The Musical Metaverse: Advancements and Applications in Networked Immersive Audio

1st Claudia Rinaldi
CNIT
Consorzio Nazionale Inter-Universitario per le Telecomunicazioni
L'Aquila, Italy
claudia.rinaldi@univaq.it

2nd Carlo Centofanti
DISIM
University of L'Aquila
L'Aquila, Italy
carlo.centofanti1@univaq.it

***Abstract*—The convergence of the Sound and Music Computing (SMC) domain with the Internet of Things (IoT) paradigm has given rise to the burgeoning field of the Internet of Sounds (IoS). Within this context, the concept of the Musical Metaverse emerges as a pivotal area of exploration, offering unprecedented opportunities for immersive audio experiences and networked musical interactions.**

This paper provides a comprehensive examination of the Musical Metaverse, delving into the technological innovations, methodologies, and applications that underpin this dynamic domain. We examine the integration of immersive audio technologies and Networked Musical Extended Reality (XR) to create collaborative virtual environments for music performance and education. We also explore advancements in smart musical instruments and ubiquitous music, focusing on their impact on accessibility, inclusiveness, and sustainability. Additionally, we address privacy, security, and large audio database management, highlighting the role of cloud services and wireless acoustic sensor networks in enhancing the efficiency of musical applications.

The aim is to provide a foundational understanding of the Musical Metaverse and its potential to influence the future of the global music industry.

***Index Terms*—Musical Metaverse, Internet of Sounds, Networked Immersive Audio, Networked Music Performance, Web Audio, Networked Musical XR, Smart Musical Instruments, Ubiquitous Music, Semantic Audio, Intelligent Music Production, Cloud-based Services, Ecoacoustics, Sonification.**

I. Introduction

The rapid evolution of digital technologies has profoundly impacted various aspects of human life, with the field of music being no exception. The convergence of Sound and Music Computing (SMC) with the Internet of Things (IoT) has paved the way for the emergence of the Internet of Sounds (IoS), a novel interdisciplinary field that harnesses networked technologies to enhance and transform musical experiences [1]. This paper focuses on the concept of the Musical Metaverse [2], a subset of IoS that encapsulates the integration of immersive audio technologies and networked musical interactions [3].

This work has been partially supported by the European Union through NextGenerationEU under the Italian Ministry of University and Research (MUR) National Innovation Ecosystem (ECS00000041 - VITALITY - CUP E13C22001060006, and partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART" - CUP E83C22004640001), and by the H2020-MSCA-RISE OPTIMIST project (GA: 872866).

The Musical Metaverse represents a virtual ecosystem where musicians, educators, and audiences can engage in real-time, interactive sound experiences across spatial and temporal boundaries. This paradigm shift not only augments traditional music production and performance but also introduces new possibilities for collaborative creation, music education, and participatory live performances. The ability to network musical instruments and environments via IoS technologies opens up innovative avenues for artistic expression and audience engagement.

The primary objective of this paper is to explore the key technological innovations, applications, and challenges associated with the Musical Metaverse. By examining networked immersive audio, networked musical extended reality (XR) [4], and semantic audio applications, we aim to provide a comprehensive overview of the current state and future potential of this field. Additionally, we will discuss the implications of these technologies on privacy, security, and ethical considerations, ensuring a holistic understanding of their impact.

This paper is structured as follows: Section 2 delves into the technological innovations driving the Musical Metaverse, including networked immersive audio and XR applications. Section 3 explores practical applications and use cases, such as participatory live music performances and music education. Section 4 addresses the management and utilization of semantic audio and large audio databases. Section 5 discusses privacy, security, and ethical considerations. Finally, Section 6 summarizes the findings and outlines future research opportunities in this burgeoning field.

By bridging theoretical insights with practical applications, this paper aims to contribute to the ongoing discourse at the 5th IEEE International Symposium on the Internet of Sounds (IEEE IS2 2024), fostering collaboration between academia and industry to further advance the development and adoption of IoS technologies [5].

II. Technological Innovations in the Musical Metaverse

The Musical Metaverse is underpinned by a range of cutting-edge technologies that enable immersive, interactive, and networked musical experiences. These innovations are crucial in transforming how music is created, performed, and experienced, offering new dimensions of connectivity and engagement. In this section, we will explore the key technological advancements that form the backbone of the Musical Metaverse, focusing on two primary areas: Networked Immersive Audio and Networked Musical Extended Reality (XR).

