unlike these functional dimensions, Persistence is treated as a transversal design principle preserving their state across successive sessions and enabling creative continuity over time. The framework is instantiated through *WAM Jam Party*, an open-source WebXR platform based on the Web Audio Modules (WAM) plugin proposal. The Interaction dimension supports modular immersive instruments and expressive gesture-sound mappings; the Collaboration dimension introduces workspace awareness, acoustic territories, and configurable collaborative roles; and the Creative Space Organization dimension structures persistent identities, projects, reusable collaborative sessions, and sharing mechanisms supporting long-term creative communities. A formative qualitative evaluation was designed and conducted with musicians and instructors from the Nice Conservatory of Music as well as Computer Science teachers to assess the artistic and pedagogical relevance of the proposed framework. Beyond its implementation, this work argues that future Musical Metaverses should be designed by jointly considering interaction, collaboration, creative space organization, and transversal persistence as complementary foundations.

## I. INTRODUCTION

The emergence of immersive technologies has considerably expanded the possibilities for collaborative music-making. Virtual and Augmented Reality, combined with low-latency networking and Web technologies, enable musicians to interact within shared virtual environments where they can perform, compose, rehearse, and experiment together regardless of their physical location. These developments have contributed to the broader vision of Persistent Musical Metaverses: immersive virtual environments supporting musical creation, performance, education, and social interaction over extended periods [29].

Existing Musical XR research has explored numerous aspects of immersive musical interaction, including Digital Musical Instruments, gesture-sound mappings, collaborative

performance, spatial audio, and networked musical experiences [2], [20], [21], [26]. These systems demonstrate that Virtual Reality can support expressive musical interaction and collaborative performance. However, most existing approaches primarily focus on real-time interaction and provide limited support for organizing creative activities over extended periods. Mechanisms such as persistent identities, reusable projects, collaborative workspaces, permissions, and long-term creative continuity remain comparatively underexplored.

Yet musical practice rarely consists of isolated performance sessions. Composition, rehearsal, teaching, and artistic research evolve through successive sessions involving changing participants and reusable creative resources. Persistent Musical Metaverses should therefore support not only immersive musical interaction, but also the organization and continuity of collaborative creative activities [5]. Achieving such continuity requires preserving not only musical artifacts but also the evolving collaborative context — identities, permissions, acoustic territories, workspace organization... — transparently and continuously across distributed clients and successive sessions.

These observations motivate the following research question:

*How can we support creativity and collaboration among users of a musical metaverse when their activities are performed across multiple sessions?*

To address this question, we propose a multi-dimensional design framework organized around three complementary functional dimensions and illustrate its application through the open-source WebXR platform *WAM Jam Party*. The Interaction dimension addresses immersive musical instruments, modular audio processing, and expressive gesture-sound mappings. The Collaboration dimension introduces mechanisms supporting coordinated musical activities through workspace awareness, acoustic territories, and configurable collaborative roles. The Creative Space Organization dimension structures persistent identities, projects, reusable collaborative sessions, and sharing mechanisms supporting long-term creative communities. Unlike these functional dimensions, Persistence is treated as a transversal design principle ensuring the transparent and continuous preservation of creative activities across successive sessions.

The proposed framework is instantiated through *WAM Jam Party*, an open-source WebXR platform built upon the Web Audio Modules (WAM) [8], [9] plugin proposal. The platform is publicly available online<sup>1</sup> and its source code is released on GitHub<sup>2</sup>. *WAM Jam Party* serves both as a proof of concept and as an extensible research testbed for Persistent Musical Metaverses, allowing researchers, musicians, educators, and developers to experiment with collaborative Musical XR scenarios and extend the platform through interoperable WAM-compatible modules.

The main contributions of this paper are:

- a multi-dimensional framework for designing Persistent Musical Metaverses, distinguishing three functional dimensions — Interaction, Collaboration, and Creative Space Organization — together with Persistence as a transversal design principle.
- the implementation of this framework within *WAM Jam Party*, an open-source, modular, and extensible WebXR platform based on interoperable Web Audio Modules;
- collaborative mechanisms including workspace awareness, acoustic territories, configurable roles, reusable creative spaces, and long-term project organization;
- a formative qualitative evaluation assessing the artistic and pedagogical relevance of the proposed framework.

## II. RELATED WORK

The design of Persistent Musical Metaverses draws upon several complementary research communities, including Networked Music Performance (NMP), New Interfaces for Musical Expression (NIME), Collaborative Virtual Reality (CVR), Computer-Supported Cooperative Work (CSCW), and the Internet of Musical Things (IoMusT). While these communities have produced significant advances in immersive musical interaction, collaborative virtual environments, and connected musical ecosystems, they have largely addressed these topics independently. In contrast, this work proposes a unified design framework integrating interaction, collaboration, and creative space organization, while considering persistence as a transversal design principle spanning these complementary dimensions.

### A. Interaction: immersive musical instruments and modular creativity

Research on Musical XR has extensively explored immersive Digital Musical Instruments, gesture-sound mappings, and embodied interaction [2], [20], [21], [28]. Systems such as *Virtuoso* [26] demonstrate the expressive potential of virtual instrument performance, *EXA* [1] explores continuous gestural control within immersive environments, while *PatchWorld* [24] and *OpenSoundLab* [32] illustrate modular music creation in social virtual spaces. Complementary work has investigated physics-based musical interaction [15], hand tracking [18], audience participation [17], and mixed-reality collaboration [31].

Together, these systems demonstrate the richness of immersive musical interaction and provide valuable design principles for expressive instrument design.

### B. Collaboration: coordinating musical activities in shared virtual environments

As Musical XR systems evolve toward shared virtual environments, collaboration becomes a central design challenge. Beyond supporting simultaneous interaction, collaborative musical activities require participants to coordinate editing actions, negotiate access to shared resources, remain aware of ongoing activities, and dynamically organize responsibilities as creative work evolves. Unlike conventional collaborative editing systems, musical collaboration additionally involves coordinating auditory perception, spatial organization, and collective performance.

These challenges have long been investigated within the CSCW community through concepts such as workspace awareness, territoriality, role management, and shared workspace coordination [6], [13], [16], [19], [27]. While originally developed for collaborative document editing and cooperative work, these concepts provide a valuable theoretical foundation for organizing interaction in immersive collaborative environments.

