Evolving the Living Looper: Artistic Research, Online Learning and Tentacle Pendula

Victor Shepardson
Intelligent Instruments Lab
University of Iceland
Reykjavík, Iceland
victorshepardson@hi.is

Halla Steinunn Stefánsdóttir
Intelligent Instruments Lab
University of Iceland
Reykjavík, Iceland
hallasteinunn@hi.is

Thor Magnusson
Intelligent Instruments Lab
University of Iceland
Reykjavík, Iceland
thormagnusson@hi.is

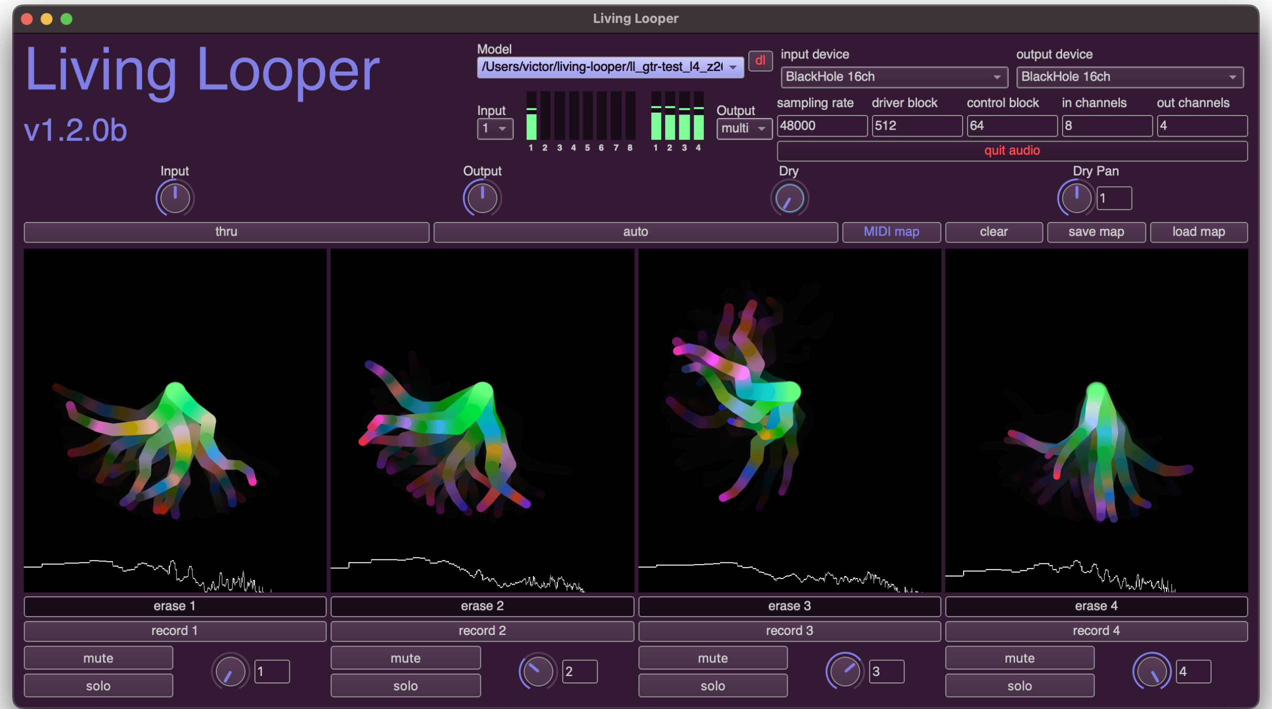

Figure 1: The Living Looper.

Abstract

The Living Looper is a neural audio synthesis looper system for live input. It combines online learning with pre-trained neural network models to resynthesize incoming audio into "living loops" that transform over time. This paper describes new features of the Living Looper and musician perspectives on its use. A new graphical interface facilitates use of the instrument by non-programmers and visualizes each loop to aid performers in tracking which loop is making which sound. We also describe a new living loop algorithm including incremental learning with partial least squares regression. Finally, we report on an artistic project using the Looper and lessons learned, resulting in an increased importance of training data and a developing sense of relationality.