Networked Immersive Audio technologies facilitate the creation of spatially rich and interactive soundscapes that can be experienced in real-time by multiple users across different locations. This approach leverages advancements in audio processing, web technologies, and networking to deliver high-fidelity, immersive audio experiences that transcend the limitations of traditional acoustic environments. The potential applications of Networked Immersive Audio range from virtual concerts and collaborative music production to educational tools and therapeutic applications, making it a cornerstone of the Musical Metaverse [6].

Networked Musical XR represents the fusion of musical interaction with extended reality technologies, including virtual reality (VR), augmented reality (AR), and mixed reality (MR). This integration allows users to engage with music in immersive virtual environments, where they can interact with virtual instruments, collaborate with other musicians, and participate in live performances without geographical constraints. The use of XR in music extends beyond entertainment, offering significant implications for music education, training, and rehabilitation.

By delving into these two areas, we aim to provide a comprehensive understanding of the technological foundations that enable the Musical Metaverse. The following subsections will examine Networked Immersive Audio and Networked Musical XR in detail, highlighting their key components, applications, and the transformative potential they hold for the future of music.

A. Networked Immersive Audio

Networked Immersive Audio (NIA) is a transformative technology that facilitates the creation and distribution of high-fidelity, spatially rich audio experiences over digital networks. This technology leverages advancements in audio signal processing, web-based audio technologies, and networking protocols to deliver interactive and immersive soundscapes to users in real-time, regardless of their physical location. NIA is pivotal to the Musical Metaverse, as it enables new forms of musical interaction and engagement that were previously unattainable.

1) Definition and Core Components: Networked Immersive Audio involves the transmission and synchronization of audio signals across multiple devices and locations to create a cohesive and immersive auditory experience. The core components of NIA include:

- **Spatial Audio Processing:** Techniques such as binaural rendering, ambisonics, and wave field synthesis that simulate three-dimensional sound fields, providing listeners with a realistic sense of spatial presence and directionality [7].
- **Low-Latency Networking:** High-speed, low-latency communication networks that ensure synchronous audio playback and interaction, critical for real-time applications such as live performances and collaborative music creation [8].
- **Web Audio Technologies:** Web-based frameworks and standards (e.g., Web Audio API) that facilitate the development and deployment of interactive audio applications directly within web browsers, enhancing accessibility and reach [9].

2) Applications and Use Cases: The applications of Networked Immersive Audio are vast and varied, encompassing both creative and practical domains. Some key use cases include:

- **Virtual Concerts:** Artists can perform live concerts in virtual environments [10], where audiences from around the world can experience high-fidelity, immersive audio as if they were physically present at the venue. These virtual concerts can include interactive elements, such as audience participation and real-time feedback.
- **Collaborative Music Production:** Musicians can collaborate remotely in real-time [11], using NIA to create and manipulate soundscapes together. This facilitates creative collaboration across geographical boundaries, enabling new forms of musical expression and production workflows.
- **Educational Tools:** NIA can be utilized in music education to create immersive learning environments [12]. Students can engage with interactive soundscapes, participate in virtual ensembles, and receive real-time feedback, enhancing their learning experience.
- **Therapeutic Applications:** Immersive audio environments can be used for therapeutic purposes, such as sound therapy and rehabilitation. The ability to create controlled, multi-sensory experiences has potential benefits for mental health and well-being [13].

3) Challenges and Future Directions:: Despite its potential, the implementation of Networked Immersive Audio faces several challenges. Ensuring consistent low-latency performance across diverse network conditions is critical for maintaining the quality of real-time interactions. Additionally, the complexity of spatial audio processing requires significant computational resources, which can be a barrier for widespread adoption.

Future research and development in NIA aim to address these challenges by optimizing network protocols, improving computational efficiency, and enhancing the scalability of immersive audio systems. As technology advances, the integration of NIA with other emerging fields, such as artificial intelligence and machine learning, is expected to further expand its capabilities and applications.

In conclusion, Networked Immersive Audio is a foundational technology for the Musical Metaverse, enabling innovative and immersive musical experiences that redefine the boundaries of traditional audio production and consumption.

