DOI: 10.61782/fa.2025.0793

FORUM ACUSTICUM EURONOISE 2025

SOUND AI: EXPLORING CULTURE-SENSITIVE AND ENVIRONMENTALLY-CONSCIOUS APPLICATIONS FOR AI-SUPPORTED SOUND-DRIVEN DESIGN

Elif Özcan[1*] **Simone Mora**[2,5] **Fabio Duarte**[2] **Sara Lenzi**[3,4]

[1] Critical Alarms Lab, Delft University of Technology, The Netherlands
[2] Senseable City Lab, Massachusetts Institute of Technology, United States of America
[3] Deusto Design Research Group, University of Deusto, Bilbao, Spain
[4] Ikerbasque Basque Foundation for Science, Bilbao, Spain
[5] Dept. of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT

In an earlier study, we explored the intersection between human perception of sound, Machine Listening, and their potential societal and environmental impact with experts in computer science, art and philosophy, design, and ethics. As a follow-up work, we explored SoundAI applications through two MSc design student projects at TU Delft on the topics of building cultural bonds by generating new community soundscapes and increasing awareness on biodiversity by focussing on other-than-human noise sensitivity. This paper presents these two use cases with societal, environmental, and ethical considerations, and suitable AI techniques. The proposed applications manifest themselves as promoting accessibility, environmental awareness, and community well-being while mitigating ethical risks in privacy and ownership. These two design explorations allow us to imagine early AI applications for sound with their opportunities and limitations. It is our intention that SoundAI responds to the real needs of individuals, natural habitats, and other species to make a positive societal and environmental impact.

Keywords: SoundAI, sound-driven design, social design, more-than-human, responsible AI

**Corresponding author*: *e.ozcan@tudelft.nl.*

1. INTRODUCTION

Human beings have a complex relationship with daily sounds. Acoustic environments, indoor and outdoor, are a major source of information and pleasure for us as human beings conditioned to listen [1]. The quality of the acoustic environment affects not only the health and wellbeing of humans but also other species and how we interact with them through sound. Hence, sound-driven interventions are being developed to overcome the unwanted effects of noise but also to leverage desirable aspects of sonic interaction. The ultimate goal is to facilitate sonic hygiene to benefit societies so that sounds that are needed and relevant are accessible and sounds that are unwanted or harmful are mitigated through design interventions [2]. Current advances in Artificial Intelligence (AI) provide opportunities to foster sound-induced wellbeing although not without pitfalls. Specifically, Machine Listening (MLi) enables new ways of listening beyond human perception, with potential applications in urban space, environmental monitoring, socio-cultural restoration, and human-computer interactions. Our paper explores this emerging technology in line with responsible AI developments for environmentally-conscious and culturally-sensitive applications.

2. ARTIFICIAL INTELLIGENCE AND SOUND

Machine Listening **(MLi)** [3] is an AI-driven technology that enables to automatically capture, identify, classify, and predict sound events in real time. MLi has its roots in audio signal processing used for music analysis and reproduction (Music Information Retrieval), speech recognition & analysis, and environmental sound detection and classification based on auditory scene analysis. By simulating human auditory processes MLi applications are becoming able to listen to the environment in order to detect sound events (e.g., environmental sounds, music, biological signals, and other auditory phenomena [4]), gather relevant information (e.g., the type of sound sources and their changes over time), and interpret their emotional quality [5]. With the advances made possible by deep learning techniques and the increased availability of labelled audio datasets, the accuracy and efficiency of sound recognition is improving. More and more, MLi supports an ecologically-relevant interpretation and possibly prediction of the acoustic environment which, in turn, informs the design of sound-driven applications [6]. Moreover, audio-processing algorithms function at different complexities in listening and interpreting that can be described in terms of general auditory perception modes [7, 8]:

- **Passive Listening**: Simply detecting and categorizing sound (e.g., noise level monitoring for reference) without acting on the information gathered.
- **Active Listening**: Interpreting and reacting to sound contextually (e.g., distinguishing between a baby crying and a dog barking; recognizing emotions in voice for providing further support).
- **Augmented Listening**: Enhancing or modifying how sounds are heard based on the listener needs (e.g., hearing aids that filter out background noise).