NIME '25, June 24–27, 2025, Canberra, Australia

Keywords

Looper, Neural Audio Synthesis, Prediction, Agency, Intentionality, Machine Learning, Violin, Datasets

1 Introduction

The Living Looper [12] is a software NIME which borrows the interface of a multichannel looper, but re-imagines the musical loop as a predictive machine learning model. The instrument consists of multiple *living loops* which are interacted with via audio input and a bank of footswitches. While holding a footswitch, an autoregressive predictive model is fit to the sound. Once released, the learned model extrapolates the sound. An encoder-decoder model is incorporated to define the timbral space of a living looper instance, and make it tractable to fit models in real time. Each living loop can include the state of each other loop as a feature in its predictive model, allowing a network of causal relationships between loops to develop as the player improvises.

Living loops are imagined as minimal models of machinic agency, living in a tiny world of latent trajectories which correspond to sounds. As a musical instrument, the Living Looper becomes a sonic interface for thinking machine agency through performance.

This paper reports on recent developments in the Living Looper project along three dimensions. In section 2, we introduce a new graphical interface for the instrument. In section 3, we describe various enhancements to the core living loop algorithm. Section 4 is an account of an extended encounter of an artist with the instrument, and in section 5 we discuss its impact on the direction of the project. Additional media relating to the Living Looper can be found on the project page.[1]

2 User Interface

The Living Looper now features a graphical user interface. Like the core audio-only version, it is built in SuperCollider [14], with the dual goal to be modular and hackable for SuperCollider users, while also making the instrument fully usable by non-programmers.

2.1 Graphical Controls

The upper part of the graphical interface includes basic controls for the SuperCollider server, a model picker which can load TorchScript files or automatically download official models, and global input and output options. Audio output can be either mono, stereo, or one channel per living loop.

Each loop has a visualization (described in section 2.2) and conventional mute, solo and pan controls. Erasing and recording functions appear as buttons. There is 'MIDI learn' functionality which can be used to map most GUI elements to a controller without writing any code; MIDI maps can be stored and loaded as files.

2.2 Tentacle Pendula

In response to user comments in our previous study about the difficulty of keeping track of which loop was making which sound, we developed a visualization for living loops.

We wanted this visualization to supply the player with an ambient, intuitive sense of loop content, which can be subconsciously connected to the sound, without inviting detailed interpretation. We also wanted the visualization to give a unique visual identity to the instrument, but not to be so interesting to look at that it would interfere with listening to the sounds.

We take advantage of the RAVE [5] encoder features already used in living loops for visualization, which has the extra effect of giving each model a subtly different visual identity. RAVE latent dimensions are ordered by importance, and the first dimension often corresponds to loudness. We use sign normalization [10] so that polarity of the loudness dimension can be presumed. This lets us use the first dimension as an overall scale for the visualization, so that very quiet sounds shrink to nothing while loud sounds appear large.

We then take inspiration from a multi-jointed pendulum to draw the RAVE latents in order of importance. Each latent dimension becomes the angular displacement of a pendulum segment. Each segment is smaller than the last, and its final position depends on the previous segments, so the earlier dimensions have greater visual importance. Nevertheless, movement in any dimension contributes to the overall pattern of motion.

The values of each latent dimension also determine the hue of each segment, ensuring that when different vectors map to similar positions, their appearances are still distinct. Previous frames of the visualization fade out gradually so that quick changes appear visible as multiple arms.

[1] https://iil.is/research/livinglooper

We refined the multiple pendulum from a prototype with rectangular segments to the current version with tapering segments smoothly linked by circles, which gives a coral- or tentacle-like appearance. These tentacle pendula are complemented by ordinary power spectra drawn beneath, which helps to differentiate different timbres moving with similar rhythms.

2.3 Software Interoperability

Any nn~ [2] model which follows the same encoder-decoder API as RAVE can be packaged into a Living Looper instance using the Python CLI[2], which can be installed from PyPI.

Living Looper instances are now themselves nn~ models, allowing the core functions to be loaded in Pure Data and Max, and allowing us to build the graphical SuperCollider version on top of NN.ar [7]. The SuperCollider-based graphical Living Looper[3] can be installed as a SuperCollider extension.

