# Speed Makes Us Oblivious: A Slow and Responsible Approach to Engaging with Audio Sample Archives

TRISHA KHALLAGHI, University of Bristol, United Kingdom
PETE BENNETT, University of Bristol, United Kingdom
ATAU TANAKA, University of Bristol, United Kingdom and Goldsmiths University of London, United Kingdom

**Abstract**

AI and the culture of speed it accelerates has significantly influenced music consumption, resulting in an economy characterised by excess, and raises important questions around consent, data provenance, and accessibility. We introduce Slow Samples, a web-based platform built on ethical principles, designed for the responsible sharing, exploration, and documentation of audio sample collections, with consent embedded from the start and accessibility beyond academic contexts as core commitments. Initial artist interviews informed the design, highlighting ongoing concerns regarding provenance, cultural appropriation, and the ethical implications of AI in music. Rooted in slow consumption, intentional friction, slow AI, and design by refusal, the platform opposes infrastructures that prioritise efficiency over context and scale over community, which is realised through extensive required metadata, timed microboundaries, provenance tracking, and constrained AI tooling. We argue that design can be a site of resistance, and that small acts of refusal in infrastructure lay the foundation for alternative practice.

Additional Key Words and Phrases: Slow Technology, Community Archives, Audio Samples, Ethical Design, Design By Refusal, Responsible AI, Slow AI, Data Provenance

## 1 Introduction

Sometimes conspicuous, sometimes covert, the integration of artificial intelligence within contemporary musical practice has become widespread in recent years – flagged in certain contexts, while concealed in others. This influence can be seen across both creative and commercial domains, from composition to production [18], and performance [60] to distribution [12, 40]. The diffusion of AI across this ecosystem raises several questions about the ethics, attribution, and creative agency in practice. Applying AI responsibly in this area encompasses a multiplicity of concerns, ranging from explainability [7] to consent [11], and accessibility to sustainability [46]. This paper primarily foregrounds one critical dimension of the many layers involved: responsible practices in the context of audio sample archives and the AI processes that transform them.

We are presently able to consume faster than ever. Modern systems, specifically those driven by platform capitalism, enable abundance at scale through algorithmic feeds that rarely run dry. This feedback loop reinforces habits of rapid consumption, often at the expense of careful attention and reflection. In musical contexts, this can mean overlooking significant context, such as authorship, historical lineage, cultural provenance and artistic intention. For audio samples, especially, this type of hyperconsumption can manifest as an erasure of these dimensions, leaving only a trail of decontextualised fragments to consume, which prompts a critical question: what are we too fast to notice?

In contrast, the philosophy of slow consumption can be understood as an antidote to this system. The idea itself is not a new phenomenon, having originated in 1986 as a protest against the opening of a McDonald's at the Spanish Steps in Rome, where chefs campaigned for "the right to pleasure" in eating [13]. What emerged was a demand to value craft over speed, context over convenience, and quality over quantity. The principles that underpin the Slow Food manifesto, and movements born from it, reflect a broader credo that prioritises intentionality, sustainability and transparency, which can now be found in fashion, technology and beyond [2, 22, 35, 45]. In music, we find *Deep Listening* [48], which occupies an adjacent space to this philosophy, reflecting a convergent response to similar forces of acceleration. Both emphasise intentionality, focused engagement and resisting passive consumption, but reject speed and distraction. They are connected practices that form a resistance to the same zeitgeist.

This work extends slow consumption principles into the realm of audio sample archives and AI-driven sample processing – a domain central to contemporary music practice, in which samples constitute an often-familiar building block of music-making. By grounding the work in the principles of slow consumption, it demonstrates how these principles can address the

Authors' Contact Information: Trisha Khallaghi, University of Bristol, Bristol, United Kingdom; Pete Bennett, University of Bristol, Bristol, United Kingdom; Atau Tanaka, University of Bristol, Bristol, United Kingdom and Goldsmiths University of London, London, United Kingdom.


*NIME '26, June 23–26, 2026, London, UK*

challenges of hyper-consumption and decontextualisation inherent in sample-based musical practice while also considering how generative AI further obscures context when processing samples.

We take a complementary, practice-driven approach to current responsible AI research in music by developing a platform where design guidelines and implementation evolved in tandem, with each shaping the other. We present an online platform, **Slow Samples**, designed to encourage thoughtful consumption of audio archives and to make processes applied to them as transparent as possible, demonstrating how slow consumption principles can be practically applied in this field to encourage ethical discovery and engagement. The platform serves multiple communities: practitioners seeking ethically sourced materials, contributors who wish to share audio on their own terms, and researchers interested in responsible AI, transparency, and data governance.

## 2 Related work

The following section situates this work both theoretically and practically by tracing a path starting from the culture around speed and audio samples, through to a brief overview on technical and responsible applications of AI in music, followed by examples of existing platforms and archives.

### 2.1 Taking it slow

Building on the philosophy of slow consumption, it is important to emphasise that the movements born from it have emerged as a form of resistance rather than a rejection of technology [51]. Slowness is not incompatible with innovation; these movements do not look backwards simply out of nostalgia, they look forward with intention. Slow food resists the industrial homogenisation of food and is rooted in the appreciation for preserving culture and modern preservation techniques [13]; slow fashion rejects fast manufacturing cycles but embraces sustainable manufacturing processes [22]. Slowness here is not about retreating from contemporary practice but about making deliberate choices within it. Unlike technical frameworks requiring specialised knowledge, slow principles connect to explainable everyday experiences.

*2.1.1 Deep listening.* Deep listening frames musical engagement as a meditative act that encourages reflection on both the surrounding sonic environment and the self. Pauline Oliveros [48] defines it as a practice "intended to expand consciousness to the whole space/time continuum of sound/silences", presenting listening as an all-encompassing experience rather than a process oriented toward any specific end goal. It shares similarities with slow consumption in its emphasis on intentionality and rejection of thoughtless consumption. They differ in that deep listening places fundamental importance on the perception of sound, whereas slow consumption directly addresses systems that shape consumption. Together, they provide a strong foundation for addressing conscious engagement at both perceptual and systemic levels.

*2.1.2 Slow technology.* Slow technology offers a framework for understanding the relationship between time and interaction design. Originally introduced by Hallnäs and Redström [28] as a provocation, it challenges dominant HCI paradigms focused on efficiency and speed, emphasising reflection over productivity, proposing that interactive technologies should not always aim to "save time", but instead "take time" in meaningful ways. A key concept is making time visible through methods such as gradual unfolding and intermittent attention. Odom et al. [45] builds on this framework but moves the discussion from the philosophical to critical reflection within contemporary HCI practice, touching on sustainability, digital materiality, and technology built to last. Together, they reposition time as a design material, framing slowness as an experiential quality and a response to technological acceleration.