These types of MLi functions are congruent with how humans listen and act in shared acoustic environments [9]: a **passive listener** acknowledges sound without action (listening-in-readiness, for example, hearing an ambulance while cooking), whereas an **active listener** anticipates sound (listening-in-search, e.g., expecting an ambulance after an emergency call). Augmented listening may support the **sound users** (listening-in-action) as the sonic world needs to be amplified while actively engaging in the sound event. AI systems can replicate these attention states, improving automation and user interaction. Furthermore, this notion also brings in multiple listener perspective as the incoming sound can be interpreted differently by different users.

On the other hand, with the rise of generative audio (i.e., generative AI for sound design and music production [10]) that rely on past data and information streams, algorithms are becoming capable of creating new, realistic audio content. This includes generating music, sound effects, and synthetic voices that mimic human speech. Current solutions that leverage Large Language Models (LLM) for the real-time generation of speech, audio, and music are available in open source and include Meta Audiocraft and Google MusicFX (for an updated list of currently available technological solutions refer to [11]). However, the more advanced and publicly available SoundAI technologies become, the more problematic it is to understand and control its function for responsible AI usage. Much has been learnt about the harms of generative AI in fields such as computer vision [12]; thus, ethical considerations should always be the guiding principle in developing new sound-driven applications in public and private domains, including public health, urban planning, conservation, healthcare, and accessibility.

3. SOUNDAI: RISKS AND OPPORTUNITIES

When Machine Listening and Generative AI for sound are integrated into broader and more advanced AI systems, it results in a new approach to audio processing that we call **SoundAI**. SoundAI will soon enable machines to detect, analyse, and generate contextualized audio data, including environmental sound analysis, auditory augmentation, and soundscape design. Thus, SoundAI represents an emerging multi-disciplinary field rather than merely an application of AI in the audio domain.

In an earlier study we explored the intersection of AI and sound, particularly through MLi, highlighting both opportunities and risks [13]. Discussions with multidisciplinary experts (artists, acousticians, sociologists, computer scientists, architects, and urban planners) and design researchers led to key insights about how AI-driven listening can be applied and its ethical implications. In general, the study underscored the need for standardized terminology and responsible development in SoundAI. As an emerging field, SoundAI should learn from the current debates in AI applications for visual and image recognition to ensure that ethical considerations such as fair human - AI relationships, data privacy, and data ownership are addressed while optimizing AI for urban environments, ecosystems, and well-being [14].

Among the envisaged benefits, experts highlighted the potential of SoundAI for *Other-than-human listening* use cases i.e., the interpretation of sounds from non-human living entities (e.g., for animal communication and environmental monitoring), which would deepen human-nature connections. In contrast, *More-than-human listening* [15] supports augmented hearing, personalised auditory experiences, and improved accessibility for individuals with hearing impairments.

Potential harms include listening biases (e.g., due to limited datasets and human labelling) that may lead to inaccurate sound representations or the stigmatisation of specific communities. In both public and private settings, passive sound monitoring raises privacy concerns, as AI may access personal conversations. Questions of data ownership remain: should audio data belong to deploying organizations, individuals, or AI developers? Finally, experts identified areas where, if designed responsibly, SoundAI has the potential to enhance human life, protect the environment, and create more responsible solutions.

Our previous study [13]lacked an analysis of the technologies behind SoundAI and the feasibility of its application to real-world use cases. This paper presents the results of an early investigation into applied solutions conducted as part of the MSc graduation studies at the Faculty of Industrial Design Engineering (TU Delft). Specifically, the two projects here described explore SoundAI-based solutions to monitor and protect biodiversity in the urban space through MLi, and foster intangible heritage and cultural integration of migrant communities by augmenting listening capabilities through GenAI.