3 Living Loop Algorithm

This section describes some enhancements to the living loop algorithm. Models are now fit incrementally to each new data point, removing restrictions on recording length (section 3.1). The ordinary least squares objective used in prior work has been replaced with partial least squares (section 3.2) which is less prone to overfitting. Improvements to the model feature (section 3.3) and a new parameterization of the encoded latents (section 3.4) are also described.

3.1 Incremental Fitting

In the living looper prototype, each loop was fit by recording into a buffer and fitting the model at the end of recording. Because of this, model fitting took place within a single processing block of audio, yet the cost to fit was proportional to the length of recording. Longer loop recordings had to be cut short or subsampled, and the total computational power available for model fitting was limited.

We replaced this 'all at once' fitting with incremental learning algorithms which distribute the computation across audio frames, processing each data point as it becomes available. This removes any limit on how long of a loop can be processed, and potentially allows more computational resources to be brought to bear on each loop.

3.2 Partial Least Squares

Living loops often need to be fit to very small datasets; consider a RAVE encoder with 16 latent dimensions, producing features at 24 Hz. A half-second loop recording might include only 12 data points, less even than the number of target dimensions. Worse, to capture temporal structure, the feature vector needs to be larger than the target vector, and it also needs to include the other loops in the looper. A 1 second time window feature in a looper with 4 loops would result in $24 * 16 * 4 = 1536$ feature dimensions. With so many more dimensions than data points, some features would likely contain noise which maps spuriously to the targets in the training dataset. The model would use these dimensions to minimize error when fit, but when extrapolating the loop, they might not reproduce any meaningful structure. That is, least squares regression is prone to overfitting.

Partial least squares regression (PLSR) [11] provides a solution to this problem. It fits a projection from the large feature

[2] https://github.com/victor-shepardson/living-looper
[3] https://github.com/victor-shepardson/living-looper-sc

space to a smaller latent space.[4] The PLSR latent space captures variance of both the feature and target data. When we choose the latent dimension to be much smaller than the feature dimension, spurious qualities of the feature are deprioritized, and overfitting is mitigated.

We implemented an online version of PLSR which works incrementally as described in section 3.1. Our algorithm is based on the CIPLS method [9] with inclusion of additional features described in earlier work [1]. It degrades gracefully to predicting the sample mean when the number of data points is very small.

3.3 Multi-Window Feature

As mentioned in section 3.2, including temporal structure in the predictive feature can make the feature dimension very large, which is computationally expensive and risks overfitting.[5]

We developed a multi-window feature, using a different window length for each latent dimension of the encoder. This has three advantages over the uniform window used previously.

The first advantage is related to the structure of variational RAVE model latent spaces. RAVE includes a principal components analysis (PCA) step which orders the latent dimensions by descending importance. If we take longer windows of the first and most important latents, we bias our living loop models toward using those more informative dimensions for prediction, while eliminating the dimensions most likely to include spurious noise.

A second advantage is that our feature dimension can be smaller while covering a longer maximum time window, reducing computational cost while enabling longer musical structures in living loops.

Finally, a third advantage is that using varying window lengths can reduce the tendency of living loop models to prefer rhythms related to the window length, since there is no single window length. This results in more varied, sometimes more organic, and perhaps more controllable loop rhythms.

We use a set of window lengths which minimizes the number of shared prime factors while distributing more feature dimensions to the more important RAVE latents.

3.4 Target Length Transform

The prototype Living Looper used an inverse squashing transform for targets, which prevents the linear predictive model from exploding toward very large numbers when used to extrapolate a loop. However, living loop models can still decay toward the target mean, which is often near to the zero vector in latent space. This tends to have a distinct sound which soon becomes overfamiliar to musicians – analogous to all the open strings ringing out on a guitar.

We found a way to mitigate this effect using a different target transform. In the forward transform, each target vector has its length concatenated as an extra element, e.g. $[a, b, c] \rightarrow [a, b, c, \sqrt{a^2 + b^2 + c^2}]$. Then in the inverse transform, the length element is reapplied: $[a, b, c, d] \rightarrow [a, b, c] \frac{d}{\sqrt{a^2+b^2+c^2}}$.