*2.1.3 Slow AI.* Little academic research has been explicitly framed as *slow AI*. Outside of academia we find AIxDESIGN, a collective from different backgrounds, who have initiated a research and community project entitled Slow AI [2]. The project has interrogated narratives associated with AI development in Silicon Valley and Big Tech, encompassing community sessions, workshops, and artist collaborations, bringing together critical AI researchers, designers, and artists. Several research strands have been explored. *Small AI* critically examines the dominance of large-scale AI models, foregrounding issues such as environmental impact and cultural erasure, advocating for community-governed, sustainable alternatives. *Esoteric AI* challenges the boundaries of enchanted and disenchanted perceptions of AI through a feminist lens [57]. *Ancestral AI* explores alternative approaches to rapid, always-on development cycles, drawing on knowledge from previous generations, including non-Western and Indigenous governance models and archiving practices [44].

### 2.2 Intentional friction & microboundaries

Traditionally, user experience design aims to maximise engagement and minimise friction. Bagnara and Pozzi [4] argue that there is an excessive focus on designing for *optimal flow*, i.e., seamless and immersive interaction, while neglecting the importance of reflection. They propose a shift from task performance to person-centred design, considering the human need for reflection, meaning-making and detachment, outlining three design directions: designing for pauses; designing

for detachment; designing for serendipity. Galak et al. [26] and Cox et al. [16] argue that frictionless design can produce mindless interaction with negative consequences, while intentional friction can redirect attention without requiring extreme interventions.

Whereas slow technology operates at the scale of whole systems, microboundaries [16] offer a more focused intervention: small, intentional moments that interrupt "automatic" behaviours at specific points within interactions, encouraging reflection without adversely affecting the user experience. A study by Mejtoft et al. [37] evaluated the impact of adding explanations to an application with built-in frictions, compared to a version without friction. Adding explanations resulted in a 67% increase in both understanding and satisfaction, with 87% of participants preferring friction. The version without friction led to more varied responses, suggesting people reacted more individually with less information. The study involved students with advanced UX knowledge who chose to participate, which the authors note may limit how widely the results apply.

### 2.3 Sample-based practice

Sampling is one of the most culturally embedded practices in contemporary music production, signifying more than the reuse of existing audio. What began as an experimental process within musique concrète in the 1940s, through cutting and looping magnetic tape, later became a core element of hip-hop. Samples come in many forms, ranging from field recordings and synthesised sounds to audio clips from film, drum breaks, and Foley. The amen break, a short drum break from "Amen Brother" by The Winstons, became so widely adopted that it gave rise to entire genres [30]. Field recordings occupy a similarly resonant position, one in which the act of gathering sound is inseparable from its environment, context, and moment in time.

Noel-Hirst et al. [42] discusses how sampling fits into generative AI systems, arguing that the "value of sampling is a meaning-making process which engages artists with their environment (i); explores boundaries of reality and fiction (ii); forms and reflects community (iii); and foregrounds agential relationships (iv)". This touches several areas of interest within this work: transparency in data provenance, multimodal representations of broader cultural context, explainability of systems, and design driven by community benefits rather than technical "sense".

### 2.4 Generative AI & music

Deep learning approaches to music generation span a range of architectures, each better suited to different musical tasks [38]. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) architectures perform well for melody and note-level composition using MIDI. Generative Adversarial Networks (GANs) have been prominent in multi-track composition and waveform synthesis [19], with models such as MuseGAN [20] excelling in multi-instrument generation. Variational Autoencoders (VAEs) [33] introduced structured latent spaces well-suited to interactive music making [52]. More recently, transformer-based models [56] such as MusicGen [15] and MusicLM [1] support multimodal conditioning such as text-to-music generation. Diffusion models offer stable training and high-fidelity outputs, excelling in tasks such as creating short samples of realistic and specific sounds [34], but require significant resources to train and run. A recurring concern across all of these systems is the datasets on which they are trained; large-scale models typically rely on data collected at volume, raising issues of consent, bias, and the marginalisation of less represented musical traditions [9].

### 2.5 Responsible AI & music

Responsible AI in music concerns the creation, evaluation, and implementation of AI systems in music, prioritising transparency, fairness, accountability, and sustainability, grounded in human-centred values. Recent research broadly focuses on the legal, social, ethical, and technical dimensions of generative systems. An analysis of Suno [55] highlights copyright, authorship and data governance challenges in the context of generative models trained on copyrighted recordings [41]. Spawning [14], launched by Holly Herndon and Mat Dryhurst, argues that artists should be able to opt out of music datasets when their work is used without consent. Empirical studies with professional musicians raise issues of exploitation, poor value distribution and job displacement as core dimensions of responsible use [31]. Research highlights that Western classical and pop music dominate training data, marginalising less represented communities [9, 58], and calls for inclusion and bias mitigation to inform policy and governance [8]. Research into trustworthiness of creative systems argues that XAI methodologies need to be context-specific and adaptable to different stakeholders [50]. The Compositional Provenance Framework [53] addresses what the author calls an "attribution crisis" in AI-generated music, arguing that existing copyright laws cannot cope with the opacity of deep generative models and that post-hoc forensic tools treat symptoms rather than causes. They recommend embedding provenance tracking into the generative pipeline, though this requires access to the model's training corpus, which is rarely available externally.

### 2.6 Audio sample platforms & archives

Commercial sample platforms such as Splice optimise for speed and volume without friction, offering straightforward access to large libraries and algorithmic recommendations. Although they provide broad access to audio samples and support rapid workflows, they are driven by different principles, representing distinct approaches.

Institutional archives such as the British Library Sound Archive [5] are driven by goals of preservation and research rather than musical practice. Thus, they are understandably not designed for creative reuse. Access and contribution are handled manually through forms and email correspondence, which is consistent with their institutional remit, meaning engagement with the material may not always be immediate.

Freesound [24] uses a community-driven model with Creative Commons licensing and optional AI features such as similarity search and tagging. The platform emphasises individual samples with optional metadata, resulting in inconsistent documentation. While Freesound tracks those who download samples, it does not offer an option to record subsequent transformations. Their recent blog post on generative AI acknowledges concerns about "AI sound flooding" and loss of context, proposing to flag AI-generated content [23]. Freesound demonstrates that community-driven sharing can coexist productively with AI research, but takes a different position on required metadata, friction, and the role of speed in ethical practice.