4. AI-SUPPORTED SOUND-DRIVEN DESIGN

Sound-driven design is a human-centred approach that identifies sound as a functional, informative, and interactive medium to shape user experiences, enhance technological systems, and optimize environments [16]. It integrates sonic elements into physical products, digital interfaces, and social systems to solve practical challenges while considering cognitive, emotional, and cultural impacts. From hospitals to entire cities, sound-driven design can optimize how humans interact with their surroundings, making the world more intuitive, efficient, and sustainable [17]. In this paper, we focus on the topic of conservation of biodiversity and cultural heritage as a sound-induced issue but also an opportunity to design for, with the support of SoundAI. Below are two projects we conceptualised as a response to such socio-technological and environmental problems of today.

4.1 Other-than-human noise sensitivity

Noise pollution in urban settings presents multifaceted challenges, impacting both human well-being and ecological stability. **Human-centric issues** include psychological and physiological distress, as noise exposure exacerbates sleep disturbances, cognitive impairment, and cardiovascular diseases. The complexity of urban soundscapes influences perception; moderate diversity in sound sources can reduce annoyance, while excessive or minimal complexity increases discomfort, particularly when dominated by traffic noise. Conversely, natural sounds have restorative effects, mitigating stress and improving overall well-being [18]. Moreover, a healthy shared acoustic environment is proven to protect communities from environmental harm and is conducive to the physical, mental and social benefit of its human and non human inhabitants [19]. On the contrary, noise pollution tends to disproportionately affects historically marginalized communities, exacerbating environmental injustice [20].

Ecological disruptions arise as anthropogenic noise interferes with wildlife communication, behaviour, and habitat use. Chronic exposure above 85 dB causes hearing loss and signal masking, preventing species from detecting predators, prey, or mates, ultimately affecting survival [21]. In the urban context, non-human species have been forced to adapt their communication strategies to the pace of the urbanised and industrialised world [22]. Noise pollution also heightens flight and vigilance responses, with animals in noisy areas exhibiting increased stress and reduced foraging efficiency [23]. **Habitat and ecosystem-level impacts** are profound. Anthropogenic noise below 250 MHz disrupts species reliant on acoustic signals, such as marine mammals using sonar [24]. Some species, like urban black-tufted marmosets, actively avoid noisy areas, despite food availability [25]. Protected areas are not immune; e.g., 63% of U.S. conservation zones experience noise levels twice the ambient background, altering predator-prey interactions and community dynamics [26].

Despite its severity, noise remains a reversible pollutant. Reduced human activity during COVID-19 lockdowns demonstrated the potential for soundscape restoration and ecological recovery [26]. Consequently, addressing noise pollution through bioacoustic and soundscape monitoring, and sound-sensitive urban design can lead to sustainable coexistence. In this project, we explored how MLi technologies could be applied to novel solutions that de-centre the discourse around noise sensitivity and noise pollution from humans, toward monitoring applications and ultimately, policies that take into account the needs of other species.

4.1.1 Bioacoustic monitoring & relevant AI models

Bioacoustic monitoring via MLi technologies has gained prominence due to technological advancements, enabling the study of animal populations, biodiversity, and ecosystem health. This method records and analyses sounds produced by animals, offering a non-invasive approach (i.e., passive listening) to data collection. Key advantages include: **minimal disturbance** (wildlife remains unaffected by monitoring); **species identification** (unique vocalizations allow researchers to differentiate species) and **behavioural insights** (monitoring reveals habitat use, population dynamics, and ecosystem interactions [27]). Recent advancements in automated analysis enhance efficiency of real-time monitoring of the acoustic environment. AI-driven models can recognize species automatically, reducing reliance on manual data processing [28]. Since 2013, bioacoustic studies have expanded to include plant life, broadening its ecological applications. Conservationists also employ this technology to detect illegal activities such as poaching and deforestation. Bioacoustic monitoring is proposed as a tool for urban planning, allowing for soundscape assessments that inform sustainable development [29]. By providing real-time data on biodiversity and environmental health, it offers a valuable approach for mitigating noise pollution and enhancing urban ecosystems.