Recent Musical XR research has begun adapting these principles to collaborative music-making. Men et al. demonstrate how fluid acoustic boundaries improve collaborative creativity compared with rigid spatial partitions [22], while we reviewed the literature and discussed configurable collaborative roles and participatory design approaches for immersive musical environments [23]. Recent co-design work by Boem et al. further shows that collaborative VR instrument design and spatial layout are tightly interdependent, with participants favoring VR-native interactions, mutual visibility, and tiered spatial arrangements for mixed-skill ensembles [4]. Together, these contributions show that effective collaboration in Musical Metaverses depends not only on shared interaction, but also on explicit organizational mechanisms regulating participation, perception, and coordination.

### C. Creative Space Organization

While interaction and collaboration have received considerable attention, the long-term organization of musical activities within immersive environments remains comparatively underexplored.

Turchet et al. introduced the Internet of Musical Things as a vision of connected musical ecosystems where instruments, performers, and services interact through network infrastructures [30]. This perspective naturally extends toward Persistent Musical Metaverses, where identities, creative projects, collaborative sessions, and musical artifacts remain available beyond synchronous interactions.

Commercial cloud-based music platforms (DAWs) already support persistent collaborative projects, but remain primarily desktop-oriented, asynchronous, and non-immersive.

Recent empirical work by Ppali et al. [25] provides an important complementary perspective. Through interviews,

<sup>1</sup><https://wamjamparty.i3s.univ-cotedazur.fr/>

<sup>2</sup><https://github.com/doriangirard9/musical-multiverse-vr>

surveys, and ethnographic observations of Social VR music communities, the authors describe how performers, organizers, world builders, and audiences continuously reconstruct collaborative spaces across fragmented platforms. Their study identifies the lack of persistent collaborative infrastructures as one of the principal barriers to sustainable creative communities.

These observations suggest that persistence should not be viewed merely as a storage mechanism, but rather as an enabling property supporting the long-term organization of creative spaces, identities, projects, and communities.

#### D. Research gap

Taken together, the literature converges toward three complementary research directions: immersive musical interaction, collaborative coordination, and the organization of persistent creative spaces. Existing systems generally make significant contributions to one or two of these aspects, yet few integrate them within a unified design framework. To address this gap, we propose a framework integrating interaction, collaboration, creative space organization, and long-term persistence within a unified design model.

### III. WAM Jam Party: INSTANTIATING A MULTI-DIMENSIONAL FRAMEWORK FOR PERSISTENT MUSICAL METAVERSES

The proposed multi-dimensional framework is instantiated through *WAM Jam Party*, an open-source WebXR platform integrating immersive musical interaction, collaborative synchronization, creative space organization, and persistent services within a unified software architecture. Figure 1 summarizes the conceptual framework. For details about the design and implementation of previous versions of *WAM Jam Party*, see [10], [11]. Unlike conventional software architectures, the proposed dimensions should not be interpreted as execution layers or software modules. Instead, they represent complementary design concerns that are simultaneously addressed throughout the platform.

#### A. Interaction dimension

The Interaction dimension provides a modular design space where expressive gesture-sound mappings, reusable Web Audio Modules, and multimodal interaction techniques can be combined to create new forms of musical expression. This dimension constitutes the foundation upon which collaborative and persistent creative activities are subsequently built.

Unlike conventional Digital Audio Workstations (DAWs), where musical structures are primarily organized along a temporal timeline, *WAM Jam Party* adopts a spatial modular paradigm. Musical entities are represented as three-dimensional objects that users can manipulate directly in the shared virtual environment and interconnect through virtual cables to construct modular audio graphs as seen in Figure 2. This approach follows the long tradition of modular synthesis while extending it to immersive collaborative environments.

The platform is built upon WAMs [8], which provide a common abstraction for musical plugins independently of their

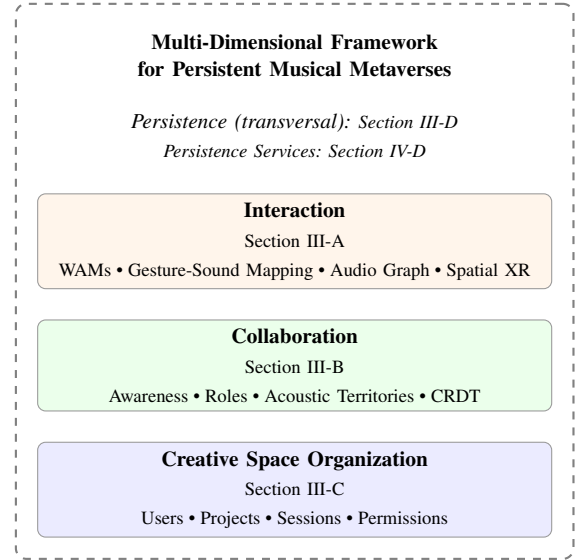


Figure 1. Proposed multi-dimensional framework for Persistent Musical Metaverses.

internal implementation. Rather than being merely a software architecture, this modularity enables heterogeneous musical entities—including synthesizers, DSP effects, AI-assisted generators, parameter controllers, spatial audio renderers, and visualization modules—to coexist within the same immersive environment. The current implementation already provides more than seventy interoperable WAM modules spanning these categories and illustrating the extensibility of the proposed framework. Figure 3 summarizes the conceptual organization of this dimension.

1) *Gesture-sound mappings as the foundation of immersive instruments*: Virtual Reality considerably enlarges the interaction space available to Digital Musical Instruments. Beyond conventional buttons and keyboards, immersive environments provide continuous input modalities including hand position, orientation, velocity, gaze, body posture, and voice. As emphasized by Hunt et al., the mapping layer—neither the controller nor the synthesis engine—constitutes the defining element of a digital musical instrument because it determines its expressiveness and playability [20]. Jordà [21] further describes this process as a form of *digital lutherie*, where instrument design consists primarily of creating meaningful relationships between human gestures and musical responses. Birnbaum et al. [2] similarly advocate analyzing instruments through multidimensional spaces describing interaction rather than hardware alone.

2) *Mapping properties*: Building upon the principles mentioned earlier, we proposed a semi-formal characterization of immersive instruments as collections of gesture-sound relationships [14], describing user inputs, musical parameters, and mapping properties. In *WAM Jam Party*, this characterization provides a systematic methodology for extending existing instruments and designing new immersive interaction techniques.

