Re-Animating the Archive: Performing a Machine Learning System as Living Memory

Jonathan Reus
j.reus@sussex.ac.uk
University of Sussex
Brighton, UK

Abstract

This paper examines machine learning systems as a form of performable archival memory, operating within long-durational musical instruments. It introduces re-animation as a conceptual framework for understanding how ML systems trained on voice and text operate as unstable, probabilistic memory systems that continually reinterpret the past through inference and training interventions. Drawing on the year-long generative radio broadcast *In Search of Good Ancestors / Ahnen in Arbeit*, the paper grounds this framework in a concrete artistic case study. Central to the conceptual work of this paper is a reframing of instrumental liveness away from immediacy or low-latency control, and towards a temporally distributed liveness unfolding across dataset creation, iterative retraining, and durational listening. The paper further positions dataset creation as a form of musicking, where collaborative data creation workshops used in the work frame voice and text contributions as gifts rather than fungible resources, embedding reciprocal address, care, and future-oriented intent directly into the instrumental system. This orientation recognizes voice as an inherently relational act whose ethical dimensions persist beyond its digitization, and offers counter-practices to extractivist data epistemologies prevalent in mainstream AI development. By treating the training and curation of an ML model as a continuous, performative, and social practice, this paper contributes a reorientation in thinking about ML instruments, arguing that musical significance in ML-based instruments can emerge through duration, care, and collective intervention rather than virtuosic immediacy.

Keywords

Machine Learning Instruments, AI Voice Synthesis, Liveness and Performance, Dataset Creation as Musicking, Archival Re-Animation, Data Ethics

1 Introduction

Machine learning systems and statistical imitation frameworks are increasingly used in musical instruments and performance contexts. Such mimetic instruments have the express goal of producing media of plausible similarity to their training data - usually corpora of recorded sound, text, or gestural sensor readings [18, 26]. These corpora are, in many cases, fixed sources from which material can be retrieved for the purpose of pattern analysis and resynthesis [37]. Within NIME discourse, such systems are typically evaluated in terms of control, responsiveness, and immediacy, suggesting a certain idea of liveness framed around low-latency interaction and performer agency [3, 24, 36].

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

This paper approaches machine learning instruments from a different angle. Using the work *In Search of Good Ancestors / Ahnen in Arbeit* as a case study, I ask what it might mean to treat a machine learning system as a durational and probabilistic archive - one whose memory is unstable and continually reinterpreted through time. This framing offers a perspective on liveness as not primarily arising from real-time manipulation, but from the totality of processes involving collection, training, forgetting, and repeated intervention that unfold across weeks and months. In doing so, it expands NIME discourse beyond immediate, low-latency control; and presents a kind of infrastructural liveness - an unfolding of encounters and interventions with a musical system's technical assemblages.

In Search of Good Ancestors / Ahnen in Arbeit (ISoGA) was a generative radio artwork in which a bespoke machine-learning voice system broadcast continuously over public channels for a full year (2022–2023)[1]. Over the course of its run, the system evolved through a series of collaborative dataset creation workshops that reshaped its training data and internal configurations. The work's concept was to create a system whereby machine learning becomes a form of living, re-animated archival memory, which allows textual and vocal materials contributed through the workshops to perpetuate, transform and degrade over time.

As much as it directly deals with the instrumental dynamics of machine learning, *ISoGA* was conceived primarily as a new form of oral history archive - a memory of past human lives transmitted through voice. Scholarship in archival theory therefore becomes invaluable for understanding this work, particularly the concept of *re-animating data*, as developed by Rachel Thomson [41–43]. Thomson describes her approach working with oral history archives as, "a method for working with the long past materials of social research to forge a method for engaging and co-producing knowledge with new audiences in the present". When such an approach is used to frame machine learning systems, it foregrounds interesting epistemological possibilities for understanding ML as a vehicle for reinterpreting accumulated traces of past material. As archive studies tells us, such reinterpretation is inherently subjective and context dependent, thus transmission is a fragile and fraught affair, and the unique mediations of a machine learning system may add instability.

Machine learning systems are not perfect memories, after all. They do not reproduce archival sources, but reconstruct probable patterns from them. They do also readily forget - the ML term "catastrophic forgetting" attests to this. However, oral history archives are also not perfect reproductions, nor is human memory itself. Interpretive instability thus may become a primary aesthetic and performative resource in the context of an archival instrumental performance system. This paper aims to present perspectives on machine learning instruments that may help to enrich NIME discussions in this area, touching on liveness and

[1] ISoGA was commissioned by CTM Festival, Deutchlandfunk Kultur and Austrian Ö1 radio.

instrumentality, memory and duration, and offers a case study of an instrument that is a socially porous, collective musical artefact whose instrumental dynamics emerge through care, iteration, community and contingency.