The most basic AI models include **reactive machine AI** and **rule-based systems.** Reactive machine AI, such as basic face detection algorithms, perform predefined tasks without learning. Basic machine learning models are used for simple tasks: **linear regression** for continuous values and **logistic regression** for binary classification. **Decision trees** offer classification and regression capabilities. In SoundAI, linear regression predicts audio features, logistic regression classifies sounds, and decision trees categorize sounds in urban ecosystems. Rule-based systems provide event-based responses [30].

4.1.2 Implications for Sound-driven Design

Long-term adaptation of wildlife to noise pollution remains under-explored, offering a research gap in biodiversity conservation through sound-based data. Additionally, the integration of bioacoustic data into urban planning is limited. AI-supported sound-driven design in the context of biodiversity preservation aims to mitigate noise pollution while enhancing desirable acoustic environments for other-than-humans. Bioacoustic monitoring offers a non-invasive method to assess biodiversity and ecosystem health, presenting opportunities for integrating sound-informed data into urban planning. However, limitations persist in current applications as existing studies tend to focus on specific geographic regions and a narrow subset of species, limiting their broader applicability. The complexity of analysing vast bioacoustic datasets poses technological challenges, requiring advanced AI-driven solutions for real-time interpretation.

The implications of these findings emphasize the urgent need for **sound-sensitive urban policies** and an opportunity for MLi-driven innovation that provides policymakers with the necessary knowledge to inform decisions. Incorporating green spaces and natural sound buffers can counteract noise pollution, supporting both human health and wildlife adaptation. However, effective implementation requires interdisciplinary collaboration between urban planners, ecologists, and acoustic scientists. Further research should address species resilience to chronic noise, explore bioacoustic data integration into smart city infrastructure, and assess noise sensitivity across diverse populations. As urbanization accelerates, designing with sound in mind is crucial for fostering healthier and more ecologically sustainable urban environments.

4.2 The Cultural Soundscape Party

Cultural sounds embody social, historical, national, and spiritual memories tied to specific places. They create a distinct sense of identity, reinforcing connections to home environments [31]. These auditory experiences, such as local dialects, religious chants, and folk ceremonies, strengthen community bonds and shape emotional atmospheres. Beyond nostalgia, these sounds serve as intangible cultural heritage, acting as mnemonic devices that remind individuals of their cultural roots [32]. For immigrants, familiar soundscapes act as sonic memories that maintain cultural identity in foreign settings [33]. However, the disruption of familiar soundscapes can cause disorientation, requiring immigrants to develop "soundscape competence" to navigate new auditory environments. This process often involves nostalgic recollections and the creation of new soundscapes that blend original and adopted cultures [34]. Soundscape studies provide insights into how immigrants negotiate their cultural identities and adapt to urban acoustic communities [35].

While people are tuned to the sounds that represent their own culture, sound artists explore how to offer newly composed soundscapes, e.g., in parks, hospitals, or other public spaces. Soundscape composition involves artistic arrangements of recorded sounds to document lived environments [36]. These compositions provide insights into immigrant experiences and socio-political dynamics [37]. Activist projects, such as *Sounds of Aliyah* [35], use soundscapes to articulate identity and protest cultural erasure [38]. Additionally, soundscape compositions facilitate cross-cultural dialogue, as seen in the *azan* (i.e., the call to prayer for Islam) debates in the Netherlands, where soundscape discussions address religious freedom and social belonging. Future soundscape compositions should actively explore ways to enhance immigrant identities and foster inclusivity in diverse urban settings. In this use case, we propose the *Cultural Soundscape Party* as an initiative to celebrate immigrant heritage through soundscape composition based on GenAI.