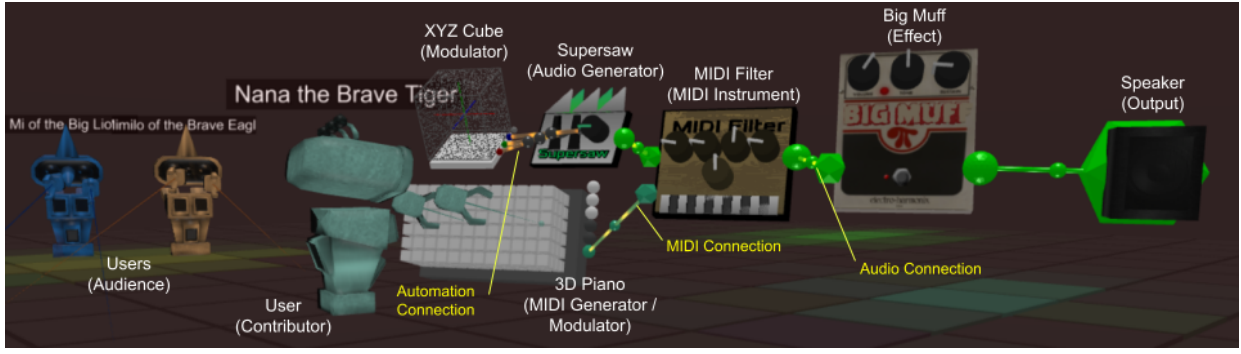


Figure 2. Typical audio chain in WAM Jam Party. Modular musical components (Web Audio Modules) are interconnected to form a collaborative audio graph. The avatar of a remote participant is visible in the shared virtual environment, manipulating a 3D Piano note generator, while two users attend the performance in the background. All avatars are represented as colored robots with their nicknames above their heads.

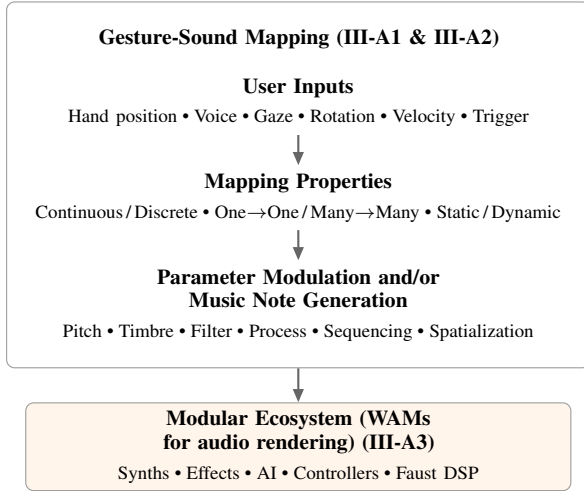


Figure 3. Interaction dimension. Multimodal user inputs pass through gesture-sound mappings to produce parameter modulation and/or music note generation for reusable WAMs organized as modular ecosystems.

3) *A modular, extensible ecosystem of interaction and visualization modules:* WAM Jam Party organizes musical functionality as an ecosystem of interoperable WAM plugins. All WAMs expose standardized audio, MIDI, and automation interfaces. Thanks to this shared API, synthesizers, DSP effects, AI generators, controllers, visualization modules, and spatial audio renderers can be freely combined within the same collaborative audio graph. Rather than selecting predefined instruments, musicians progressively construct reusable musical systems by assembling modular components adapted to immersive three-dimensional environments.

WAM Jam Party has been developed incrementally since 2024 through successive research contributions introducing new interaction techniques. Recent additions include immersive instruments (e.g., 3D Piano, 3D Harp, and Virtual Drumkit [12]), continuous spatial controllers inspired by the Theremin, gaze- and voice-controlled interfaces [14], new audio effects, parameter controllers, reusable DSP components and AI-assisted MIDI generation components based on Google’s

Magenta library. Together, these modules support assisted composition, improvisation, rhythmic pattern generation, and musical variation directly within collaborative modular audio graphs. Rather than prescribing a fixed collection of instruments, the framework supports the continuous evolution of immersive musical interaction as new WAM-compatible components become available.

A curated YouTube playlist<sup>3</sup> documents many of these recent additions and illustrates the continuing evolution of the platform.

### B. Collaboration dimension

While the Interaction dimension focuses on individual musical expression, the Collaboration dimension addresses how multiple participants coordinate their activities within shared virtual environments. Beyond supporting simultaneous interaction, collaborative Musical Metaverses require mechanisms for coordinating shared activities while preserving both creative flexibility and individual autonomy. These challenges have been extensively studied in CSCW, which provides conceptual foundations for coordinating shared workspaces, regulating access to common resources, and supporting group awareness [6], [16], [19], [27].

Musical Metaverses introduce sound itself as a shared organizational resource. Collaboration therefore concerns not only who manipulates musical objects, but also who can hear, access, and participate in evolving musical activities. Building upon these principles, WAM Jam Party adapts CSCW concepts to immersive musical collaboration through dedicated mechanisms described in the following sections.

1) *Workspace awareness:* Workspace awareness refers to the up-to-the-moment understanding of collaborators’ activities within a shared workspace [6], [19]. In immersive environments, this awareness naturally extends beyond document editing to encompass body position, gaze direction, gestures, object manipulation, and spatial proximity. WAM Jam Party continuously synchronizes participants’ avatars, controller positions, gaze direction, object manipulations, and editing operations through

<sup>3</sup><https://www.youtube.com/playlist?list=PLNUy77gRLMkw>

the collaborative Conflict-free Replicated Data Type (CRDT) infrastructure presented in Section IV-B. Musical interactions therefore become immediately observable by every connected participant, facilitating implicit coordination without requiring continuous verbal communication.

Visual cues further reinforce awareness by displaying user selections, temporary object ownership during manipulation, and ongoing editing operations. Because modular audio graphs are spatially organized, participants progressively develop common spatial conventions regarding workstation organization, instrument placement, and collaborative responsibilities.

2) *Acoustic territories*: Territoriality has long been recognized in CSCW as a fundamental mechanism supporting ownership, coordination, and parallel work [27]. More recently, Men et al. demonstrated that fluid acoustic boundaries improve collaborative creativity in immersive musical environments compared with rigid spatial partitions [22]. Building upon these observations, *WAM Jam Party* introduces acoustic territories, where territorial organization emerges directly from the spatial propagation of sound.