Radio Aporee [43] is an initiative dedicated to the artistic investigation of sound, place, and spatial conditions, particularly recognised for its map feature, where users can upload and access location-based samples. The archive contains field and location recordings spanning urban, rural, and natural environments, as well as public and private spaces. Contributions are open to everyone, with multiple modes for exploring sounds.

## 3 Conversations with artists

We set out to develop a sample-sharing archive that enables musicians to share their work while adhering to slow and responsible principles. The process followed a participatory design and development cycle, with users taking part in exploratory interviews that provided intuition for the platform's initial design. Participants were selected from the author's existing networks. The conversations involved two rounds of semi-structured interviews, each with two practising music producers. In each round, one participant had experience with generative AI through abstracted interfaces, while the other did not.

### 3.1 Round I: Attitudes toward AI and the ethics of sampling

The first round followed a structured format, beginning with an introduction to the research topic. Initial questions focused on participants' attitudes toward AI, followed by open-ended questions addressing sampling ethics, cultural context, and perspectives on generative audio tools.

Participants were asked about their current practices: whether they cared about the source of their samples, if they had felt uneasy using a particular sound, and what ethical sampling meant to them. Both shared frustration about losing context and concerns about careless appropriation. P2, talking about the Amen break, said, "just calling it the Amen break isn't enough really," and, speaking more generally about AI tools, said, "this is kind of a problem with producing music in general in that anyone can appropriate sounds [...] from someone else's culture without having the context of why the sound originated in the first place."

These concerns mirror those found in responsible AI music research, where Western styles dominate training data and become overrepresented. As P1 explained, "If you don't train a generative music system with music from certain parts of the world, then those kinds of sounds won't be represented in the model," and added, "Every culture would have different things that they want you to capture.", highlighting the need for flexible metadata systems to support community-based documentation.

This concern extended to a preference for specialised over generalised tools. P2 characterised larger tools as ones that encourage people with the wrong intentions to "make jungle songs for Spotify that are every break mashed together." The ethical and quality arguments were seen as aligned: P1 noted that "you are more likely to get good results ethically and accurately if you use smaller models rather than trying to do it all in a big model." The importance of understanding tools also came up, with P2 noting that "making sure that people understand what is happening [...] with the things they are inputting and making is definitely the start of it." Both expressed that ethical tools would be valued because of their constraints: as P1 observed, "producers would feel good, like using a tool or a plugin that had integrity."

### 3.2 Round II: Sample library mapping and metadata brainstorming

The second round centred on a sample library exercise where participants mapped and categorised sounds from their own libraries, organising them by origin, source type, geographic context, adn other dimensions they found meaningful,

followed by metadata schema brainstorming. Participants were asked to consider what metadata would allow them to identify the origins of a sound ten years on, and what accompanying data would be required for generative audio systems to operate transparently. P1 argued that surface-level attribution was insufficient: "the origin of the instrument or the technique or the object that they used to create that sound [...] that in itself would have its own sort of author, potentially, or creator or inventor". They also noted that metadata should account for sounds displaced from their original contexts, arguing that displacement is not always ethically neutral, pointing to colonial appropriation as a potential factor.

## 4 Slow Samples: Towards a slow and responsible audio sample archive

This project presents **Slow Samples**, a web-based platform for the responsible curation, documentation and processing of audio samples. The culture of collecting audio samples through recording, extraction, exchange, or purchase is longstanding and diverse, representing a largely informal yet deeply embedded set of practices around the accumulation of context. Personal sample libraries represent assemblages over time, often embodying contextually rich and curated audio. The platform seeks to extend this practice with ethics and intentionality, offering a place to document collections as a creative resource, in the shape of a highly contextual community archive.

The platform is motivated by a set of questions it asks the user to hold: who created this sample? Where does it come from? Why does it matter? These questions form the foundations of this work. The sample collection is the primary unit of the platform: users assemble, document and share samples as collections. Users can upload and process collections and samples, using both conventional audio operations and AI-based tools. A minimum threshold of contextual information is required for every collection, which captures the origin of the sound, its contributors, and its significance. The decision to focus on collections is intentional, as they encourage a relational way of working, drawing attention to how sounds sit alongside one another in terms of theme, timbre, or palette. This also aligns with calls for small, opt-in datasets in responsible AI music research [9], where such collections could serve as the basis for curated, opt-in datasets in the future.

Collections requiring context also encourage curation. This curation is labour: the platform makes this labour visible and celebrates its value through the use of microboundaries. Usage and provenance are documented at the collection level, seen in Figure 1, creating a relational narrative that traces how sounds pass through different hands and contexts over time.

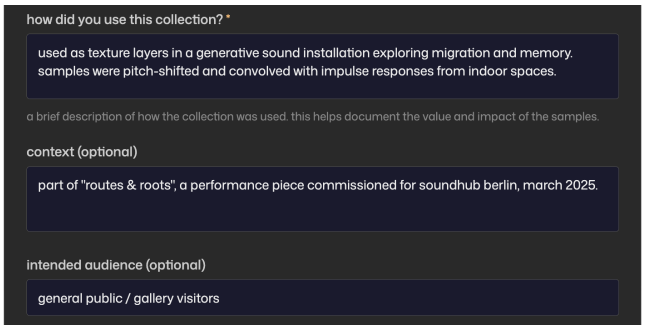


Fig. 1. Excerpt from usage form

### 4.1 Design by refusal

We are inspired by Zong and Matias [59] framework of "data refusal from below", which centres the perspective of those affected by data collectors, rather than from the standpoint of people who influence them, such as policymakers and designers; "writing from the standpoint of people who refuse, rather than the institutions that seek their compliance" [59]. They characterise refusal across four facets: *autonomy* distinguishes between individual consent mechanisms (e.g. personal opt-outs) and collective interests; *time* examines whether refusal is reactive to past harms, or proactively prevents future ones; *power* assesses whether refusal merely allows for the choice between options defined by data collectors, or creates new possibilities to systemic change; *cost* addresses whether refusal accepts individual burdens, or if it able to redistribute efforts through collective infrastructure and organisation. Thus, these refusals are an intentional stance; they are design statements. Each refusal enables an alternative practice, pointing to a feature in the system. A selection of these features and how they have been mapped to refusals are described in Table 1.