B. Networked Musical XR

Networked Musical Extended Reality (XR) represents a cutting-edge intersection of music and extended reality technologies, including virtual reality (VR), augmented reality (AR), and mixed reality (MR). This convergence enables musicians, educators, and audiences to interact with music in immersive virtual environments, breaking free from the limitations of physical spaces. Networked Musical XR is instrumental in the evolution of the Musical Metaverse [14], offering novel ways to create, perform, and experience music.

1) Definition and Core Components: Networked Musical XR involves the use of XR technologies to facilitate real-time, interactive musical experiences across distributed environments. The core components of Networked Musical XR include:

- **Virtual Reality (VR):** Fully immersive environments where users can interact with digital musical instruments, virtual performance spaces, and other musicians as if they were in the same physical location.
- **Augmented Reality (AR):** Overlaying digital musical elements onto the physical world, allowing users to interact with both virtual and real-world instruments and environments. This can enhance live performances and music education by integrating digital augmentation into the physical space.
- **Mixed Reality (MR):** Seamlessly blending the physical and digital worlds, enabling interactions with both physical and virtual musical elements in real-time. MR offers a hybrid experience that combines the benefits of both VR and AR.

2) Applications and Use Cases: The integration of XR in musical contexts opens up a myriad of applications and use cases, enhancing creativity, collaboration, and engagement. Key use cases include:

- **Virtual Music Performances:** Musicians can perform live in virtual venues, where audiences can attend from anywhere in the world. These performances can incorporate interactive elements, such as audience participation and dynamic environments that respond to the music.
- **Collaborative Composition and Rehearsal:** Musicians can collaborate in virtual spaces to compose and rehearse music together. XR technologies enable real-time interaction and visualization of musical ideas, facilitating creative collaboration across distances.
- **Music Education and Training:** XR can revolutionize music education by providing immersive learning environments where students can interact with virtual instruments, participate in ensemble practices, and receive real-time feedback from instructors [15]. AR can enhance traditional music lessons by overlaying instructional content onto physical instruments.
- **Interactive Music Experiences:** XR allows for the creation of interactive music installations and experiences, where users can explore and manipulate soundscapes in immersive environments. This can be used for artistic installations, therapeutic applications, and entertainment.

3) Challenges and Future Directions: Implementing Networked Musical XR presents several challenges. Technical limitations such as latency, bandwidth, and computational power are critical factors that affect the quality of real-time interactions. Ensuring seamless integration of physical and virtual elements requires sophisticated tracking and synchronization technologies. Additionally, the accessibility of XR hardware and the development of user-friendly interfaces are essential for widespread adoption.

Future research in Networked Musical XR aims to overcome these challenges by advancing XR hardware capabilities, optimizing network protocols for low-latency performance [16], and developing intuitive user interfaces. The integration of artificial intelligence and machine learning can further enhance XR experiences by providing intelligent responses and adaptive interactions within virtual environments.

As technology continues to evolve, the potential of Networked Musical XR will expand, offering new possibilities for musical creativity and interaction. By merging the physical and virtual realms, XR technologies have the potential to transform the way we create, perform, and experience music, making the Musical Metaverse a reality.

In conclusion, Networked Musical XR is a transformative component of the Musical Metaverse, providing immersive and interactive musical experiences that push the boundaries of traditional music practices. Its applications span from virtual performances and collaborative creation to innovative educational tools, paving the way for a new era of musical engagement.

III. Applications and Use Cases

The integration of networked immersive audio and extended reality technologies in the Musical Metaverse has paved the way for innovative applications and use cases that transform traditional musical experiences. This section explores two primary areas where these advancements have significant impact: participatory live music performances and music education within the Internet of Sounds (IoS) context.

A. Participatory Live Music Performances

One of the most exciting applications of the Musical Metaverse is the facilitation of participatory live music performances. Traditionally, live performances are confined to physical venues, limiting audience size and engagement. However, networked immersive audio and XR technologies allow for the creation of virtual concert experiences that transcend these physical boundaries.