Each modular musical installation generally terminates with one or several spatial audio output nodes, defining the propagation region of the corresponding musical content. Inside this region, participants perceive the complete immersive sound field. Beyond its limits, sound progressively attenuates until becoming inaudible, naturally creating an acoustic territory around the installation. The radius of this territory can be adjusted dynamically according to the collaborative context.

Acoustic territories regulate perception before interaction. By default, neighbouring musical installations remain audible, allowing participants to discover ongoing activities, move naturally between workspaces, and progressively join collaborative groups. However, musicians may temporarily transform an open territory into a protected acoustic territory, isolating the corresponding workspace whenever private creative activity becomes desirable.

Whereas territoriality in conventional CSCW systems is primarily expressed through visual organization, Musical Metaverses introduce auditory territoriality as an additional organizational mechanism. Sound propagation itself becomes a means of coordinating collaborative activity, allowing creative spaces to evolve dynamically as musical practice unfolds. Figure 4 illustrates two neighbouring musical installations. The central drumkit is also associated with a private acoustic territory (only the user playing hears it) that lets the user record a drum pattern before propagating it to other users, allowing multiple collaborative activities to coexist without explicit visual partitioning.

3) *Configurable collaborative roles*: Collaboration also requires regulating who may perform particular actions within shared creative spaces. Recent work distinguishes between musical / social roles (performer, audience, teacher) and interaction authority roles defining editing permissions within collaborative music systems [3], [23].

*WAM Jam Party* deliberately separates musical intention from interaction authority. A performer, teacher, technician,

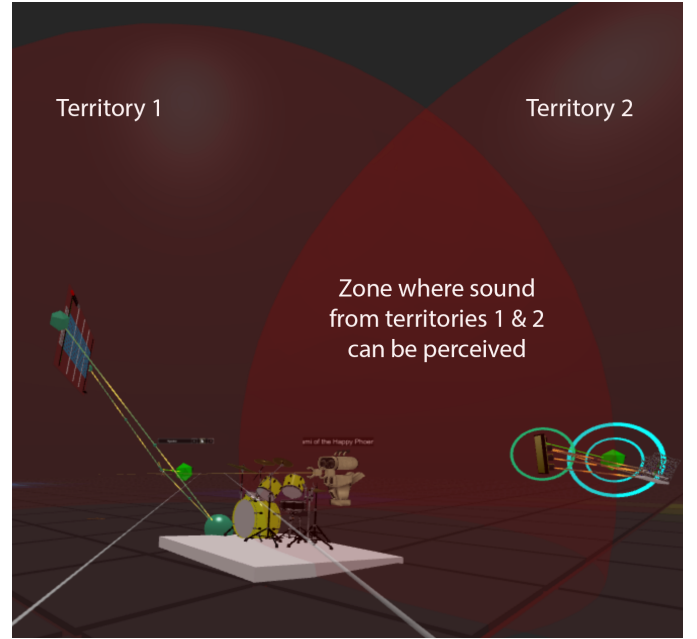


Figure 4. Two neighbouring acoustic territories in *WAM Jam Party*. The semi-transparent red spheres represent the acoustic territories associated with each musical installation; their radius defines the maximum propagation range of the corresponding spatial audio. Such territories allow multiple collaborative workspaces to coexist while preserving natural sound-based coordination.

or student may therefore receive different editing privileges depending on the collaborative situation without changing their musical role.

Following our previous work [7], [23], the platform currently supports three permission levels:

- **Audience**, allowing navigation and observation without modifying the environment.
- **Contributor / Technician**, allowing creation, manipulation, editing, and deletion of musical objects;
- **Owner** additionally managing projects, invitations, sessions, and other players' permissions.

These permissions are independently associated with categories of collaborative actions, including object creation, deletion, parameter editing, graph rewiring and project administration. Such fine-grained control enables collaboration models ranging from completely open public jam sessions to tightly supervised educational workshops or curated artistic performances. Table I summarizes the context-sensitive permission model. The three core permission levels (Audience, Contributor / Technician and Owner) are fully implemented, while the context-specific role specializations (Moderator, Performer, Student) shown in Table I are currently being integrated.

### C. Creative Space Organization

While the Interaction dimension enables musical expression and the Collaboration dimension supports coordinated multi-user creativity, sustainable musical activities also require mechanisms for organizing participants, projects, workspaces, and collaborative sessions over extended periods. This pattern

| Context | Role        | Add nodes | Edit params | Rewire/delete | Grant/revoke | Kick/ban |
|---------|-------------|-----------|-------------|---------------|--------------|----------|
| Common  | Owner       | ■         | ■           | ■             | ■            | ■        |
| Common  | Moderator   | □         | □           | ■             | ■            | ■        |
| Common  | Audience    | □         | □           | □             | □            | □        |
| Perf.   | Performer   | ■         | ■           | □             | □            | □        |
| Perf.   | Technician  | ■         | ■           | □             | □            | □        |
| Jam     | Contributor | ■         | ■           | ■             | □            | □        |
| Lesson  | Student     | ■         | □           | □             | □            | □        |

Table I

PERMISSION MATRIX. ■ = RIGHT GRANTED; □ = RIGHT DENIED. ADAPTED FROM [23].

echoes musical practice more broadly, where creative activity evolves through successive sessions rather than isolated encounters. Recent participatory co-design studies similarly emphasize that spatial layout and instrument design are inseparable in collaborative VR music-making [4].

The Creative Space Organization dimension addresses these organizational aspects by structuring collaborative activity around persistent identities, projects, sessions, and sharing mechanisms. Rather than considering collaborative sessions as temporary communication channels, *WAM Jam Party* treats them as organized creative spaces supporting long-term artistic, educational, and research activities. Figure 5 summarizes the conceptual organization of this dimension. Persistent identities provide continuity across projects, which in turn contain reusable collaborative sessions. Sharing mechanisms and configurable permissions regulate access to these creative spaces, allowing communities to evolve progressively while preserving a coherent organizational structure.

1) *Persistent identities*: Creative communities begin with persistent participants. While anonymous users may immediately join temporary public sessions without registration, authenticated users obtain persistent identities allowing them to create projects, manage collaborative spaces, configure permissions, and receive invitations. This dual access model balances accessibility with continuity: casual visitors discover the platform without account creation, while authenticated users progressively establish persistent identities preserving ownership, responsibilities, and collaborative relationships across multiple projects.