| Refusals | Derived Platform Features |
|---|---|
| *A refusal to be instant; a refusal to be frictionless* | Timed microboundary before download, processing times displayed explicitly |
| *A refusal to strip away context; a refusal to obscure origins and processes* | Required metadata fields on upload, provenance tracking, visual transformation history, plain-language explanations for all AI tools, conventional and AI processing presented separately |
| *A refusal to extract without consent; a refusal to to use without permission* | Consent status field, restrictions and intended use field, opt-in for AI training data use |
| *A refusal to facilitate passive engagement; a refusal to discourage reflection* | Collections as primary unit encouraging curation, usage documentation prompted after download, opt-in rather than default AI use |

Table 1. Mapping of design refusals to platform features

## 5 Implementing refusals

The implementation of refusals was shaped by a set of guiding concerns: how to encourage ethical engagement with highly contextual audio and the processes that shape it; how to document and present said transformations over time; how principles from slow consumption might inform more responsible approaches to music archives that directly interact with AI; how to centre the needs of practitioners and casual users, rather than only experts; whether adding intentional friction could encourage thoughtful engagement without feeling punitive; and how to best build opt-in/opt-out mechanisms from the start.

### 5.1 Metadata: Required context

Contributors must complete a set of required metadata fields on upload to ensure detailed contextual information. They are encouraged to upload sounds they have created themselves, though sounds derived through other means are permitted, provided all required fields can be completed, and consent has been obtained. The final schema was derived from several sources: interviews conducted during the early stages of this research, metadata standards in archival practice, and frameworks in responsible AI music research. Together, they inform a consent architecture that operates at several levels: documenting the provenance of a sound during its creation and the terms under which it might be used in the future.

| Field | Description |
|---|---|
| **Date of recording** | When the sound was captured |
| **Creator(s) and contributor(s)** | Name or pseudonym of those who contributed to the creation of the sample |
| **Location of recording** | Geographical location, or digital environment for synthesised sounds |
| **Source description** | What is being sampled |
| Recording method | How it was captured: e.g. field recording, synthesis, live performance, or extraction |
| Additional context | Significance, background, or meaning of the sounds |
| Consent status | Confirmation that necessary permissions were obtained |
| Opt-in for training data | Whether the contributor consents to future use of the sample as AI training data |
| Restrictions and intended use (optional) | Any further limitations on how the collection can be used |
| Tags (optional) | Descriptive keywords |

Table 2. Required metadata fields for uploaded samples

The core fields, noted in bold, reflect established archival practice derived from the Dublin Core Metadata Element Set [17], with recording method options drawn from Freesound's Broad Sound Taxonomy [25]. Consent and restrictions field go beyond these standards, and include an explicit opt-in for AI training data use. This takes inspiration from Spawning and its associated tools [32], where much of the concern is whether consent has been obtained fairly. Collections with a non-permissive consent status or training opt-out cannot be processed, as shown in **??** and **??**. These are not conceived merely as administrative measures, but as foundational conditions for future use.

*5.1.1 Trust in user inputs.* The metadata schema is built on the idea that contributors will act in good faith. While certain fields are constrained through categorical inputs, freeform text fields are largely open. The platform does not validate the quality of contributions. This is a known limitation that would become more important as the platform grows. Consent, once given, is not entirely fixed, but comes with a caveat: contributors can update the intended usage after upload, though this does not revoke local downloads already made. In this instance, a notification is shown on the dashboard to let the downloader know that the permissions have since changed, and the onus is on them to delete it locally. This raises a point of contention, and an open question: what if the material has already been used in a published project?

Further improvements, such as community moderation, have been considered, but are not part of the current prototype, which focuses on showing the value of documentation rather than controlling how it is done. Trusting contributors is a deliberate choice: the platform is intended for a community that shares common values, designed to support responsible behaviour rather than assume bad faith.

## 5.2 Digest-before-access

Before downloading samples, users encounter intentional pauses requiring engagement with context in the form of timers that cannot be skipped. The time required has been chosen with the intention of redirecting attention without feeling punitive, as a behavioural interruption rather than a comprehension tool, though this balance remains to be validated through future user testing. Processing times are displayed explicitly rather than hidden. This temporal threshold, as shown in Figure 2, is one method of implementing a microboundary, to encourage thoughtful engagement with the material.


Fig. 2. Mandatory pause before downloading collection

## 5.3 Constrained tooling

The role of AI in this project is deliberately constrained; its presence serves to assist rather than dominate. Many existing AI systems in music operate as opaque "black boxes" that prioritise novelty over accountability. This project takes the opposite position: if an AI process cannot be explained in simple, clear terms, it will not be included. Well-documented, open source and low-resource AI methods are favoured, using Model Explorer [9] as a practical resource.

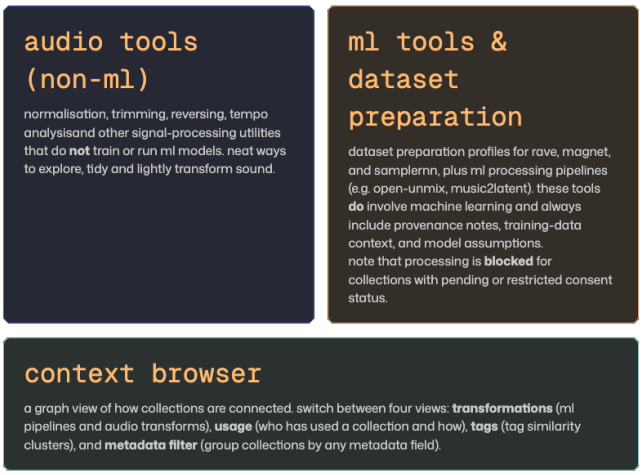


Fig. 3. Panel showing a selection of tools (audio processing, machine learning-based, context browser)

The tool menu, as shown in Figure 3, includes basic audio processing operations such as normalisation, trimming, tempo detection, and audio effects including reverb, chorus, and compression. A small set of AI-based processors are included alongside: source separation via Open-Unmix [54] and latent space audio encoding and decoding via Music2Latent [49]. Harmonic percussive source separation (HPSS) [21] is included as a non-learned source separation method. Dataset preparation pipelines for RAVE [10], MAGNet[1], and SampleRNN [36] allow collections to be exported in formats ready for training outside the platform. Every transformation is documented in the archive and associated with a specific sample collection. Each AI tool includes information of how it works, assumptions, authorship, and references to documentation and/or code. The platform has been built with extension in mind, though the goal is not to highlight the most advanced AI features, but to demonstrate how such techniques can be applied responsibly, with respect for contributors and consent.