In a participatory live music performance, audiences are not merely passive listeners but active participants in the musical experience. Using virtual reality (VR) or augmented reality (AR), attendees can immerse themselves in a virtual concert hall or an open-air festival, interacting with the environment and the performers in real-time. These technologies enable dynamic, responsive environments where visual and auditory elements can be manipulated based on audience interaction.

For example, a virtual concert might feature interactive elements where the audience can influence the visual effects or the setlist through real-time feedback [17]. Musicians can perform from different locations around the world, yet appear to be on the same virtual stage, offering a seamless and unified performance experience. This not only enhances the engagement of the audience but also provides new opportunities for artists to experiment with their performances and connect with fans in innovative ways.

Furthermore, these virtual platforms can host larger audiences than traditional venues, reaching a global audience without the constraints of geographical location. This democratizes access to live music, allowing people from various backgrounds and regions to participate in and enjoy high-quality musical performances.

B. Music Education in the IoS Context

The Internet of Sounds (IoS) significantly impacts music education by providing immersive, interactive, and accessible learning environments [18]. Traditional music education often relies on in-person instruction and physical instruments, which can be limiting in terms of accessibility and reach. IoS technologies, combined with networked immersive audio and XR, offer novel solutions to these challenges.

In an IoS-enabled music education setting, students can engage with virtual instruments and environments that mimic real-world scenarios. Virtual reality (VR) classrooms allow students to practice and perform in simulated concert halls or orchestras, providing a realistic and immersive learning experience. Augmented reality (AR) can be used to overlay instructional content onto physical instruments, offering real-time guidance and feedback as students practice.

These technologies also facilitate remote learning and collaboration. Students and instructors can connect from different locations, participating in live, interactive lessons and ensemble practices [19]. This is particularly beneficial for students in remote areas or those who do not have access to high-quality music education resources locally. The ability to record and review performances in these virtual environments provides additional learning opportunities, enabling students to self-assess and improve their skills.

Moreover, IoS technologies support inclusive music education by catering to diverse learning needs. Adaptive interfaces and customizable learning environments can be designed to accommodate students with different abilities and learning styles, ensuring that music education is accessible to all.

The integration of intelligent music production tools within these environments further enhances the learning experience. Students can experiment with composition and production, using AI-driven tools to receive real-time suggestions and feedback. This not only aids in the development of technical skills but also fosters creativity and innovation.

In conclusion, the applications and use cases of the Musical Metaverse in participatory live music performances and music education highlight the transformative potential of IoS technologies [20]. By breaking down traditional barriers and creating immersive, interactive environments, these advancements are reshaping the way music is performed, experienced, and learned, paving the way for a more inclusive and engaging musical future.

IV. Semantic Audio and Data Management

Semantic audio and data management [21] are crucial components of the Musical Metaverse, enabling the intelligent processing, organization, and utilization of audio data. These technologies facilitate a range of applications, from enhancing music retrieval systems to powering advanced audio analysis and synthesis. This section delves into the key aspects of semantic audio and data management, focusing on semantic audio applications, audio datasets for deep learning and data mining, the challenges in managing large audio databases, and the protocols and exchange formats essential for efficient data handling.

A. Semantic Audio Applications

Semantic audio refers to the analysis and interpretation of audio content to extract meaningful information, such as genre, mood, instrumentation, or even specific musical events. These applications are integral to the functionality of the Musical Metaverse, providing the foundation for intelligent music services and enhanced user experiences.

Key semantic audio applications include:

- **Music Information Retrieval (MIR):** Techniques that enable users to search and retrieve music based on various attributes, such as melody, rhythm, or timbre. This improves the discoverability of music and enhances user interaction with digital music libraries.
- **Automatic Music Transcription:** Systems that convert audio recordings into symbolic representations [22], such as MIDI files or musical scores. This is useful for educational purposes, music analysis, and automated music production.
- **Emotion and Mood Detection:** Algorithms that analyze the emotional content of music, facilitating applications in music therapy, adaptive playlists, and interactive installations.
- **Audio Content Recommendation:** Personalized recommendation systems that suggest music based on user preferences and listening habits, leveraging semantic analysis to improve the accuracy and relevance of recommendations.

B. Audio Datasets for Deep Learning and Data Mining

The development of sophisticated semantic audio applications relies heavily on large, annotated audio datasets, which are essential for training deep learning models and conducting data mining to uncover patterns and insights within audio content.