2) *Projects and collaborative sessions*: Creative work is organized around projects and collaborative sessions serving complementary purposes. A project represents a long-term creative objective such as preparing a concert, conducting a classroom activity, or developing a research prototype... Each project contains one or more collaborative sessions, corresponding to composition, rehearsals, performances, pedagogical activities, technical preparation, or alternative experimental branches.

Sessions constitute immersive collaborative workspaces preserving not only modular audio graphs but also their collaborative organization, including participant roles, permissions, acoustic territories, and workspace layouts. As illustrated in Figure 5, they become reusable organizational units that can be duplicated, adapted to new contexts, shared with collaborators,

locked to prevent mutations, or published as reusable templates preserving complete collaborative environments rather than isolated musical artifacts.

This organization naturally supports iterative creative life-cycles extending from experimentation and rehearsal to performance, teaching, and subsequent reuse across projects, allowing creative knowledge to progressively accumulate within persistent musical communities.

3) *Sharing and collaborative communities*: Long-term collaboration requires mechanisms allowing creative spaces to be shared, discovered, and progressively opened to new participants. Projects and sessions may remain private, be shared through invitation links, or be explicitly opened to selected registered users. Combined with the configurable collaborative roles introduced in Section III-B, these mechanisms support a broad spectrum of collaborative situations ranging from small artistic collectives to educational workshops, research collaborations, and public performances.

This organization also facilitates the gradual evolution of creative communities. A musical project may initially involve only a few collaborators during composition, progressively expand to additional performers during rehearsals, and eventually welcome audiences or external contributors during public performances.

#### D. Persistence as a transversal design principle

Persistence extends the three functional dimensions across time by preserving both musical artifacts and their collaborative context. Unlike the previous dimensions, it does not introduce new interaction capabilities but ensures the long-term continuity of creative activities.

Supporting persistence in a collaborative Musical Metaverse is considerably more challenging than saving conventional project files. The platform must continuously maintain a coherent distributed representation of a shared creative environment while multiple participants simultaneously manipulate modular audio graphs, move objects in three-dimensional space, modify permissions, create or reorganize collaborative sessions, and interact through low-latency network synchronization. In this context, persistence becomes a distributed collaborative persistence problem, where consistency must be maintained not only across time but also across distributed clients.

Traditional collaborative music systems generally preserve isolated artifacts such as audio files, patches, presets, project documents, or session settings. Persistent Musical Metaverses must instead preserve the complete collaborative environment in which musical activities take place, including modular audio graphs, parameter values, spatial layouts, participant roles, permissions, acoustic territories, and the organizational relationships between projects and collaborative sessions.

In *WAM Jam Party*, persistence therefore operates across the entire framework rather than as an isolated subsystem. Every modification of the shared environment is continuously synchronized and transparently persisted without requiring any explicit save operation. This distributed collaborative persistence provides a seamless user experience: after reloading

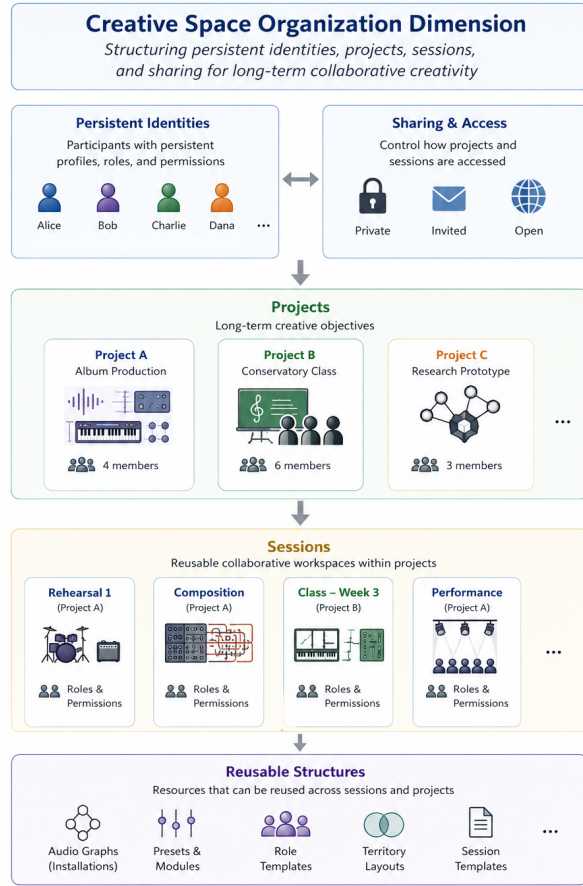


Figure 5. Conceptual organization of the Creative Space Organization dimension.

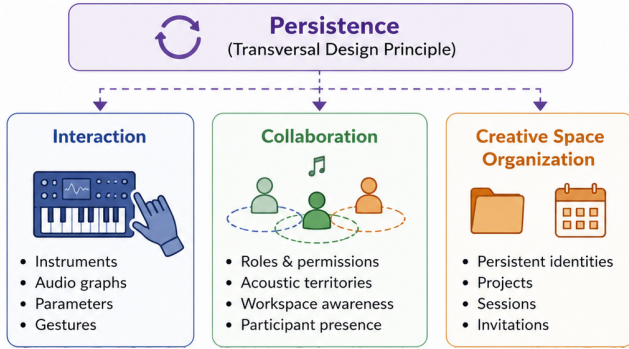


Figure 6. Persistence as a transversal design principle.

the application, recovering from an unexpected crash, reconnecting following a network interruption, or simply returning to a project days or weeks later, users recover either the state in which they left the environment, or, if collaborators continued working during their absence, the latest consistent shared state. Musical activities therefore resume immediately without reconstructing either the musical environment or its collaborative organization. By preserving the complementary aspects of the Interaction, Collaboration, and Creative Space

Organization dimensions together, persistence supports iterative creative processes spanning composition, rehearsal, teaching, experimentation, and performance across successive sessions (Figure 6).

## IV. IMPLEMENTATION

### A. Immersive WebXR client

The client application is implemented entirely using open Web technologies. *Babylon.js* provides three-dimensional rendering and WebXR support, while audio processing relies on the Web Audio API together with WAMs. Every musical object is instantiated dynamically as a reusable WAM directly inside the virtual environment.