### 5.4 Provenance tracking & transparent archives

The platform provides a way to track all processes the audio sample undergoes. Users are prompted to document their use of collections, as shown in Figure 4, and any processes/usage applied to the audio via the platform are also recorded, as shown in Figure 6 and Figure 7. Collections are represented on a separate page that visually displays these provenance changes. They can also be viewed on a map, seen in Figure 5. The aim is to represent a living archive of transparent records, which grows and evolves as collections move through different realms, capturing the relationships surrounding them.

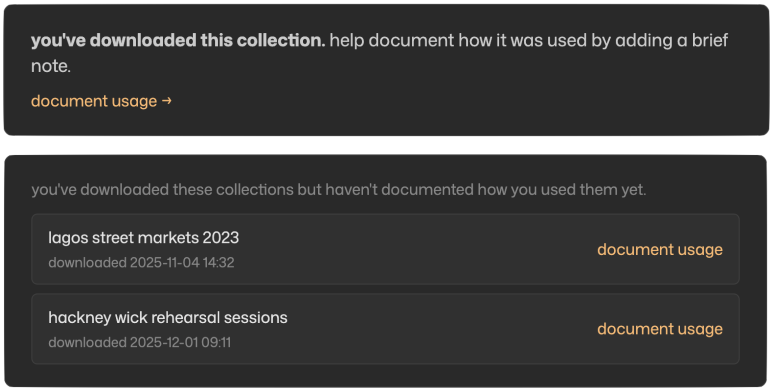


Fig. 4. Prompts reminding users to document usage of sample collections

[1]MAGNet,https://github.com/Louismac/MAGNet, Accessed 1 March, 2026.

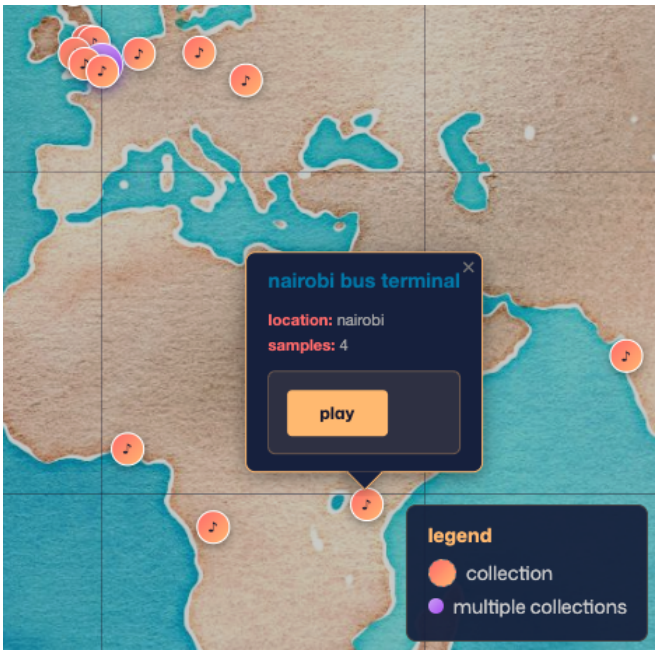


Fig. 5. Interactive map of collections


Fig. 6. Context browser: Interactive graph displaying collections and documented usage

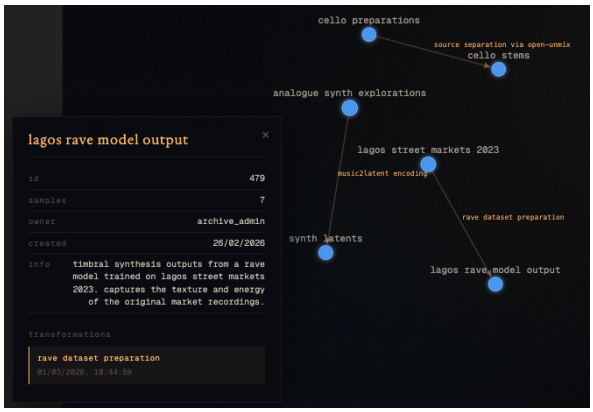


Fig. 7. Context browser: Interactive graph displaying collections and transformations

### 5.5 A desire to be open

Investigating how to publish the platform as an open-source project is a longer-term goal, and reflects a commitment to broader community focus. Open sourcing alone does not resolve issues of inclusion or substitute for community governance, but it is a starting point. The requirement for documentation means that users, regardless of background, can scrutinise available tools and make informed decisions about their use. Future plans include a "suggestions box" for users who want to propose suitable tools but lack the technical background to contribute them directly, aiming to make complexity accessible. The small number of tools currently integrated is therefore not a limitation but a look into the future; it is ongoing work, and the hope is that expansion will follow as growth around the platform continues, further discussed in Section 6.3.4.

### 5.6 Technical architecture

The platform has been built as a web application to make it accessible across multiple devices, with support for both audio upload and in-browser recording, seen in Figure 8. Building this involved a combination of traditional programming and

AI-assisted development, using the Cursor IDE [2] to generate scaffolding code. The tension this creates in relation to the platform's critique around consent is addressed in section 6.

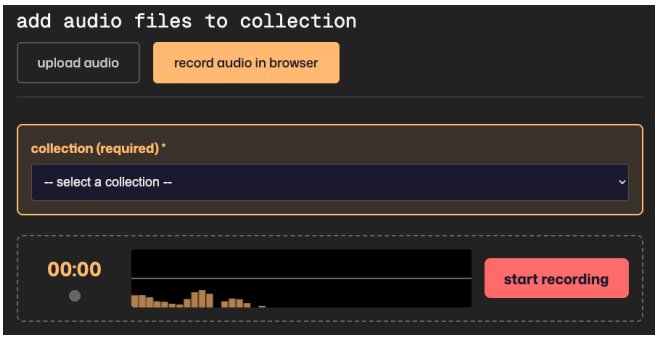


Fig. 8. Interface for recording audio in browser

The backend was developed using Flask [3], a lightweight Python framework, due to its compatibility with Python-based audio and machine learning libraries. Audio processing uses librosa [4] for analysis and feature extraction, while pydub [5] is employed for basic manipulations such as trimming and normalisation. Celery [6] is a Python task manager used for asynchronous processing, and makes slowness visible within the platform. Audio processing and ML inference do not happen instantly. Trimming silence might take seconds, but running audio through a music2latent process could take considerably longer. Tasks are queued as separate processes that are executed asynchronously. This serves two purposes: making processing times visible (giving a rough idea of computational cost) and keeping the interface usable whilst tasks run. Celery uses Redis [7] as a queue, allowing multiple users to run processing tasks simultaneously without burdening the web server. Rather than locking users out while they wait, it makes waiting visible and manageable.