2 Re-Animation as a Concept for Machine Learning Instruments

2.1 ML as Archive

In conventional archival frameworks, sound recordings are stored, indexed, and annotated for later retrieval. That is, the recordings are replayed with the goal of maintaining fidelity to an original source for a new listener, one who is often far removed from the original listening context. But while the engineering framing of these recordings as "data" might suggest that the resulting recordings are now free from the burdens of context, archival theory maintains the opposite epistemological stance. As Moore and Thomson observe from their re-animating data work, "it is not always necessary to have been 'there' to care" [30]. Archive studies differs fundamentally from the beliefs of data science in that it does not understand the act of encountering "data" as pure retrieval. The act of re-listening to archival recordings is never passive recovery; it is already an act of re-animation, a transformation of the past which requires care.

When I ask that we consider vocal machine learning systems as a kind of oral history archive, this is not only an epistemological metaphor, but an ontological assertion. Machine learning systems plus their datasets are categorically more like archives than not. Jacques Derrida's *Archive Fever* is crucial here. Derrida argues that the archive is not merely a record of the past, a resource to be accessed, but a "commandment" that shapes the future [15]. This orientation is, for Derrida, not a question of preserving the past so much as "the question of a response, of a promise, and of a responsibility for tomorrow". Archives are not a neutral store of records, but a system organised around *arkhē* - the etymological root of "archive". *Arkhē* means both commencement and commandment, origin and command. It is both where things begin and where authority is exercised. The archive always involves a double gesture: selection (what is preserved) and exclusion (what is left out, what decays). An ML system shares exactly this structure - it is constituted by what was in the training corpus and how its statistical features are modeled (its arkhē). Everything outside that corpus and statistical capacity to reproduce is, in a real sense, unarchivable by the system. The model does not simply reflect and reproduce its training data; it is instituted by it, in the same way an archive is instituted by the decisions of the archivist and the power relations that funded the collection. The mimetic ability of such a system echoes Derrida's commencement in that it samples new material from a generative statistical model, and commandment in that it is governed both by laws (analysis and sampling algorithms) and exercises power (the power to include and exclude what is heard and propagated).

Oral history sharpens this comparison further. Oral history archives are already defined by a tension between the live, embodied vocal event and its archival trace. Oral historians have long noted that transcription produces a loss of aural information - that the subjectivity, hesitation, and breath in a recorded voice escape the documentary record [32]. A voice AI system trained on oral recordings repeats this loss at a higher level of abstraction: not only is the live voice reduced to a recording (first archival reduction), but the recording is then reduced to statistical features (second archival reduction). Each reduction is an act of appraisal - a decision about what is worth preserving - and each is irreversible.

2.2 ML as Memory

Unlike a traditional oral history archive, a trained ML model has no indexical location where a given vocal event "lives." Instead, it encodes dispositions toward certain sounds, transitions, and timbres that only become audible in the act of generation. Every output is always a generation rather than a retrieval. This reconstructive character aligns ML systems with long-standing insights from cognitive science about human memory. While the notion that human memory is a generative process goes back at least to the seminal work of Frederic Charles Bartlett [4], more recent work in computational neuroscience has gone so far as to propose that human memory consolidation itself operates through generative networks analogous to variational autoencoders, describing recall after consolidation as a "generative process" rather than the retrieval of a copy [39]. The representational space of a voice model is, in this light, a set of statistical tendencies that manifest differently each time they are called upon. Complementary to the archival framing, mimetic ML can be thought of as a mobilized memory of listening, an active process of re-animation and reconstruction. What is often problematic about the rhetoric of "generation" surrounding mimetic ML is that it invisibilizes the archival and memory-like conditions of its constitution. The concept of re-animation thus does valuable work in deflating the rhetoric of generation.

2.3 ML as Reanimation

Human memory is not only cognitive, however; it is also bodily and social. For ethnographer Paul Connerton, social memory is transmitted through "commemorative ceremonies and bodily practices" that are performed rather than retrieved [13]. In this sense, the social re-animation of memory happens not only when mediated by re-synthesis, but also in mediated bodily practices and human encounters with technical infrastructures of memory. These include, for example, acts of dataset creation, reflective listening and model fine-tuning - which all fall within the purview of memory's instrumental dynamics.