4.2.1 AI technology supporting soundscape composition

In order to compose new soundscapes, both analytic and generative methods are needed such as data acquisition and processing, sound source identification and classification, and sound synthesis and modification. AI facilitates soundscape composition by analysing and processing audio data (i.e., active listening) based on signal processing methods. Techniques such as spectral analysis, rhythmic pattern recognition, and feature extraction allow the identification of cultural sounds. Emotional evaluation models using spectrograms help assess affective qualities of soundscapes (e.g., nostalgic), enhancing their relevance for immigrant communities. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) enable the identification of cultural sounds in complex environments [39]. Using labelled datasets of traditional and urban sounds, AI reconstructs culturally meaningful soundscapes, allowing diaspora communities to reconnect with their auditory heritage (i.e., augmented listening). More specific AI methods that support soundscape composition are based on generative and text-guided methods. By integrating these technologies, AI tools facilitate the reconstruction of dynamic cultural soundscapes that reflect the evolving identities of immigrant communities: **Generative Adversarial Networks (GANs)** create realistic soundscapes by learning intricate patterns in audio data. **Variational Autoencoders (VAEs)** create diverse and controllable audio samples, useful for exploring latent soundscapes. **Diffusion Models** generate audio from prompts, such as text or other audio inputs, enabling creative exploration of sound. **Text-to-Sound Synthesis** used models like MusicLM to synthesize audio from text descriptions by leveraging shared embedding spaces. **Sound Mixture Editing Systems** allow users to modify sound mixtures based on textual instructions, such as extracting specific sounds or adjusting volumes.

4.2.2 Implications for Sound-driven Design

Sound is a critical component of cultural identity, acting as a repository of collective memory, particularly for immigrant communities being on-boarded into new environments. Traditional songs, spoken narratives, and ambient cultural sounds provide deep insight into heritage and belonging. However, urban nightlife often lacks diverse cultural representation, limiting the ability of communities to share and sustain their sonic heritage. As a response, this project curates soundscapes that merge traditional and contemporary elements. AI tools enhance this process by processing sound contributions and assisting sound artists and DJs in creating adaptive, participatory compositions.

The **Cultural Soundscape Party** has been conceptualised as an AI-driven immersive event that preserves and celebrates cultural heritage through participatory sound collection, AI-assisted composition, and interactive storytelling. Community members contribute personal audio memories, which AI processes to create high-quality sound materials. Using GAN-generated audio and text-guided sound editing, DJs and sound artists blend traditional sounds with contemporary beats, ensuring authenticity while fostering engagement. The event unfolds through storytelling, live DJ performances, and real-time audience interaction, enhancing cultural exchange. A post-event digital archive preserves playlists, narratives, and highlights, serving as an educational resource for long-term cultural soundscape accessibility. This initiative promotes cross-cultural dialogue, enriches nightlife experiences, and leverages AI for cultural preservation.

We acknowledge the technological and design limitations: AI-generated compositions rely on mass training data, which, if insufficiently diverse, risks misrepresenting cultural sounds. Moreover, automated sound blending may not fully capture the emotional and contextual depth of human storytelling. While techniques such as GAN-generated audio and text-guided sound editing improve authenticity, they cannot entirely replace human artistic intuition and cultural expertise. The integration of AI in cultural soundscapes introduces ethical challenges (such a data bias and cultural representation, authorship and copyright, and cultural sensitivity and appropriation) that require proactive solutions. AI models must be trained on culturally diverse datasets to prevent biased representations that oversimplify or distort traditional sounds. Guidelines must clarify intellectual property rights for AI-generated cultural compositions, ensuring fair recognition of community contributions and artistic labour. AI-generated soundscapes must avoid stereotypical portrayals and instead prioritize contextual authenticity to prevent cultural appropriation and ensure ethical storytelling.

By integrating speculative design approaches, AI-driven sound interventions can challenge stereotypes and broaden the range of cultural sounds represented. Future research should explore participatory frameworks that allow cultural communities greater control over AI-enhanced compositions, ensuring respectful and nuanced sonic representations.