## 6 Discussion

Slow AI remains largely absent from academic research, which is perhaps unsurprising given AI's typical associations. The relationship between slowness and AI is therefore unconventional, and the emergence of slow AI outside academia is often situated within a broader political context, frequently informed by feminist and decolonial perspectives [6, 39]. It is a form of resistance to the dominant strokes of speed, scale and extraction that characterise mainstream AI development. Foundational work by Haraway [29] argues that knowledge is always situated, and that context is integral to meaning rather than supplementary to it. Decolonial perspectives extend this by arguing that archival and metadata systems reproduce colonial relationships by extracting value from marginalised communities without reciprocity [27, 39, 47], a concern that participants also raised during interviews. Together, they point to the same argument: what is treated as universal is always shaped by the interests of those with authority to define the norm. Thus, frictionless design, large generalised models, and decontextualised samples represent systemic issues broader than the platforms that encompass them. By rooting design in refusals, we embrace these ideas of resistance; build to support plural cultural interpretations; and resist the flattening of knowledge.

Our refusals represent a redistribution of costs from individuals to infrastructure. Rather than requiring individuals to ethically navigate systems with no support, the platform intends to bear some of the cost by surfacing context. This does not eliminate individual burden; users must still engage thoughtfully with the information provided. However, it attempts to shift the burden from resisting within systems designed for extraction, to engaging with information designed for understanding. Whether these refusals translate into intended behaviours in real use remains an open question, and will be addressed through future user studies, described in subsection 6.3. Refusing frictionlessness challenges the values that underpin many commercial platforms; refusing extraction without consent is our way to resist the treatment of creative labour as something to be freely appropriated; refusing opacity is a way to challenge the power in systems that cannot be scrutinised; refusing scale resists the erasure of context that comes with abundance. Systemic change requires action at

[2] *Cursor IDE*, https://cursor.com/, accessed March 1, 2026.
[3] *Flask*, https://flask.palletsprojects.com/, accessed March 3, 2026.
[4] *librosa*, https://librosa.org/doc/latest/index.html, accessed March 3, 2026.
[5] *pydub*, https://www.pydub.com/, accessed March 1, 2026.
[6] *Celery*, https://docs.celeryq.dev/en/stable/, accessed March 1, 2026.
[7] *Redis*,https://redis.io/, accessed March 1, 2026.

multiple levels. In this work, we illustrate that tensions surrounding music, AI, and consent can be addressed through design in ways accessible to those outside academic research.

### 6.1 Exempt from nothing

Finally, the use of an LLM-assisted IDE for scaffolding code raises questions regarding consent. We acknowledge this tension without fully resolving it. The use of the tool is disclosed explicitly, and all generated code has been reviewed carefully. However, this does not address the underlying question of whether the models used consented data. Therefore, we highlight this as a compromise that the platform's own values would otherwise resist.

### 6.2 It is a privilege to be slow

Slowness, although well-intentioned, is often enabled by people's specific socioeconomic conditions, which Atanasova [3] frames as a form of temporal privilege, one that is not universally afforded. They argue that slowness is not a neutral quality; it is unevenly distributed across society. The slow movement has been critiqued in a similar way: it tends to be expensive to engage with, and sometimes culturally coded in ways that align with middle-class Western values. Engaging with slowness typically requires money, time, and the absence of structural pressures that enforce haste. Where the platform thoughtfully requires time, it implicitly assumes that users have time to spare. Hence, those without the luxury of time may experience intentional friction as a barrier, rather than an invitation to reflect.

### 6.3 Future work

The limitations of this work are shaped by its scale and scope. Initial interviews were conducted with a small group as an exploratory step, intended to inform the early stages of the design process. Examples of tooling have been included in the current version, but there is a scope to enhance this aspect through future contributions, which in turn can help build networks that support responsible practice. Finally, sharing these findings beyond academia remains a priority.

*6.3.1 Interviews.* Through semi-structured interviews, we plan to examine different perspectives on metadata and data provenance, aiming to balance thoroughness with flexibility across contexts. User testing with practising artists will help assess whether the design frictions are well-balanced across different users, use cases and workflows, and to better understand how the platform can fit within creative and archival practice. We will include a broad range of participants, extending beyond music and AI researchers to practising musicians, archivists, and individuals representing diverse cultural and musical backgrounds.

*6.3.2 Building networks.* A longer-term goal is to build networks of people who care about these issues, not only within NIME, but across a broader spectrum of communities. This includes archivists, metadata practitioners, and those working from critical perspectives on AI and technology, such as those engaged with the slow AI movement. Bringing people into conversation through special interest groups, workshops, or informal settings is a good starting point.

*6.3.3 Pocket zine.* We plan to turn the guidelines and findings from this work into a small, pocket-sized zine. Each page will feature a short statement based on the main themes of our research. The goal is to spark conversations about responsibility, its relationship to provenance, and how quickly AI is changing music. Since music communities sometimes find it hard to talk openly about AI, we hope this friendly format will make it easier. As a physical resource, the zine can reach practitioners who might not read a conference paper.

*6.3.4 Open source.* Open-sourcing the platform raises questions beyond technical considerations; decisions about which tools are included and how standards for inclusion evolve as the field changes are not straightforward. It is also crucial to consider how the communities whose material is represented can participate in these decisions. The current requirement for documentation and plain-language explanation is only an initial step. Meaningful community governance needs formal structures for input and review, which would benefit from collaboration with practitioners, archivists, and cultural representatives beyond the immediate NIME community.

*6.3.5 Data provenance.* Contributors can document usage of single samples in a free text field, but these sounds do not yet appear as nodes in the provenance graph. We would like to add provenance tracking for individual samples in the future, but more work is needed to find a way to show detailed relationships without making the visualisation too complex.

A related limitation is that the current provenance model is bound to the platform. Once the audio is downloaded, the metadata does not travel with the file. Investigating how this could be embedded at the file level, or bundled as a file attached to a collection, is also a consideration for future work.

## 7 Conclusion

Slow Samples is both a platform, an argument, a tool, and an artefact. It demonstrates that slowness can be incorporated into infrastructure as a structural condition, not only as an individual ethical choice. Via research through design, we suggest that ethical data practices can be integrated from the outset. By centring on slowness, we ground the platform in values that extend beyond academia, drawing attention to what is often absent from commercial tools, and do so in terms that practitioners can engage with directly. These are not merely technical problems requiring technical solutions, but human ones; demanding deliberation and ongoing engagement. Thus, the platform does not solve all the structural problems it identifies, but it makes some of them tangible.