With this reparameterization, distance from the latent zero vector is preserved even when a loop 'averages out' toward the mean, which often makes for a wider diversity of sounds within a loop.

4 Extended Encounter

The second author, Halla Steinunn Stefánsdóttir, is a professional violinist and postdoctoral researcher. She encountered the Living Looper as part of a wider project aimed at exploring what creative engagement with an intelligent violin performance platform may afford a musician. This resulted in an experiment involving AI models developed from a variety of different sound archives. Stefánsdóttir's project was a collaboration with Shepardson (the first author) and a third colleague.

This section is written from Shepardson's perspective as a participant in the project, with reference to later communications with Stefánsdóttir. All quotes are from Stefánsdóttir unless otherwise cited.

Throughout the project, Stefánsdóttir played mainly her 1780s violin in a baroque setup as a live sound source for the Living Looper, via a clip-on DPA microphone.

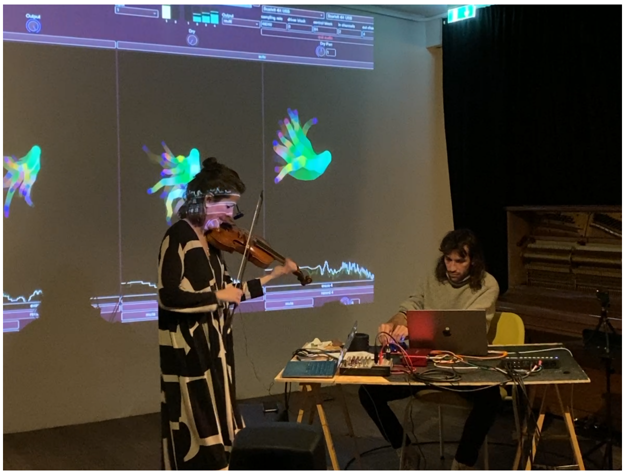

Figure 2: Stefánsdóttir and Shepardson performing with the Living Looper.

4.1 Living Looper as Confluence of Agencies

Stefánsdóttir's project emphasizes developing an ethical and affective relationship to the various archives used. She became interested in the Living Looper both as an easy point of entry into performance with generative AI, and for how its distinct instrumentality made an "invite" to go further.

As she began testing the system, Stefánsdóttir ran up against a "double agency" of the encoder-decoder model interpreting the sound, on one hand, and the living loops animating and mutating it, on the other. Stefánsdóttir also spoke of the Living Looper as "like a composition", identifying the "robust" agency of Shepardson shaping the interaction. Finally, Stefánsdóttir became increasingly concerned with the agencies embedded in datasets as they surfaced through model training.

Stefánsdóttir met these agencies in a series of performative encounters, adopting a variety of stances toward them. At one point, she spoke of lending a different agency to the looper, reducing it to a "recording machine" as she explored the rapport between violin and model. At the same time, a persistent feature in Stefánsdóttir's work with the Looper was to "hold" for

[4] note this is a different latent space from the RAVE encoder latent space, which itself is the target space for living loop models.

[5] overfitting is not necessarily bad in the living looper – but we can perhaps distinguish 'good' overfitting which reproduces a loop from 'bad' overfitting which responds in an arbitrary way to noise which isn't perceptible by the musician

moments when "something surprising and spectacular" would suddenly emerge from the living loop behaviors. At other times Stefánsdóttir acted to "negate its agency" by using only one living loop. A repertoire of techniques began to take shape, now resisting certain agencies, now letting them run wild, now changing them for effect.

RAVE and the Living Looper became practical tools for realizing Stefánsdóttir's project as her interest grew toward selecting and comparing different archives. She felt that each one "almost requires its own interface", yet developing a completely unique technical approach to each dataset would have exceeded the scope of the project. We settled on training RAVE models for Stefánsdóttir to then play through her violin in the Living Looper as a starting point for each archive. Stefánsdóttir found this setup sufficiently rich, yet approachable as a musician rather than technician, for her to develop a distinct approach to each archive through the Looper.