Building comprehensive audio datasets involves collecting diverse audio samples and meticulously annotating them with relevant metadata, such as genre, mood, or specific audio events. Crowdsourcing and automated annotation tools can aid significantly in this process. Deep learning techniques, such as convolutional neural networks (CNNs) [23] and recurrent neural networks (RNNs), require extensive datasets to achieve high performance. These models can learn to identify complex audio patterns and perform tasks like genre classification, instrument recognition, and audio synthesis. Analyzing large audio datasets can reveal valuable insights, such as trends in music consumption, common structural elements in different music genres, and correlations between audio features and listener preferences. These insights can inform music production, marketing strategies, and user interface design, driving innovation and enhancing user experiences in the Musical Metaverse.

C. Challenges in Managing Large Audio Databases

Managing large audio databases presents several challenges that need to be addressed to ensure efficient storage, retrieval, and processing of audio data [24].

- **Scalability:** As audio datasets grow in size, scalability becomes a critical issue. Efficient storage solutions, such as cloud-based services, and robust indexing techniques are necessary to handle the increasing volume of data.
- **Data Quality and Consistency:** Ensuring the quality and consistency of audio data and its annotations is vital for reliable analysis and model training. This involves rigorous data cleaning, validation, and standardization processes.
- **Privacy and Security:** Protecting the privacy of users and the security of audio data is paramount. Implementing encryption, access control, and anonymization techniques can help safeguard sensitive information.
- **Real-Time Processing:** Many applications, such as live music performances and interactive installations, require real-time processing of audio data [25]. Achieving low-latency performance while maintaining high accuracy and reliability is a significant technical challenge.

D. Protocols and Exchange Formats

Effective management and interoperability of audio data in the Musical Metaverse depend on standardized protocols and exchange formats. Protocols like Open Sound Control (OSC) and MIDI (Musical Instrument Digital Interface) facilitate communication between different audio devices and software, enabling seamless integration and interaction. Common audio file formats, such as WAV, MP3, and FLAC, are essential for storing and exchanging audio data. Additionally, formats like Audio Definition Model (ADM) support the complex metadata required for immersive audio experiences [26]. Adopting metadata standards, such as ID3 tags for MP3 files and Extensible Metadata Platform (XMP) for various media types, ensures that audio files are accompanied by comprehensive and consistent metadata, enhancing their usability and discoverability.

V. Privacy, Security, and Ethical Considerations

As the Musical Metaverse continues to expand, addressing privacy, security, and ethical considerations becomes increasingly critical. These aspects ensure that the technologies not only function effectively but also align with broader societal values and regulations. This section explores the key issues related to privacy and security in acoustic sensor networks, ethical and sustainability aspects, and accessibility and inclusiveness.

A. Privacy and Security in Acoustic Sensor Networks

Acoustic sensor networks play a vital role in the Musical Metaverse by enabling the capture, processing, and transmission of audio data across various devices and platforms. However, these networks pose significant privacy and security challenges. The data captured by acoustic sensors can include sensitive information, such as conversations and environmental sounds, which must be protected to prevent unauthorized access and misuse.

To address these concerns, robust encryption methods are essential for securing data transmission within acoustic sensor networks. Implementing advanced encryption standards (AES) and secure communication protocols can help ensure that data remains confidential and tamper-proof during transit [27]. Additionally, access control mechanisms, such as multi-factor authentication and role-based access control, can limit data access to authorized personnel only.

Data anonymization techniques are also crucial in protecting user privacy. By removing or obfuscating personally identifiable information (PII) from audio data, it is possible to mitigate the risk of privacy breaches. Regular security audits and vulnerability assessments can help identify and address potential security weaknesses in acoustic sensor networks, ensuring ongoing protection against emerging threats.

B. Ethical and Sustainability Aspects

The deployment of IoS technologies in the Musical Metaverse raises several ethical considerations that must be addressed to ensure responsible use. One primary concern is the potential for bias in semantic audio applications and AI-driven music production tools [28]. Biases in training datasets can lead to unfair or discriminatory outcomes, impacting the inclusivity and fairness of these technologies.