We hope that this work is the beginning of a conversation, rather than the end of one. We urge you not to build blindly for speed. Instead, question the true cost of efficiency, and who pays. Consider consent as part of your architecture from the start; identify the sacrifices required to achieve it. Think about who your tool is for; then think again, think further. If it draws on other cultures, can you provide anything in return? Access to abundance obscures the fact that every act of consumption has a cost, and is a reminder worth keeping. Question your processes, make them visible, and if they cannot be explained, think again about who you are building for, and ask whether they should be included. Generative music tools can only be as ethical as the data they are built on, and responsibility means keeping track of context rather than obscuring it. Refusals are a form of resistance, and resistance is inherently political. Small acts, although in isolation can feel tiny, can accumulate into a broader infrastructure. With this in mind, we encourage you to contribute to that infrastructure, whether through your own practice or by helping expand what is proposed here. Slow tools may not "win", but that isn't the point.

## 8 Ethical standards

Semi-structured interviews with practising artists were undertaken during the initial design phase of this study, with both verbal and written informed consent obtained from all participants.

## References


[1] Andrea Agostinelli, Timo I Denk, Zalán Borsos, Jesse Engel, Mauro Verzetti, Antoine Caillon, Qingqing Huang, Aren Jansen, Adam Roberts, Marco Tagliasacchi, et al. 2023. Musiclm: Generating music from text. *arXiv preprint arXiv:2301.11325* (2023).

[2] AIxDESIGN. 2025. *Slow AI.* AIxDESIGN. https://aixdesign.co/projects/slow-ai

[3] Aleksandrina Atanasova. 2021. *Digital Nomadism as a Critique of Modern Life: The Role of Consumption.* Ph. D. Dissertation. Royal Holloway, University of London.

[4] Sebastiano Bagnara and Simone Pozzi. 2012. Design for reflection. *Work* 41, S1 (2012), 1108–1113.

[5] British Library. n.d.. *Sound and Vision: Introduction.* https://bl.libguides.com/reference-services/sound-and-vision/introduction

[6] Jude Browne, Stephen Cave, Eleanor Drage, and Kerry McInerney. 2023. *Feminist AI: Critical perspectives on algorithms, data, and intelligent machines.* Oxford University Press.

[7] Nick Bryan-Kinns, Berker Banar, Corey Ford, Courtney N Reed, Yixiao Zhang, and Jack Armitage. 2024. Explainable AI and music. In *Artificial intelligence for art creation and understanding*. CRC Press, 1–29.

[8] Nick Bryan-Kinns, Olga Sutskova, Rebecca Fiebrink, Phoenix Perry, Elizabeth Wilson, Anna Wszeborowska, et al. 2025. Bringing people into ai. (2025).

[9] Nick Bryan-Kinns, Anna Wszeborowska, Olga Sutskova, Elizabeth Wilson, Phoenix Perry, Rebecca Fiebrink, Gabriel Vigliensoni, Rikard Lindell, Andrei Coronel, and Nuno N Correia. 2025. Leveraging small datasets for ethical and responsible AI music making. In *Proceedings of the 20th International Audio Mostly Conference*. 70–81.

[10] Antoine Caillon and Philippe Esling. 2021. RAVE: A variational autoencoder for fast and high-quality neural audio synthesis. *arXiv preprint arXiv:2111.05011* (2021).

[11] Seyhan Canyakan et al. 2024. The role of AI in creative processes: ethical and legal perspectives in the music industry. *Journal of Music Theory and Transcultural Music Studies* (2024), 143–158.

[12] Filippo Carnovalini and Antonio Rodà. 2020. Computational creativity and music generation systems: An introduction to the state of the art. *Frontiers in Artificial Intelligence* 3 (2020), 14.

[13] Janet Chrzan. 2004. Slow food: What, why, and to where? *Food, Culture & Society* 7, 2 (2004), 117–132.

[14] Martin Clancy. [n. d.]. The Artist: Reflection 2.0 (2024): Interview with Holly Herndon:'Here Comes Everybody'. In *Artificial Intelligence and Music Ecosystem*. Focal Press, 57–72.

[15] Jade Copet, Felix Kreuk, Itai Gat, Tal Remez, David Kant, Gabriel Synnaeve, Yossi Adi, and Alexandre Défossez. 2023. Simple and controllable music generation. *Advances in neural information processing systems* 36 (2023), 47704–47720.

[16] Anna L Cox, Sandy JJ Gould, Marta E Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design frictions for mindful interactions: The case for microboundaries. In *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*. 1389–1397.

[17] DCMI Usage Board. 2012. *Dublin Core™ Metadata Element Set, Version 1.1: Reference Description.* Dublin Core Metadata Initiative. https://www.dublincore.org/specifications/dublin-core/dces/ DCMI Recommendation.

[18] Emmanuel Deruty, Maarten Grachten, Stefan Lattner, Javier Nistal, and Cyran Aouameur. 2022. On the development and practice of AI technology for contemporary popular music production. *Transactions of the International Society for Music Information Retrieval* 5, 1 (2022).

[19] Chris Donahue, Julian McAuley, and Miller Puckette. 2018. Adversarial audio synthesis. *arXiv preprint arXiv:1802.04208* (2018).

[20] Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang. 2018. Musegan: Multi-track sequential generative adversarial networks for symbolic music generation and accompaniment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.

[21] Jonathan Driedger, Meinard Müller, and Sascha Disch. 2014. Extending Harmonic-Percussive Separation of Audio Signals.. In *Ismir*. 611–616.

[22] Kate Fletcher. 2010. Slow fashion: An invitation for systems change. *Fashion practice* 2, 2 (2010), 259–265.

[23] Eduardo Fonseca, Jordi Pons Puig, Xavier Favory, Frederic Font Corbera, Dmitry Bogdanov, Andres Ferraro, Sergio Oramas, Alastair Porter, and Xavier Serra. 2017. Freesound datasets: a platform for the creation of open audio datasets. In *Hu X, Cunningham SJ, Turnbull D, Duan Z, editors. Proceedings of the 18th ISMIR Conference; 2017 oct 23-27; Suzhou, China.[Canada]: International Society for Music Information Retrieval; 2017. p. 486-93.* International Society for Music Information Retrieval (ISMIR).

[24] Frederic Font, Gerard Roma, and Xavier Serra. 2013. Freesound technical demo. In *Proceedings of the 21st ACM international conference on Multimedia*. 411–412.