This then leads us to the question of liveness in archival practice. Liveness in the archival sense is not as much about real-time as it is about encounters and performativity. A machine learning system is live when it is performing interpretation of the archive; it is live when someone interprets and contributes to its underlying data corpus; it is live as it actively includes and excludes during training. I use the term re-animation to describe this multifaceted liveness in regard to an archival corpus, echoing Rachel Thomson's work, which emphasizes the ongoing performativity of an archival source. Thomson and McGeeney have usefully articulated re-animation as four linked performative strategies: **collaging**, the recombination of archival fragments to surface new meanings; **re-asking**, the reopening of past questions to new audiences across temporal disjunctures; **re-voicing**, the performance of another's archived words in a new register; **re-collecting**, the invitation to new participants to contribute to an archive still in the making [42].

Thomson's framework, originally developed for human engagement with oral history archives, maps usefully onto the compositional logic of a machine learning system designed around

archival vocal memory. In *ISoGA*, the system continuously collages statistical fragments from accumulated corpora; its probabilistic inference re-asks learned patterns in new configurations with each synthesis; its voice models re-voice contributed texts through accumulated timbres; and its workshop cycle is explicitly re-collecting - inviting participants to contribute to an archive whose future shape they cannot fully anticipate. Dataset creation thereby became a form of musicking under technological conditions [21, 38], embedding social relationships, intentions, and constraints into the system's evolving memory.

There is a tension here between a purist notion of live performance and the archive. Indeed, this tension has its own rich history. Live performance has been a difficult object of archival study precisely because archives are inevitably incomplete [40]. Peggy Phelan famously posited that "performance cannot be saved" without betraying its ontology [12], echoing similar sentiments from oral history scholars [32]. What re-animating data and *ISoGA* attempt to stage, however, is the act of the archive itself performing - enacting a kind of uncanny temporality. As Colby describes, "moving between archive and stage is, ultimately, an uneasy exercise in time, in making the duration of the past live again in whatever sense this can be accomplished, and as such there is a strangeness about it which manifests itself primarily in the realm of sensation" [12]. The artistic effect of *ISoGA* operates in this realm of uncanny sensation; producing a juxtaposition of re-animating acts where the past is not retrieved but brought out-of-joint through the contingencies of live synthesis. This aesthetic is in alignment with the broader "archival turn" in performance studies that treats documents as performative reiterations to be acted upon rather than fixed, immutable traces of a live event [11].

Framing machine learning instruments through re-animation inevitably shifts how liveness is understood. Rather than prioritizing immediacy, transparency, or fine-grained control, re-animation emphasizes multi-temporal relationships - allowing systems to change publicly, to remember imperfectly, and to reveal their instability over time. This opens space for machine learning instruments less concerned with real-time virtuosity and more attentive to social and interpersonal relationships mediated by dataset creation and training regimens. In the sections that follow, this concept is examined through the elements of *ISoGA*: the design of the BroadCaster system, the workshop-based interventions that reshaped it, and the listening experiences that emerged from its long-durational public performance.

3 The BroadCaster

3.1 Voice and Language Synthesis

The BroadCaster, the generative system at the core of *ISoGA*, was designed as a machine learning instrument whose musical composition unfolded over the time span of weeks and months. The goal of the system was to operate an assemblage of interacting processes that continuously reinterpreted accumulated vocal and textual material (see Figure 1). The technical components of The BroadCaster can be understood as a variable, dynamic score - shaped by social interventions - rather than a deterministic pipeline, an example of the 21st century move from "composition of works" to "invention of systems" [29].

At a high level, the system integrates four functional layers: probabilistic text generation, neural voice synthesis, long-durational parameter modulation, and audio orchestration. These processes are loosely coupled, operating asynchronously and allowing the system to tolerate technical delays, interruptions, and modular reconfiguration without collapsing. This architectural looseness was a deliberate design choice, supporting instability, drift and serendipitous compositional alignments.

Table 1: Three Magic Words

Rule	Action
__S L O W__	Read this phrase very slowly.
__SHUFFLE__	Shuffle the words while reading.
__TLDR__	Summarize in as few words as possible.

The **text generation module** produces short individual utterances that are remixed and prepared for synthesis into voice by the orchestration module. The text generation module accepts asynchronous requests with parameters controlling stochasticity of generation, continuity with prior text outputs, contextual prompts controlling the subject matter of the utterances, and control over textual polyphony (multiple parallel scripts with related context). The values of these parameters were controlled primarily by the parameter modulation module, allowing the orchestration module to treat the generated textual material as a flexible libretto for vocal performance. Each instance of text synthesis constitutes a re-performance of textual tendencies learned through model pre-training, with a single GPT-2 family text model pre-trained on the ubiquitous open LLM training dataset The Pile [20]. The initial text model was then gently fine-tuned on texts contributed by participants both in workshops and via the radio broadcast's web portal interface.