To mitigate these issues, it is important to ensure that training datasets are diverse and representative of various musical genres, cultures, and user demographics. Transparency in the development and deployment of AI models is also essential, enabling users to understand how decisions are made and ensuring accountability for any biases or errors.

Sustainability is another critical aspect, as the energy consumption and environmental impact of large-scale IoS deployments can be significant. Adopting energy-efficient technologies and practices, such as measuring the energy consumption [29] to develop optimizing algorithms for lower power consumption and using renewable energy sources for data centers, can help reduce the environmental footprint of the Musical Metaverse.

C. Accessibility and Inclusiveness

Ensuring accessibility and inclusiveness in the Musical Metaverse is vital for providing equal opportunities for all users, regardless of their abilities or backgrounds. This involves designing IoS technologies and applications that cater to diverse user needs and preferences.

For instance, creating user interfaces with customizable features, such as adjustable text sizes, color contrasts, and audio feedback, can make applications more accessible to users with visual or auditory impairments. Additionally, providing multilingual support and culturally relevant content can enhance the inclusiveness of IoS applications, allowing users from different linguistic and cultural backgrounds to engage with the Musical Metaverse.

Collaborating with user communities and accessibility experts during the design and development process can help identify and address potential barriers, ensuring that IoS technologies are usable by as many people as possible. Implementing accessibility standards, such as the Web Content Accessibility Guidelines (WCAG) [30], can further enhance the inclusivity of IoS applications.

In summary, addressing privacy, security, and ethical considerations is essential for the sustainable and responsible development of the Musical Metaverse. By prioritizing these aspects, we can create a more secure, fair, and inclusive environment for all users, fostering trust and promoting the widespread adoption of IoS technologies.

VI. Conclusion and Future Directions

As the Musical Metaverse evolves, it continues to redefine the landscape of music creation, performance, and consumption through innovative technologies and applications. This section provides a summary of the key findings from our exploration of the Musical Metaverse and highlights emerging trends and future research opportunities that will shape its development.

A. Summary of Findings

The integration of Sound and Music Computing (SMC) with the Internet of Things (IoT) has given rise to the Internet of Sounds (IoS), a dynamic field that leverages networked technologies [31] to enhance musical experiences. Networked Immersive Audio and Networked Musical Extended Reality (XR) are at the forefront of this transformation, enabling real-time, interactive, and immersive soundscapes that transcend physical boundaries.

Participatory live music performances and IoS-enabled music education are two significant applications of the Musical Metaverse. These technologies democratize access to music, allowing audiences and students worldwide to engage with high-quality musical experiences. The development of semantic audio applications and the management of large audio datasets further enhance the capabilities of IoS technologies, enabling intelligent music services and advanced audio analysis.

Addressing privacy, security, and ethical considerations is crucial for the responsible deployment of IoS technologies. Ensuring robust data protection, mitigating biases, promoting sustainability, and enhancing accessibility and inclusiveness are essential for fostering trust and widespread adoption of the Musical Metaverse.

B. Emerging Trends and Future Research Opportunities

Several emerging trends are poised to drive the future of the Musical Metaverse. The continued advancement of artificial intelligence (AI) and machine learning (ML) will play a critical role in enhancing semantic audio applications, enabling more sophisticated and personalized music experiences. AI-driven tools can assist in music composition, production, and performance, opening new creative possibilities for artists and producers.

The convergence of IoS with other emerging access network technologies, such as 5G [32], optical network and edge computing [33], will further enhance the capabilities of networked immersive audio and XR applications. These technologies will enable ultra-low-latency communication and real-time processing, making interactive and immersive music experiences more seamless and accessible.

Future research opportunities lie in optimizing the scalability and efficiency of IoS technologies. Developing more energy-efficient algorithms and exploring renewable energy sources for data centers can help reduce the environmental impact of large-scale IoS deployments. Additionally, ongoing efforts to enhance data privacy and security will be essential in maintaining user trust and protecting sensitive information.

Exploring new use cases and applications of the Musical Metaverse beyond entertainment and education is another promising area of research. For instance, therapeutic applications of immersive audio environments can offer significant benefits for mental health and rehabilitation. Similarly, the use of IoS technologies in social and cultural contexts can promote cross-cultural exchange and collaboration.