5. DISCUSSION

The two concepts presented in this paper are meant to further define real-world use cases where SoundAI can have a positive impact and provide solutions that benefit humans, society, and the broader environment. The projects illustrate distinct applications of SoundAI grounded in different modes of machine listening. *Noise Sensitivity* employs passive and active MLi to monitor urban biodiversity through bioacoustic data, exemplifying Other-than-human listening. It addresses ecological well-being by detecting stress signals and behavioural changes in wildlife, informing sound-sensitive urban planning. *Cultural Soundscapes* by contrast, employs augmented listening and GenAI to recreate and archive cultural soundscapes for displaced communities, enabling More-than-human listening. It amplifies human auditory experience, fostering belonging and memory among migrants. While both projects showcase opportunities for sustainable and inclusive design, they also surface key risks: *Noise Sensitivity* must navigate ecological data bias and ownership, while *Cultural Soundscapes* raises concerns about cultural sensitivity, consent, and responsible use of generative sound tools. Together, they demonstrate SoundAI's capacity to reshape how we listen, design, and coexist—across species, cultures, and technologies.

The two projects, albeit embryonic, share a vision of SoundAI as a facilitator for sound-driven design strategies grounded in the following principles:

- **Sound as an Environmental and Social Factor** shaping human and non-human environments.
- **Sound's Influence on Behaviour** enabling species to migrate or human interactions and rituals.
- **Interdisciplinary Nature of SoundAI** including biology, environmental science, anthropology, urban studies, and technology.
- **Ethical and Policy Considerations** regarding conservation, heritage, urbanisation, environmental well-being.

Future developments of both projects would necessarily involve the analysis of currently available solutions for both technologies or else the development of customised solutions. Such technologies, however, cannot be identified or designed without the broader consequences in mind: technologies that limit the risks for privacy violation should be excluded (hence privileging, for instance, so called edge-solutions that process audio data on local devices), and issues of cultural ownership and copyright should be addressed at an earlier stage of the design.

6. CONCLUSIONS AND FUTURE DIRECTIONS

Both the *Cultural Soundscape* and *Noise Sensitivity* projects exemplify SoundAI-based design as a promising tool for addressing today's socio-technological and environmental challenges, specifically in the context of conservation of cultural heritage and biodiversity. Each project explores how sound—beyond sensory and aesthetic appreciation —functions as an interface that can inform, interact, and intervene in complex systems, offering intuitive and emotionally relevant experiences through AI-supported design.

The *Cultural Soundscape* project embraces the cognitive and emotional impact of sound, focusing on how AI-generated soundscapes preserve and celebrate the intangible heritage of migrant communities. It transforms cultural sounds into interactive, immersive experiences—bridging tradition and technology—while also providing long-term archival and educational value. Here, sound is used not only for entertainment, but as a mnemonic and social tool that promotes identity, inclusion, and inter-cultural dialogue.

Conversely, the *Noise Sensitivity* project expands the scope of sound-driven design beyond the human-centric focus. It reframes noise pollution as a multispecies issue, applying bioacoustic monitoring and AI classification systems to design ecologically urban acoustic environments. In this case, sound becomes a functional and diagnostic medium—not just expressive—used to assess biodiversity, detect environmental threats, and inform urban policy. This project highlights how sound-driven design can align technological innovation with environmental ethics, advocating for more inclusive, multispecies futures.

While both projects leverage SoundAI technologies to analyse and synthesize complex acoustic environments, they face similar limitations—such as data bias, technological scalability, and the need for interdisciplinary integration. Yet, they also reveal sound's adaptive and generative potential in making cities more humane and ecologically responsive. Altogether, these projects demonstrate that sound-driven design can operate at the intersection of human condition and acoustic environments, shaping not only how we hear the world, but how we care for it.

7. ACKNOWLEDGMENTS

This ongoing research is partially funded by the MIT MISTI Global Seed Fund. SL's work is partially funded by the Marie Skłodowska-Curie Cofund Programme of the European Commission (H2020-MSCA-COFUND-2020-101034228-WOLFRAM2).

8. REFERENCES

[1] J. Kang and B. Schulte-Fortkamp, *Soundscape and the built environment*, vol. 525. CRC press Boca Raton, FL, USA:, 2016.

[2] A. Findeli, "Design history and design studies: Methodological, epistemological and pedagogical inquiry," *Design Issues*, vol. 11, no. 1, pp. 43–65, 1995.

[3] J. E. K. Parker and S. D. and, "'all possible sounds': speech, music, and the emergence of machine listening," *Sound Studies*, vol. 9, no. 2, pp. 253–281, 2023.