[25] Freesound.org. 2026. *The Broad Sound Taxonomy*. https://freesound.org/help/faq/#the-broad-sound-taxonomy

[26] Jeff Galak, Justin Kruger, and George Loewenstein. 2013. Slow down! Insensitivity to rate of consumption leads to avoidable satiation. *Journal of Consumer Research* 39, 5 (2013), 993–1009.

[27] Jamila J Ghaddar and Michelle Caswell. 2019. "To go beyond": towards a decolonial archival praxis. *Archival Science* 19, 2 (2019), 71–85.

[28] Lars Hallnäs and Johan Redström. 2001. Slow technology–designing for reflection. *Personal and ubiquitous computing* 5, 3 (2001), 201–212.

[29] Donna Haraway. 2013. Situated knowledges: The science question in feminism and the privilege of partial perspective 1. In *Women, science, and technology*. Routledge, 455–472.

[30] Nate Harrison. 2014. Reflections on the Amen Break: A Continued History, an Unsettled Ethics. In *The Routledge Companion to Remix Studies*. Routledge, 444–452.

[31] Jonathan Herington, Raffaella Borasi, Benjamin J Guerrero, David E Miller, Blaire Koerner, Yu Jung Han, Zenon Borys, and Rachel Roberts. 2025. Musicians' ethical concerns about AI: an interview study. *AI & SOCIETY* (2025), 1–14.

[32] Holly Herndon and Mat Dryhurst. [n. d.]. *Spawning*. https://spawning.ai/ Accessed on 2025-20-02.

[33] Diederik P Kingma and Max Welling. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).

[34] Haohe Liu, Yi Yuan, Xubo Liu, Xinhao Mei, Qiuqiang Kong, Qiao Tian, Yuping Wang, Wenwu Wang, Yuxuan Wang, and Mark D Plumbley. 2024. Audioldm 2: Learning holistic audio generation with self-supervised pretraining. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 32 (2024), 2871–2883.

[35] Heike Mayer and Paul L Knox. 2006. Slow cities: Sustainable places in a fast world. *Journal of urban affairs* 28, 4 (2006), 321–334.

[36] Soroush Mehri, Kundan Kumar, Ishaan Gulrajani, Rithesh Kumar, Shubham Jain, Jose Sotelo, Aaron Courville, and Yoshua Bengio. 2016. SampleRNN: An unconditional end-to-end neural audio generation model. *arXiv preprint arXiv:1612.07837* (2016).

[37] Thomas Mejtoft, Sarah Hale, and Ulrik Söderström. 2019. Design friction. In *Proceedings of the 31st European Conference on Cognitive Ergonomics*. 41–44.

[38] Rohan Mitra and Imran Zualkernan. 2025. Music generation using deep learning and generative AI: a systematic review. *IEEE Access* 13 (2025), 18079–18106.

[39] Shakir Mohamed, Marie-Therese Png, and William Isaac. 2020. Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology* 33, 4 (2020), 659–684.

[40] Jan Mycka and Jacek Mańdziuk. 2025. Artificial intelligence in music: recent trends and challenges. *Neural Computing and Applications* 37, 2 (2025), 801–839.

[41] Vilasini Nayar. 2025. The ethics of ai generated music: A case study on suno ai. *GRACE: Global Review of AI Community Ethics* 3, 1 (2025).

[42] Ashley Noel-Hirst, Charalampos Saitis, Nick Bryan-Kinns, et al. 2025. Sampling the Latent Space: Exploring the Creative Potential of Generative AI Through the Lens of Sample-Based Music Making. (2025).

[43] Udo Noll. 2013. Radio Aporee. *Berlin: Udo Noll* (2013).

[44] Adriaan Odendaal. [n. d.]. Crafting Alternative AI Narratives Through Co-Speculative Zine-Making. *Available at SSRN 5503409* ([n. d.]).

[45] William Odom, Richard Banks, Abigail Durrant, David Kirk, and James Pierce. 2012. Slow technology: critical reflection and future directions. In *Proceedings of the Designing Interactive Systems Conference*. 816–817.

[46] Sertaç Oğul. 2024. In tune with ethics: Responsible artificial intelligence and music industry. (2024).

[47] Angela Okune. 2020. Open ethnographic archiving as feminist, decolonizing practice. *Catalyst: Feminism, Theory, Technoscience* 6, 2 (2020).

[48] Pauline Oliveros. 2005. *Deep listening: A composer's sound practice*. IUniverse.

[49] Marco Pasini, Stefan Lattner, and George Fazekas. 2024. Music2latent: Consistency autoencoders for latent audio compression. *arXiv preprint arXiv:2408.06500* (2024).

[50] Nicola Privato and Jack Armitage. 2023. A context-sensitive approach to XAI in music performance. *arXiv preprint arXiv:2309.04491* (2023).

[51] Larissa Pschetz, Michelle Bastian, and Chris Speed. 2016. Temporal design: Looking at time as social coordination. (2016).

[52] Adam Roberts, Jesse Engel, Colin Raffel, Curtis Hawthorne, and Douglas Eck. 2018. A hierarchical latent vector model for learning long-term structure in music. In *International conference on machine learning*. PMLR, 4364–4373.

[53] Sridharan Sankaran. 2026. Beyond the black box: a sociotechnical framework for ethical attribution in AI-generated music. *AI and Ethics* 6, 1 (2026), 135.

[54] Fabian-Robert Stöter, Stefan Uhlich, Antoine Liutkus, and Yuki Mitsufuji. 2019. Open-unmix-a reference implementation for music source separation. *Journal of Open Source Software* 4, 41 (2019), 1667.

[55] Suno, Inc. 2026. *Suno – AI Music Generation Platform*. https://suno.com/ Accessed 2026-02-26.

[56] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).

[57] Gaston Welisch and Natalia Stanusch. 2025. Summoning Esoteric AI. In *AIxDESIGN Festival: On Slow AI*.

[58] Elizabeth Wilson, Anna Wszeborowska, Nick Bryan-Kinns, et al. 2025. A Short Review of Responsible AI Music Generation. (2025).

[59] Jonathan Zong and J Nathan Matias. 2024. Data refusal from below: A framework for understanding, evaluating, and envisioning refusal as design. *ACM Journal on Responsible Computing* 1, 1 (2024), 1–23.

[60] Harun Zulić. 2019. How AI can change/improve/influence music composition, performance and education: three case studies. *INSAM Journal of Contemporary Music, Art and Technology* 2 (2019), 100–114.