In conclusion, the Musical Metaverse represents a transformative shift in how music is experienced, created, and shared. By continuing to advance the underlying technologies, addressing ethical and sustainability concerns, and exploring new applications, we can unlock the full potential of the Musical Metaverse, creating a vibrant and inclusive future for music in the digital age.

References

[1] L. Turchet, M. Lagrange, C. Rottondi, G. Fazekas, N. Peters, J. Østergaard, F. Font, T. Bäckström, and C. Fischione, "The internet of sounds: Convergent trends, insights, and future directions," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 264–11 292, 2023.

[2] M. Cobos, J. Ahrens, K. Kowalczyk, and A. Politis, "An overview of machine learning and other data-based methods for spatial audio capture, processing, and reproduction," *EURASIP Journal on Audio, Speech, and Music Process.*, 2022.

[3] L. Turchet, "Musical metaverse: vision, opportunities, and challenges," *Personal and Ubiquitous Computing*, vol. 27, 01 2023.

[4] Y. Wang, Z. Su, N. Zhang, R. Xing, D. Liu, T. H. Luan, and X. Shen, "A survey on metaverse: Fundamentals, security, and privacy," *IEEE Commun. Surv. Tutorials*, vol. 25, no. 1, pp. 319–352, 2023. [Online]. Available: https://doi.org/10.1109/COMST.2022.3202047

[5] H. Wang, H. Ning, Y. Lin, W. Wang, S. Dhelim, F. Farha, J. Ding, and M. Daneshmand, "A survey on the metaverse: The state-of-the-art, technologies, applications, and challenges," *IEEE Internet of Things Journal*, vol. 10, no. 16, pp. 14 671–14 688, 2023.

[6] C. Rinaldi, F. Franchi, A. Marotta, F. Graziosi, and C. Centofanti, "On the exploitation of 5g multi-access edge computing for spatial audio in cultural heritage applications," *IEEE Access*, vol. 9, pp. 155 197–155 206, 2021.

[7] F. Zotter and M. Frank, *Ambisonics : A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, 1st ed., ser. Springer Topics in Signal Processing, 19. Cham: Springer Nature, 2019.

[8] L. Turchet, C. Rinaldi, C. Centofanti, L. Vignati, and C. Rottondi, "5g-enabled internet of musical things architectures for remote immersive musical practices," *IEEE Open Journal of the Communications Society*, pp. 1–1, 2024.

[9] H. Choi, "Audioworklet: The future of web audio," in *ICMC*, 2018.

[10] T. Potter, Z. Cvetković, and E. De Sena, "On the relative importance of visual and spatial audio rendering on vr immersion," *Frontiers in Signal Processing*, vol. 2, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frsip.2022.904866

[11] M. A. Camilleri, "Metaverse applications in education: A systematic review and a cost-benefit analysis," *Interactive Technology and Smart Education*, 06 2023.

[12] L. Freina and M. Ott, "A literature review on immersive virtual reality in education: state of the art and perspectives," in *The international scientific conference elearning and software for education*, vol. 1, no. 133, 2015, pp. 10–1007.

[13] E. Bisso, M. S. Signorelli, M. Milazzo, M. Maglia, R. Polosa, E. Aguglia, and P. Caponnetto, "Immersive virtual reality applications in schizophrenia spectrum therapy: a systematic review," *International journal of environmental research and public health*, vol. 17, no. 17, p. 6111, 2020.

[14] R. Gupta, J. He, R. Ranjan, W.-S. Gan, F. Klein, C. Schneiderwind, A. Neidhardt, K. Brandenburg, and V. Välimäki, "Augmented/mixed reality audio for hearables: sensing, control, and rendering," *IEEE Signal Processing Magazine*, vol. 39, no. 3, pp. 63–89, 2022.

[15] A. Boem and L. Turchet, "Musical metaverse playgrounds: exploring the design of shared virtual sonic experiences on web browsers," in *2023 4th International Symposium on the Internet of Sounds*, 2023, pp. 1–9.

[16] C. Centofanti, W. Tiberti, A. Marotta, F. Graziosi, and D. Cassioli, "Taming latency at the edge: A user-aware service placement approach," *Computer Networks*, vol. 247, p. 110444, 2024.