We worked with six different archives, each model requiring "a different negotiation" and "weighing of agencies":

Baroque violin. Dry recordings of Stefánsdóttir's violin playing from a previous project. This model was a starting point, and Stefánsdóttir found it modulated the relationship to her violin in interesting ways. According to Stefánsdóttir, the timbral model had a deadened "matte" quality, making materialities such as horsehair, gut and wood appear to her differently. Far from transparently augmenting the affordances of her violin, the model provoked her to find different ways of playing and listening. Yet at times, the compositional nature of living loops foregrounded the Looper's agency, also obstructing engagement with the phenomenological properties of the violin model.

Guitar. Shepardson's electric guitar dataset as described in [12]. This model became a point of comparison for Stefánsdóttir's violin model and a channel of communication between collaborators. Stefánsdóttir noted a particular "rapport" between the guitar model and her violin, making it an "easy entry" into performance.

Saxophone. A solo soprano sax recording previously created by improviser Franziska Schroeder. Stefánsdóttir was surprised when the model uttered watery sounds, which turned out to be traceable to techniques used by Schroeder. This was a powerful reminder of the origins of the model, spurring Stefánsdóttir to juxtapose recordings from an older archive of "minuscule detailed sounds" of the voices of her female colleagues, friends and family. With the Looper, Stefánsdóttir was "attentively performing" to make it "a recording device", developing a practice to reliably draw out certain sounds.

Glitch. A dataset created by Shepardson, Stefánsdóttir, and a colleague from minuscule, accidental, disposable, abject, or erroneous sounds which were byproducts of our various AI and recording projects. This dataset emphasized the agency of the technologies themselves; Stefánsdóttir's performance focused on improvisation and the system's capacity to surprise.

Circle of friends. A dataset mixing instrumental recordings of Shepardson, Stefánsdóttir, and several associated musicians. Performance with this model emphasized an algorithmic listening [6] through the models and datasets. Testing the models revolved around the challenge of trying to hear the ultimate source of each living loop behavior and identify hybridization of the participants. This became a way for Stefánsdóttir to honor the contributors, i.e., her colleagues. The performance was accompanied by a visual of the same performers rehearsing together to underline the altered reality produced by the system.

Great grandma. Voice recordings of Stefánsdóttir's great grandmother, Halla Lovísa Loftsdóttir, edited by Stefánsdóttir from ethnographic recordings preserved in an Icelandic archive of music and folklore.[6] Stefánsdóttir described the effect of modeling Loftsdóttir's voice as "semantic depletion", which became "grotesque" when animated by the living loop agencies. In this case, we concluded that the Looper was *not* an appropriate interface to the archive, at least not directly. Performance with this model became about increasing Loftsdóttir's agency in the performance and "making strange" [4], given the thematics of her singing about a man's encounter with the uncanny world of elves. The animation of Loftsdóttir's voice with the mindless machinic agency of living loops was in far too strange a register for this to work. Instead, Shepardson developed and performed an alternate way of playing Loftsdóttir's RAVE model permitting lighter shades of the uncanny, while Stefánsdóttir chose to use Shepardson's guitar model in the Living Looper.

Throughout, Stefánsdóttir's reliance on the graphical interface evolved. She used it to a large extent during her phase of familiarization with the looper. During later stages it sometimes faded into the background, but at other times became a dynamic part of the performance, emphasizing "the mode of search" as she tried to locate which loop had gone into a particular state.

5 Discussion

Throughout the project, Stefánsdóttir provided feedback as the technical innovations described in section 3 were implemented in new models. Though Stefánsdóttir sometimes approvingly noted differences in the behavior of new models, she also sometimes chose not to 'upgrade' an existing model in the available time because she had already developed practice around it, even if it meant tolerating failures, for example the specific "mechanical hissing" of a living loop collapsed to the zero vector before introduction of the length transform (section 3.4).

The Living Looper was conceived as a "microscope for machinic agency" [12]. To an extent this played out during the project described in section 4. However, our notion of machinic agency is less narrowly focused on living loop behaviors following Stefánsdóttir's experiences; the metaphor of a "microscope" is perhaps no longer apt. The encoder-decoder component of the Living Looper, and the datasets used to train it, were consciously de-emphasized in the original work. It was intended to focus on the loop algorithm, as a foil to other projects which would focus on the data. However, this conceit didn't survive contact with Stefánsdóttir's priorities as an artist.