A distinctive element of the text generation system was a post-processing layer of "magic words". Originally devised by the artist collective varia, magic words are a playful method of collaborative writing, reading and revision where writers are able to introduce text-based triggers directly into shared documents[2]. These text notations, written using double underscores (e.g. **__REPEAT__** or **__DISSOLVE__**), represent computational instructions for text processing. In the workshops of *ISoGA*, participants were invited to design new magic words that could be read by other participants in live group performances, and later, interpreted by the text generation system itself, indicating how a text should be fragmented, repeated, or altered. Following the first workshop, the magic words vocabulary developed during the workshop was incorporated into the script generation system as a controllable post-processing layer. After generating initial text for a script passage, the system stochastically inserted magic words into the generated raw text, after which a secondary parsing stage reprocesses portions of text according to the vocabulary conceived during the workshop. This introduced composed poetic variations into generated scripts - enabling micro-compositions and distinct performance styles to emerge within the broadcast, and ensuring that the collective decisions about how language should behave in the system were embedded directly into the instrument's generative logic. Three example magic words are provided in Table 1.

Voice synthesis was handled by a module composed of a growing collection of neural voice models, beginning with a single

[2]The method had its origins in varia's Read & Repair series - a practice of collective reading and annotation on shared Etherpads that first took shape in a session hosted by amy pickles, and was subsequently developed through the Minimal Viable Learning edition of Read & Repair, put together by amy pickles, Manetta Berends, Luke Murphy, Joana Chicau, Cristina Cochior, and Silvio Lorusso [44].