*WAM Jam Party* adopts a spatial modular paradigm in which musical components are represented as manipulable three-dimensional entities interconnected through virtual cables to construct collaborative audio graphs (Figure 2).

### B. Collaborative synchronization

Supporting collaborative music-making requires maintaining a coherent shared representation of both musical objects and participant activities. *WAM Jam Party* leverages CRDT with *Yjs* and peer-to-peer communication through WebRTC to synchronize collaborative interactions while preserving eventual consistency.

Beyond synchronizing musical graphs, the platform continuously propagates participant positions, avatars, controller movements, gaze direction, and editing operations. This shared awareness enables musicians to coordinate their actions naturally while interacting within the same virtual workspace.

### C. Creative space services

Long-term musical activities require organizational mechanisms extending beyond individual collaborative sessions. *WAM Jam Party* therefore introduces dedicated services managing persistent identities, projects, collaborative sessions, permissions, and invitations through a REST API backed by a relational database. Projects organize long-term musical activities, while sessions represent reusable collaborative workspaces preserving their social organization and associated access rights.

Rather than considering projects and sessions as simple storage containers, the platform treats them as persistent organizational structures supporting artistic productions, pedagogical activities, rehearsals, performances, and research experiments.

### D. Persistence Services

Unlike the previous components, persistence is not implemented as an isolated subsystem but operates across the complete software architecture. Every service contributing to the collaborative environment — including interaction, collaboration, and creative space organization — participates in maintaining the persistent state of the platform.

Accordingly, whenever users modify an instrument, a collaborative session, or the organization of a creative space, the corresponding state is automatically propagated to all connected users through the *Yjs* synchronization mechanism

operating several times per second. An elected user is then in charge of communicating periodically (every 30 seconds) the state of the session to the database for storage, with a re-election mechanism in case of failure. This persistent representation includes modular WAM graphs, parameter values, spatial layouts, collaborative metadata, permissions, and project organization. Upon reconnection, the platform transparently restores this state, allowing collaborative activities to resume without requiring manual reconstruction of either the musical environment or its organizational context.

## V. FORMATIVE QUALITATIVE EVALUATION

Hereafter, we present the results of a formative study aimed at providing a qualitative assessment of *WAM Jam Party*. The evaluation focused on the Interaction dimension and the general usability of the platform. Due to time and scheduling constraints, collaborative and persistence aspects were not formally evaluated within this study and remain supported by the system description and architectural design presented in the preceding sections. The goal of this evaluation was to identify usability issues, assess the perceived relevance of the proposed interaction concepts, and collect requirements for future developments.

### A. Evaluation protocol

We have designed a user study based on semi-structured interviews that includes a period of direct exposure and trial of *WAM Jam Party*. The evaluation steps included: i) welcome and introduction to the study, ii) demographic survey, iii) demonstration of the creative space organization mechanisms (account, projects, sessions, etc.), iv) onboarding through tutorial completion, v) trial of three musical instruments implementing diverse interaction modalities: 3D Piano, XYZ Cubes (similar to a 3D Theremin) with parameter modulation [14], and Virtual Drumkit [12], vi) debriefing. The steps iv and v were performed while participants were experiencing *WAM Jam Party* through a Meta Quest 3 Head-Mounted Display (HMD). Readers are referred to the YouTube playlist cited in Section III-A3 for video demonstrations of the immersive instruments, interaction techniques, and other WAM components evaluated in this study.

The semi-structured interviews were organized around four thematic blocks: (1) professional profile and prior experience with music, VR, and collaborative music-making; (2) onboarding experience and tutorial usability; (3) feedback on each tested instrument (3D Piano, XYZ Cubes, Virtual Drumkit); and (4) overall impressions regarding the platform's creative and pedagogical potential. Responses were analyzed by identifying recurring themes across participants and distinguishing widely shared observations from individual suggestions.

### B. Participants

We recruited a convenience sample of seven participants. All participants were male; their ages ranged from 26 to 65 years ( $M = 46.2$ ,  $SD = 12.7$ ). Two participants were members of the electroacoustic orchestra of the Nice Conservatory of Music,

two were instructors at the same conservatory, and three were Computer Science teachers at Université Côte d'Azur. Only one of the seven participants reported having no prior musical background. Regarding prior VR experience, two participants reported having no prior experience, one reported having "very little" experience, three reported having had "several" prior experiences, and one did not report his level of experience.

We acknowledge that the participant sample is limited in both size and demographic diversity. All seven participants were male, which restricts the range of perspectives represented in the findings. Future evaluations should recruit larger and more gender-balanced participant groups to improve the generalizability of the results.

The following results are organized around three themes reflecting the main categories of feedback that emerged across participants: general VR interaction difficulties, instrument-specific observations, and system-level usability.

### C. Results

Each session with participants lasted about 90 minutes, with approximately 1 hour dedicated to trials of the tool. We could identify some usability problems and gather feedback and suggestions to enhance the tool. We have identified 13 recurring usability issues. Among them, 3 are VR-related, 3 are feedback on instruments, and 7 concern the whole system.

1) *VR-related issues*: Several participants experienced difficulties with VR locomotion, object positioning, and controller-based navigation, particularly those like P4 who were unfamiliar not only with Virtual Reality but also with the gamepad interaction techniques commonly used in 3D video games. They suggested adding teleportation and automatic repositioning mechanisms to facilitate interaction with musical modules.

Currently, users can grab and rotate any module, but an automatic behavior to facilitate this would be appreciated by all participants. For instance, P5 stated "*I want it [the module] to face me...*". Participants with low VR literacy confused the controller's trigger and A-button, or the trigger and the grab button (which was judged too sensitive by P1). Finally, breaks were required in the middle of test sessions as participants (notably P3) suffered visual fatigue due to the experience. Fatigue and maneuverability issues further reinforce the importance of persistence for accessibility, reducing the quantity of interactions. These problems are clearly related to interaction with VR rather than specific to the Musical Metaverse.

2) *Instrument-related feedback*: Considering the instruments, the most appreciated (6 out of 7 participants) were the XYZ Cubes. Participants highlighted their relevance for mobilizing the full VR potential. P4 (music instructor) notably mentioned they "*tried to create such instruments with Leap Motion or ROLI interfaces*" outside of VR but were lacking visual feedback on modulation.