In reality, the timbral model is hugely important to the Living Looper behavior, for three reasons. First, it largely determines the sonic palette, making it an irresistible dimension of control for artists, who are often more interested in radically transforming the sound of their instrument than in working with a deadened imitation of it. Second, the encoder is subject to a fundamental technical constraint: however realistic the instrument sound, it imposes some latency on the reconstruction, and so fails to approach the rhythmic affordances of a 'normal' looper. Third, the datasets involved substantially alter the meaning of a Living

[6] https://ismus.is

Looper performance, as Stefánsdóttir demonstrated with great sensitivity.

Stefánsdóttir introduced a postphenomenological vocabulary of 'intentionality' to the project, which complements the previously deployed concept of agency. If intentionality is a directedness toward the world, we can speak of the intentionality in recording a dataset, in selecting and editing it, that of Shepardson composing the Living Looper, or that of RAVE interpreting sounds and living loops animating them.[7] There can also be composite intentionalities; when Stefánsdóttir experiments with the instrument, she directs herself toward the technology's particular direction toward the world.

Ihde [8] categorizes relations to technology by how intentionality operates with the mediating artifact – *through* the artifact in an embodiment relation, *on* the artifact in an alterity relation, *interpreting* the artifact in a hermeneutic relation, or passively influenced by it in a background relation. Verbeek [13] adds classes of composite relation to the hermeneutic. In the constructive relation, intentionality operates through the mediation *into* a virtual world which it conjures. In the augmenting relation, the scope of intentionality is expanded into realms which can only be sensed via technological means, as in the example of a long-exposure photograph.

These relations provide vocabulary to discuss how our Baradian [3] agents intra-act and where we make the agential cuts between them. Stefánsdóttir found the augmenting and constructive relation particularly relevant; indeed, listening to machine listening can both augment one's perception of sound, and direct it into the virtual world of algorithms. We can identify shades of each relation within the project. A few examples are drawn from our exploratory sessions:

Embodiment. As a test, Stefánsdóttir projects an intended sound through her violin, embodying it; we sense them as one agent moving within the scope of the ongoing music.

Alterity. Shepardson erases a living loop, acting on it in an alterity relation; he acts to stop the music within the scope of the session.

Hermeneutic. Shepardson hears a living loop begin 'mechanical hissing' and knows it has entered an attractor state; he 'reads' this fact in the scope of software development.

Augmenting. In concert, Stefánsdóttir performs on the violin 'in parallel' to the violin's role in the software system; audience and performer's perception is augmented by amplification of the violin sound.

Constructive. Shepardson perceives the simulated behavior of living loops and model *through* the software in a constructive relation; living loops merge and separate in the scope of their virtual world.

Background. Our approach to Loftsdóttir's model is confused by the strictures of the Living Looper in a background relation; we struggle to place Loftsdóttir and the Living Looper in the same world.

Ultimately, we find the complex tangle of agency and intentionality in the Living Looper stimulating. Tracing the status of machine, operator, and data actants within an AI system is inevitably difficult; Stefánsdóttir found the Looper compelling precisely as a means to stage these challenges as art music. While training a RAVE model and exploring it via the Looper provides an easy starting point, unpicking our relations to each data source, data set, model, musician, instrument, composer and living loop is an ongoing project which progressively deepens our understanding of each archive.

[7] note that technological 'intentionality' does not imply that artifacts have intentions in the same way a human does.

6 Conclusion

This paper reported on development of the Living Looper project, covering design, algorithmic, and artistic aspects. We documented a musician's extended encounter with the instrument and resultant changes in direction for the project.

We see two broad directions for the project to go in. One is to embrace the influence of training datasets and imagine the Living Looper as an instrument for examining our relation to them, as we did in Stefánsdóttir's project. Complementarily, we can continue to experiment with non-data driven encoder-decoder processes, expecting to either produce a version of the project which does foreground living loop behaviors, or to again expose contradictions in the concept.