[4] C.-E. Cella, "Machine listening intelligence," *arXiv preprint arXiv:1706.09557*, 2017.

[5] S. Lenzi, S. Spagnol, and E. Özcan, "Improving the quality of the acoustic environment in neonatal intensive care units: a review of scientific literature and technological solutions," *Frontiers in Computer Science*, vol. 5, p. 1156693, 2023.

[6] K. Zaman, M. Sah, C. Direkoglu, and M. Unoki, "A survey of audio classification using deep learning," *IEEE Access*, vol. 11, pp. 106620–106649, 2023.

[7] C. Salter, *Entangled: Technology and the Transformation of Performance*. Cambridge, MA: MIT Press, 2010.

[8] B. Kane, *Sound Unseen: Acousmatic Sound in Theory and Practice*. Oxford, UK: Oxford University Press, 2014.

[9] E. Özcan, C. L. H. Broekmeulen, Z. A. Luck, M. van Velzen, P. J. Stappers, and J. R. Edworthy, "Acoustic biotopes, listeners and sound-induced action: a case study of operating rooms," *International Journal of Environmental Research and Public Health*, vol. 19, p. 16674, 2022.

[10] R. Mitra and I. Zualkernan, "Music generation using deep learning and generative ai: A systematic review," *IEEE Access*, vol. 13, pp. 18079–18106, 2025.

[11] S. Latif, M. Shoukat, F. Shamshad, M. Usama, H. Cuayáhuitl, and B. W. Schuller, "Sparks of large audio models: A survey and outlook," *arXiv preprint arXiv:2308.12792*, 2023.

[12] A. Sarkar, "Enough with "human-ai collaboration"," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI EA '23, (New York, NY, USA), Association for Computing Machinery, 2023.

[13] S. Lenzi, E. Özcan, S. Mora, M. Mazzarello, A. Haatveit, and F. Duarte, "What in the world do we hear? understanding public and private spaces through soundai," in *DRS2024: Boston, 23–28 June* (C. Gray, P. Hekkert, L. Forlano, and P. Ciuccarelli, eds.), (Boston, USA), 2024.

[14] A. Sarkar, "Enough with "human-ai collaboration"," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23)*, (New York, NY, USA), pp. Article 415, 1–8, Association for Computing Machinery, 2023.

[15] S. Lenzi, J. Sádaba, and A. Retegi, "Climate adaptation in urban space: the need for a transdisciplinary approach," *Frontiers in Sustainable Cities*, vol. 7, p. 1562066, 2025.

[16] S. D. Monache, N. Misdariis, and E. Özcan, "Semantic models of sound-driven design: Designing with listening in mind," *Design Studies*, vol. 83, p. 101134, 2022.

[17] E. Özcan, W. J. Rietdijk, and D. Gommers, "Shaping critical care through sound-driven innovation: introduction, outline, and research agenda," *Intensive Care Medicine*, vol. 46, no. 3, pp. 542–543, 2020.

[18] A. Mitchell, M. Erfanian, C. Soelistyo, T. Oberman, J. Kang, R. Aldridge, J.-H. Xue, and F. Aletta, "Effects of soundscape complexity on urban noise annoyance ratings: a large-scale online listening experiment," *International journal of environmental research and public health*, vol. 19, no. 22, p. 14872, 2022.

[19] U. Ruiz Arana, "Multispecies soundscapes: a more-than-human approach," in *Proceedings of the Inter-Noise Conference 2024*, (Nantes, France), pp. 2743–2749, 2024.

[20] J. Nelson-Olivieri, T. Layden, E. Antunez, A. Khalighifar, M. Lasky, T. Laverty, K. Sanchez, G. Shannon, S. Starr, A. Verahrami, *et al.*, "It's not just noise: The consequences of inequitable noise for urban wildlife," 2023.