The Virtual Drumkit received mixed feedback: P2 was very enthusiastic about it but P4 (percussionist) was doubtful about the use of percussion instruments without tangible feedback and stick rebounds. Most users appreciated the fact that they could

approach the drumkit and sit on its throne to get automatically positioned and locked in place in front of the drumkit. Some participants would have liked to have a similar automatic positioning in front of all interactive modules. On the downside, participants found the drumkit too big and most of them had trouble resizing it as desired with the proposed two-handed grab interaction.

The 3D Piano was finally deemed interesting but considered unfinished by some users. Many had a hard time positioning themselves properly to perform with it, and generally preferred the XYZ Cubes to modulate parameters.

Overall, all participants were able to use the instruments shown to them on their own and play or experiment with them, even though certain users mentioned that the evaluator's help was welcome during the process, otherwise they might not have understood everything they could do in the environment or each interaction modality.

3) *System results*: The project management system seemed relevant to the participants, as illustrated by P1 who compared *WAM Jam Party*'s system to Ableton's. P1, P2, and P4 asserted that the management interface would suit their needs. P3 noted it was important to him for the project management interface to be accessible from desktop. Overall, users seemed moderately to highly interested in using the application for musical purposes. P2, who enjoyed the "virtual hardware" aspects, rated his interest in it 5/5 while P4, who was the most critical overall, still rated it 3/5 and also envisioned museum-like applications to exhibit and try vintage systems. In contrast, P5 deemed the presented instruments not appropriate for the musical workflow "he was used to", which is "*too acoustic*", but mentioned that with a bit of practice he might reconsider his position. One of the main benefits identified by P5 was the possibility of instrumental practice without owning real instruments, while he also suggested applications of the Musical Metaverse for sound engineers to create digital twins of real setups.

The feedback received allows us to correct certain aspects before moving to the collaborative tests and envision future improvements in general.

## VI. DISCUSSION

The proposed framework aims not only to describe the design of *WAM Jam Party* but also to identify broader principles for designing Persistent Musical Metaverses. Beyond its implementation, three observations emerge from this work.

### A. A conceptual framework rather than a software architecture

Although instantiated through *WAM Jam Party*, the same design dimensions could guide other Musical XR platforms, including non-WebXR-based systems built upon Unity, Unreal Engine, or native XR development kits. More generally, the framework provides a common vocabulary for describing, comparing, and designing collaborative Musical XR systems beyond their implementation choices.

### B. Designing expressive yet accessible Musical Metaverses

Developing Persistent Musical Metaverses requires balancing two complementary objectives that are often in tension: expressive flexibility and ease of use. Supporting modular audio graphs, reusable creative spaces, configurable collaboration mechanisms, and persistent organizational structures enables a broad range of artistic, educational, and collaborative workflows, but also increases the complexity of the user experience.

Rather than simplifying the underlying conceptual model, future work should investigate interaction techniques that progressively expose this richness through improved onboarding, contextual guidance, adaptive interfaces, and progressive disclosure of advanced functionality. More generally, this trade-off is unlikely to be specific to *WAM Jam Party* and will probably remain a central design challenge for future Musical Metaverses combining modular creativity with long-term collaborative organization.

### C. Open standards as a foundation for sustainable Musical Metaverses

Our experience also suggests that open Web standards provide a particularly suitable technological foundation for Persistent Musical Metaverses. Technologies such as WebXR, WebRTC, Web Audio, WebGL, and WebGPU offer interoperable building blocks evolving independently of specific XR devices while remaining broadly supported across modern browsers.

Combined with WAMs, this approach enables new musical instruments, effects, controllers, AI-assisted tools, and visualization components to be integrated incrementally without modifying the underlying collaborative infrastructure. Beyond portability and extensibility, open standards and open-source reference implementations facilitate reproducible research, comparative evaluation, and community-driven evolution. In this perspective, *WAM Jam Party* is intended not only as an implementation of the proposed framework, but also as an open experimental platform supporting future research on collaborative and persistent Musical XR systems.

## VII. CONCLUSION AND FUTURE WORK

This paper introduced a multi-dimensional framework for designing Persistent Musical Metaverses, structured around three complementary functional dimensions — Interaction, Collaboration, and Creative Space Organization — supported by Persistence as a transversal design principle.

The framework was instantiated through *WAM Jam Party*, an open-source Web-based Musical Metaverse built upon WebXR, WebRTC, Web Audio, and the Web Audio Modules. This implementation demonstrates how immersive interaction, collaborative organization, reusable creative spaces, and transparent persistence can be integrated within a unified environment supporting artistic creation, education, and experimentation.

The formative evaluation presented in this paper focused on individual interaction with the proposed immersive instruments and the general usability of the platform. Because a substantial part of each evaluation session was devoted

to familiarizing participants with the environment, time and scheduling constraints did not allow extended multi-user collaborative experiments within this study. While the formative evaluation highlights opportunities to further improve usability and onboarding, evaluating the collaborative aspects through dedicated multi-user sessions constitutes an important direction for future work, as well as formally evaluating the usability of the platform once corrections derived from the formative evaluation are implemented. Persistence and creative space organization aspects will be further explored through long-term studies in which participants can create, organize, and resume collaborative musical activities over extended periods of time.

Beyond the proposed framework itself, *WAM Jam Party* provides an extensible reference implementation that researchers, educators, artists, and developers can freely use, study, and extend. We hope that both the conceptual framework and its open-source implementation will contribute to the emergence of interoperable, sustainable, and reproducible Musical Metaverses, supporting future research on collaborative musical creation in immersive environments.

**Acknowledgments.** This work was supported by the French government, through the France 2030 investment plan managed by the ANR, under the “UCA DS4H” project (ref. ANR-17-EURE-0004).