In ongoing work, we also hope to address a persistent user comment about the unsuitability of the soft foot controller used thus far, by developing a physical controller.

7 Ethical Standards

All machine learning models and data used in this research was created by participants, freely supplied by collaborators, and/or part of a public archive.

Acknowledgments

The Intelligent Instruments Lab is supported by the European Research Council (ERC) as part of the Intelligent Instruments project (INTENT), under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101001848). This research was also supported by an NVIDIA hardware grant.

References

[1] Hervé Abdi. 2010. Partial least squares regression and projection on latent structure regression (PLS Regression). *WIREs Computational Statistics* 2, 1 (2010), 97–106. https://doi.org/10.1002/wics.51 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/wics.51.

[2] ACIDS. 2023. nn~: a max/Pd external for real-time ai audio processing. https://github.com/acids-ircam/nn_tilde

[3] Karen Barad. 2007. *Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning.* duke university Press.

[4] Genevieve Bell, Mark Blythe, and Phoebe Sengers. 2005. Making by making strange: Defamiliarization and the design of domestic technologies. *ACM Trans. Comput.-Hum. Interact.* 12, 2 (June 2005), 149–173. https://doi.org/10.1145/1067860.1067862

[5] Antoine Caillon and Philippe Esling. 2021. RAVE: A variational autoencoder for fast and high-quality neural audio synthesis. *arXiv:2111.05011 [cs, eess]* (Nov 2021). http://arxiv.org/abs/2111.05011 arXiv: 2111.05011.

[6] Miguel Carvalhais and Rosemary Lee. 2019. Soundwalking and Algorithmic Listening. https://doi.org/10.14236/ewic/RESOUND19.8

[7] Gianluca Elia. 2024. NN.ar() nn_tilde adaptation for SuperCollider. https://github.com/elgiano/nn.ar

[8] Don Ihde. 1990. *Technology and the lifeworld: from garden to earth.* Indiana University Press.

[9] Artur Jordao, Maiko Lie, Victor Hugo Cunha De Melo, and William Robson Schwartz. 2021. Covariance-free Partial Least Squares: An Incremental Dimensionality Reduction Method. *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)* (Jan. 2021), 1420–1428. https://doi.org/10.1109/WACV48630.2021.00146 Conference Name: 2021 IEEE Winter Conference on Applications of Computer Vision (WACV) ISBN: 9781665404778 Place: Waikoloa, HI, USA Publisher: IEEE.

[10] Nicola Privato, Victor Shepardson, Giacomo Lepri, and Thor Magnusson. 2024. Stacco: Exploring the Embodied Perception of Latent Representations in Neural Synthesis. In *Proceedings of the International Conference on New Interfaces for Musical Expression.*

https://www.nime2024.org/program/255/Stacco__Exploring_the_Embodied_Perception_of_Latent_Representations_in_Neural_Synthesis.pdf

[11] Roman Rosipal and Nicole Krämer. 2006. Overview and Recent Advances in Partial Least Squares. In *Subspace, Latent Structure and Feature Selection*, Craig Saunders, Marko Grobelnik, Steve Gunn, and John Shawe-Taylor (Eds.). Vol. 3940. Springer Berlin Heidelberg, Berlin, Heidelberg, 34–51. https://doi.org/10.1007/11752790_2 Series Title: Lecture Notes in Computer Science.

[12] Victor Shepardson and Thor Magnusson. 2023. The Living Looper: Rethinking the Musical Loop as a Machine Action-Perception Loop. https://doi.org/10.5281/zenodo.11189164 Pages: 224–231 Publication Title: Proceedings of the International Conference on New Interfaces for Musical Expression Publisher: Zenodo.

[13] Peter-Paul Verbeek. 2008. Cyborg intentionality: Rethinking the phenomenology of human–technology relations. *Phenomenology and the Cognitive Sciences* 7, 3 (Sept. 2008), 387–395. https://doi.org/10.1007/s11097-008-9099-x

[14] Scott Wilson, Nick Collins, and David Cottle (Eds.). 2011. *The SuperCollider book.* MIT Press, Cambridge, Mass.