[21] G. Arcangeli, L. G. Lulli, V. Traversini, S. De Sio, E. Cannizzaro, R. P. Galea, and N. Mucci, "Neurobehavioral alterations from noise exposure in animals: a systematic review," *International Journal of Environmental Research and Public Health*, vol. 20, no. 1, p. 591, 2022.

[22] L. Pschetz, M. Gebker, S. Wieland, and M. Bastian, "Designing temporal ecologies: Reframing multispecies temporalities through design," in *DRS2024: Boston, 23–28 June* (C. Gray, E. Ciliotta Chehade, P. Hekkert, L. Forlano, P. Ciuccarelli, and P. Lloyd, eds.), (Boston, USA), 2024.

[23] K. A. Zeller, M. A. Ditmer, J. R. Squires, W. L. Rice, J. Wilder, D. DeLong, A. Egan, N. Pennington, C. A. Wang, J. Plucinski, *et al.*, "Experimental recreationist noise alters behavior and space use of wildlife," *Current Biology*, vol. 34, no. 13, pp. 2997–3004, 2024.

[24] S. Ambrose, "Effects of noise pollution on urban wildlife," January 2022. Retrieved on 2025-04-04.

[25] K. Iveson and D. Hochuli, "How urban soundscapes affect humans and wildlife — and what may have changed in the hush of lockdown," *The Conversation*, n.d. Retrieved from The Conversation, accessed on 2025-04-04.

[26] A. C. Kok, B. W. Berkhout, N. V. Carlson, N. P. Evans, N. Khan, D. A. Potvin, A. N. Radford, M. Sebire, S. Shafiei Sabet, G. Shannon, *et al.*, "How chronic anthropogenic noise can affect wildlife communities," *Frontiers in Ecology and Evolution*, vol. 11, p. 1130075, 2023.

[27] P. Holderried, "Bioacoustics – keeping your ears open in nature," 2023. Online version published on 25.09.2023.

[28] J. Salamon, J. P. Bello, A. Farnsworth, M. Robbins, S. Keen, *et al.*, "Towards the automatic classification of avian flight calls for bioacoustic monitoring," *PLoS ONE*, vol. 11, p. e0166866, 2016.

[29] Carbon Rewild, "Bioacoustic monitoring," n.d. Retrieved on 2025-04-04.

[30] GeeksforGeeks, "Common ai models and when to use them," *GeeksforGeeks*, April 2024. Retrieved on 2025-04-04.

[31] H. Westerkamp, "The local and global 'language'of environmental sound," in *Proceedings of Sound Escape: An International Conference on Acoustic Ecology*, vol. 1, pp. 1–8, 2000.

[32] K. Kato, "Soundscape, cultural landscape and connectivity," *Sites: a journal of social anthropology and cultural studies*, vol. 6, no. 2, pp. 80–91, 2009.

[33] B. Martin, "Soundscape composition: Enhancing our understanding of changing soundscapes," *Organised Sound*, vol. 23, no. 1, pp. 20–28, 2018.

[34] L. U. Marks and D. Polan, *The skin of the film: Intercultural cinema, embodiment, and the senses*. Duke University Press, 2020.

[35] C. J. Weleminsky, *Sonic Diaspora: Exploring Migration Through Interdisciplinary Soundscape Composition*. PhD thesis, Goldsmiths, University of London, 2021.

[36] M. Thorogood, P. Pasquier, and A. Eigenfeldt, "Audio metaphor: Audio information retrieval for soundscape composition," *Proc. of the Sound and Music Computing Cong.(SMC)*, pp. 277–283, 2012.

[37] A. McCartney, "Circumscribed journeys through soundscape composition," *Organised Sound*, vol. 7, no. 1, pp. 1–3, 2002.

[38] T. Rennie, "Socio-sonic: An ethnographic methodology for electroacoustic composition," *Organised Sound*, vol. 19, no. 2, pp. 117–124, 2014.

[39] S. Garg, T. Sehga, A. Jain, Y. Garg, P. Nagrath, and R. Jain, "Retracted: Urban sound classification using convolutional neural network model," in *IOP Conference Series: Materials Science and Engineering*, vol. 1099, p. 012001, IOP Publishing, 2021.