## REFERENCES

- [1] Z. K. Aesthetic Interactive. Exa: The infinite instrument. [https://store.steampowered.com/app/606920/EXA\\_The\\_Infinite\\_Instrument/](https://store.steampowered.com/app/606920/EXA_The_Infinite_Instrument/), 2019.
- [2] D. Birnbaum, R. Fiebrink, J. Malloch, and M. M. Wanderley. 2005: Towards a dimension space for musical devices. In *A NIME Reader: Fifteen Years of New Interfaces for Musical Expression*, pages 211–222. Springer, 2017.
- [3] A. Boem, D. Dziwis, M. Tomasetti, S. Etezazi, and L. Turchet. “It Takes Two” — shared and collaborative virtual musical instruments in the musical metaverse. In *2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)*, pages 1–10. IEEE, 2024.
- [4] A. Boem, S. Skouras, G. B. Hinkis, M. Mousset, and L. Turchet. Co-designing virtual reality musical instruments and spatial layouts for collaborative music-making. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 820–831. Zenodo, June 2026.
- [5] A. Boem, M. Tomasetti, A. Gabriele, A. di Scipio, A. Micalizzi, and L. Turchet. Imagining the musical metaverse: A study on the needs and creative practices of composers and musicians. *Digital Creativity*, pages 1–27, 2025.
- [6] N. Bryan-Kinns and F. Hamilton. Identifying mutual engagement. *Behaviour & Information Technology*, 31(2):101–125, 2012.
- [7] M. Buffa and T. Burns. Real-time collaborative music creation on the web: exploiting web audio modules for interactive performance and composition. In *2025 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–8, 2025.
- [8] M. Buffa, T. Burns, C. Ferrari, J. Kleimola, O. Larkin, J. Lebrun, S. Letz, Y. Orlarey, S. Ren, and S. Yi. Ten years of web audio modules: Audio plug-ins for the web. *Journal of the Audio Engineering Society*, 73(12):804–826, Dec. 2025.
- [9] M. Buffa, S. Ren, O. Campbell, T. Burns, S. Yi, J. Kleimola, and O. Larkin. Web audio modules 2.0: An open web audio plugin standard. In *Companion Proceedings of the Web Conference 2022 (WWW ’22)*. ACM, 2022.
- [10] M. Buffa, M. Winckler, Q. Escobar, S. Demont, A. Hofr, and A. Mir-Sadjadi. Interactive Audio Sculpting: Plugin Customization and UI Affordances in Immersive Environments. In *2025 IEEE 6th International Symposium on the Internet of Sounds (IS2)*, pages 1–10, L’Aquila, Italy, Oct. 2025. IEEE.
- [11] M. Buffa, M. Winckler, Q. Escobar, S. Demont, A. Hofr, and A. Mir-Sadjadi. Wam jam party: a collaborative real-time music creation experience in vr. In *WAC 2025-Web Audio Conference*. Zenodo, 2025.
- [12] M. Buffa, M. Winckler, and A. Mir-Sadjadi. Embodied Virtual Instruments in Web-Based Multi-User VR: A Case Study with a 3D Drum Kit and Web Audio Modules. *WAC 2025 - 9th Web Audio Conference*, 2025.
- [13] C. Çakmak, A. Çamcı, and A. G. Forbes. Networked virtual environments as collaborative music spaces. In *Proceedings of the 2016 Conference on New Interfaces for Musical Expression (NIME 2016)*, pages 106–111, 2016.
- [14] S. Demont, M. Buffa, M. Winckler, and F. Robert. Vers la formalisation geste-son pour la conception d’instruments musicaux immersifs. In *JIM 2026 – Journées d’Informatique Musicale*, Lille, France, 2026. hal-05663276.
- [15] R. DuPlessis. A virtual instrument for physics-based musical gesture: CHON. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2022)*, 2022.
- [16] R. Fencott and N. Bryan-Kinns. Hey man, you’re invading my personal space! privacy and awareness in collaborative music. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010)*, pages 198–204, 2010.
- [17] C. Frame. Concerts of the future: Designing an interactive musical experience in VR. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2024)*, Utrecht, The Netherlands, 2024.
- [18] M. Gioiosa, F. Avanzini, L. A. Ludovico, S. Brambilla, and L. Ripamonti. Sound blocks VR: An accessible virtual reality musical instrument. In *Computer-Human Interaction Research and Applications*, pages 316–330. Springer Nature Switzerland, 2025.
- [19] C. Gutwin and S. Greenberg. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)*, 11(3-4):411–446, 2002.
- [20] A. Hunt, M. M. Wanderley, and M. Paradis. The importance of parameter mapping in electronic instrument design. *Journal of New Music Research*, 32(4):429–440, 2003.
- [21] S. Jordà. *Digital Lutherie: Crafting Musical Computers for New Musics’ Performance and Improvisation*. PhD thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2005.
- [22] L. Men and D. Zhao. Designing privacy for collaborative music making in virtual reality. In *Audio Mostly 2021*, pages 93–100. ACM, 2021.
- [23] A. Mir-Sadjadi, M. Buffa, M. Winckler, and F. Robert. A review of user roles in collaborative vr music-making environments: Towards role design guidelines with WAM Jam Party as a case study. In *JIM 2026 – Journées d’Informatique Musicale*, Lille, France, 2026. hal-05663345.
- [24] PatchXR. Patchworld (version 77.1.6) [software]. <https://www.patchxr.com/>, 2022. [Online].
- [25] S. Ppali, A. Boem, A. Covaci, M. Constantinides, F. Liarokapis, and L. Turchet. Looking inside the VR music scene: Mapping platforms, events and people. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*, pages 1–22, Barcelona, Spain, 2026. ACM.
- [26] Really Interactive AB. Virtuoso (version 1.1.8) [software]. <https://fasttravelgames.com/games/virtuoso>, 2022. [Online].
- [27] S. D. Scott, M. Sheelagh, T. Carpendale, and K. M. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 294–303, 2004.
- [28] S. Serafin, M. Geronazzo, C. Erkut, N. C. Nilsson, and R. Nordahl. Sonic interactions in virtual reality: State of the art, current challenges, and future directions. *IEEE Computer Graphics and Applications*, 38(2):31–43, 2018.
- [29] L. Turchet. Musical metaverse: vision, opportunities, and challenges. *Personal and Ubiquitous Computing*, 27(5):1811–1827, Oct. 2023.
- [30] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet. Internet of musical things: Vision and challenges. *IEEE Access*, 6:61994–62017, 2018.
- [31] P. Uro, F. Berthaut, T. Pietrzak, and M. M. Wanderley. Decoupling physical and virtual spaces for new collaboration strategies in co-located mixed reality instruments. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME 2025)*, Canberra, Australia, 2025.
- [32] L. Zeller and A. Bauer. The walkable instrument: Modular patches as entangled environments in opensoundlab. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, pages 36–39, 2025.