[17] D. Fantini, G. Presti, M. Geronazzo, R. Bona, A. G. Privitera, and F. Avanzini, "Co-immersion in audio augmented virtuality: the case study of a static and approximated late reverberation algorithm," *IEEE Transactions on Visualization and Computer Graphics*, 2023.

[18] J. Wang, K. Qian, Y. Qiu, H. Zhang, and X. Xie, "A multi-attribute subjective evaluation method on binaural 3d audio without reference stimulus," *Applied Acoustics*, vol. 200, p. 109042, 2022.

[19] Z. Li, J. Wang, H. Zhang, S. Hasan, and J. Li, "Non-reference subjective evaluation method for binaural audio in 6-dof vr applications," in *2023 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. IEEE, 2023, pp. 422–427.

[20] M. Schoeffler, A. Silzle, and J. Herre, "Evaluation of spatial/3d audio: Basic audio quality versus quality of experience," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 1, pp. 75–88, 2016.

[21] K. Tsioutas, I. Doumanis, and G. Xylomenos, "A framework for understanding and defining quality of musicians' experience in network music performance environments," in *Proceedings of the Audio Engineering Society Convention 146*, 04 2019.

[22] F. Brinkmann and S. Weinzierl, "Audio Quality Assessment for Virtual Reality," in *Sonic Interactions in Virtual Environments*, M. Geronazzo and S. Serafin, Eds. Cham: Springer International Publishing, 2023, pp. 145–178. [Online]. Available: https://doi.org/10.1007/978-3-031-04021-4_5

[23] C. Schneiderwind, M. Richter, N. Merten, and A. Neidhardt, "Effects of modified late reverberation on audio-visual plausibility and externalization in ar," in *2023 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, 2023, pp. 1–9.

[24] T. Thiede, "PEAQ–The ITU Standardfor ObjectiveMeasuremenot f Perceived Audio Quality," *J. Audio Eng. Soc.*, vol. 48, no. 1, 2000.

[25] ETSI, "5g; service requirements for video, imaging and audio for professional applications (viapa)," accessed: 8-3-2024.

[26] J.-M. Jot, R. Audfray, M. Hertensteiner, and B. Schmidt, "Rendering spatial sound for interoperable experiences in the audio metaverse," in *2021 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*. IEEE, 2021, pp. 1–15.

[27] D. Thery and B. F. G. Katz, "Auditory perception stability evaluation comparing binaural and loudspeaker Ambisonic presentations of dynamic virtual concert auralizations," *The Journal of the Acoustical Society of America*, vol. 149, no. 1, pp. 246–258, 01 2021. [Online]. Available: https://doi.org/10.1121/10.0002942

[28] L. Turchet, M. Lagrange, C. Rottondi, G. Fazekas, N. Peters, J. Østergaard, F. Font, T. Bäckström, and C. Fischione, "The internet of sounds: Convergent trends, insights, and future directions," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 264–11 292, 2023.

[29] C. Centofanti, J. Santos, V. Gudepu, and K. Kondepu, "Impact of power consumption in containerized clouds: A comprehensive analysis of open-source power measurement tools," *Computer Networks*, vol. 245, p. 110371, 2024.

[30] H. Yu, M. Shokrnezhad, T. Taleb, R. Li, and J. Song, "Toward 6g-based metaverse: Supporting highly-dynamic deterministic multi-user extended reality services," *IEEE Network*, vol. 37, no. 4, pp. 30–38, 2023.

[31] L. Vignati, G. Nardini, M. Centenaro, P. Casari, S. Lagén, B. Bojovic, S. Zambon, and L. Turchet, "Is music in the air? evaluating 4g and 5g support for the internet of musical things," *IEEE Access*, 2024.

[32] L. Turchet, C. Rinaldi, C. Centofanti, L. Vignati, and C. Rottondi, "5g-enabled internet of musical things architectures for remote immersive musical practices," *IEEE Open Journal of the Communications Society*, 2024.

[33] A. Marotta, L. Valcarenghi, K. Kondepu, C. Centofanti, C. Rinaldi, F. Graziosi, and D. Cassioli, "Fiber to the room challenges and opportunities," in *2024 16th International Conference on COMmunication Systems & NETworkS (COMSNETS)*. IEEE, 2024, pp. 1139–1142.