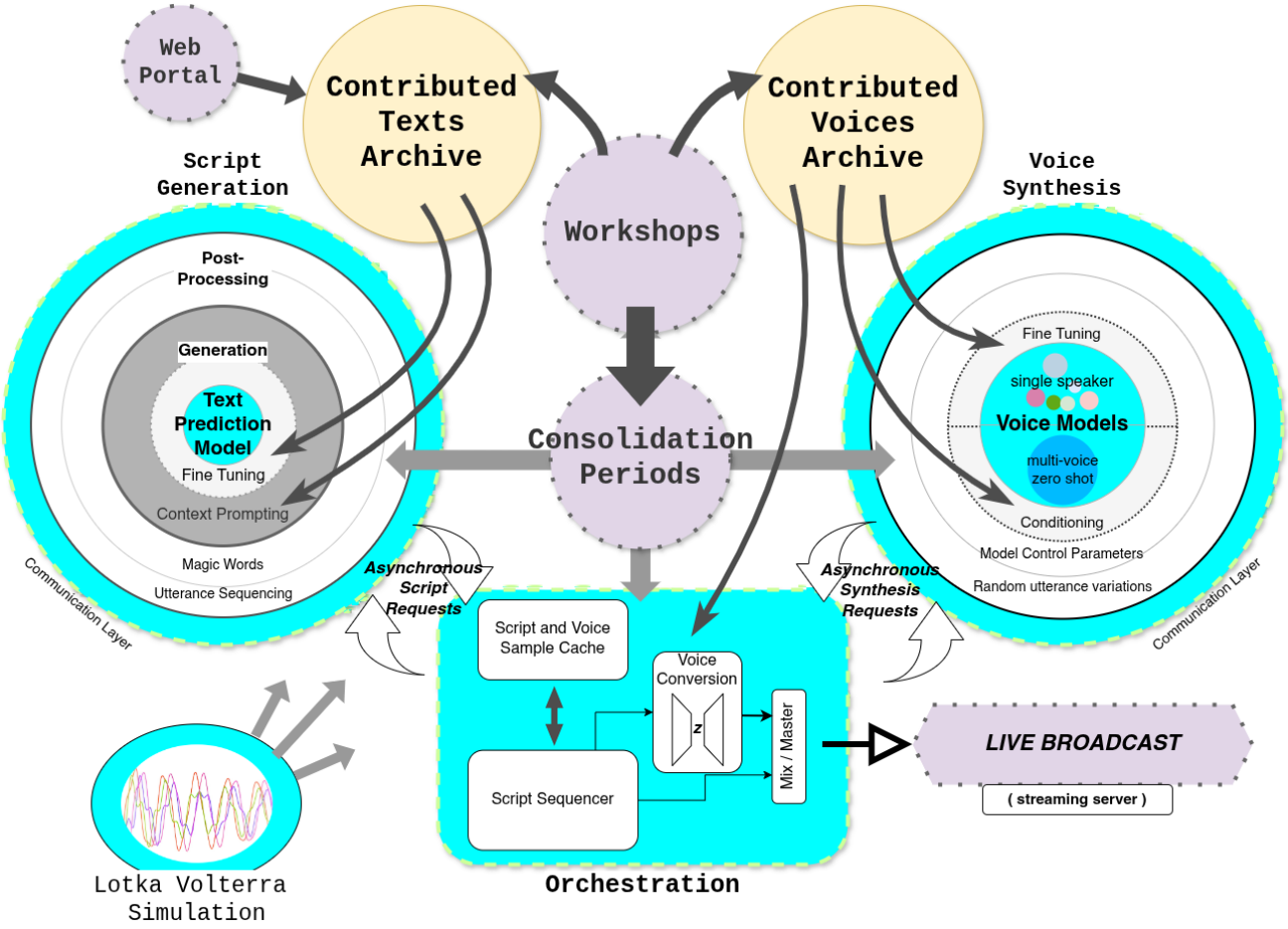

Figure 1: An illustration of The BroadCaster assemblage. The BroadCaster was conceived as a re-animation instrument rather than a linear generative pipeline. Text generation, voice synthesis, long-durational modulation, and audio composition operate asynchronously and feed back into one another over time. Past outputs remain active within the system, workshops and consolidation periods function as interventions that reconfigure the instrument's internal conditions.

Tacotron2 model trained on the voice of audiobook reader Linda Johnson (LJ Speech dataset [23]), with an additional network bending feature added to the alignment generator section of the network to allow for modest control over linguistic versus non-linguistic sound generation. Additional model architectures, such as VITS [27] and YourTTS [10] were added throughout the year, and trained and fine-tuned at different moments during and post workshops. As new voices were added to the archival corpus, the system's vocal identity became increasingly heterogeneous, producing a polyphonic field in which multiple timbral and prosodic characteristics coexisted.

Influenced largely by the shifting simulation values of the modulator, the selection, configuration and alignment warping of voices were treated as compositional parameters, allowing aesthetics of vocal identity to remain fluid, polyvocal and unstable. The creation of polyvocality was further supported by a post-processing layer of voice conversion models [9] within the orchestration module. These voice conversion models were trained on the archival vocal corpus of the work, alongside additional vocal recordings from my own archive.

The initial training conditions of the BroadCaster on two of the most widely used open voice and text training corpora was an artistic decision. Certain datasets acquire momentum and authority through repeated use, even when their limitations or ethical complexities become apparent. By deliberately using two of the most widely used datasets in voice and text, I had intended here to make visible the dependence of contemporary generative systems on generic, pre-existing datasets, while also creating a stable point of aesthetic departure from which transformation of the broadcast over the year could be understood.

3.2 Long-form Composition through Predator-Prey Simulation

The challenge of creating a non-repeating composition over the time span of a year was addressed through implementing a multi-parameter complex system simulation based on the multi-species Lotka-Volterra equations, a long-form compositional approach suggested by Eldridge in 2007 [17]. The multi-species Lotka-Volterra equations extend the classic predator-prey model to a community of N interacting species, producing aperiodic, chaotic dynamics that are highly sensitive to initial conditions and resistant to long-term prediction - a mathematical behavior that echoed *ISoGA*'s concern with the fragility of memory and the unknowability of the future. Multiple simulations were run in parallel at different timescales using the Runge-Kutta integration method, and their continuously streaming population values

were routed as modulation parameters across all four modules of The BroadCaster.

In the text generation module, the simulation governed two competing dynamics: forgetting and accretion. A persistent context cache - a rolling block of text drawn from the archival corpus and previously generated material - served as the contextual prompt for the GPT-2 model. The rate at which old context was removed from the top of this cache (forgetting) and the rate at which new material from the archive was added (accretion) were continuously modulated by the LV parameters, causing the system's textual memory to expand, contract, and shift its frame of reference over time. The same simulation also governed the density of magic words inserted during the post-processing of generated scripts: periods of high density introduced dense poetic interruption and structural variation, while lower-density periods allowed smoother, more continuous monologues to emerge.

In the voice synthesis module, the LV equations operated differently depending on the underlying architecture of each model. For duration-predictor-based models such as VITS and YourTTS, the simulation modulated the duration predictor directly, stretching or compressing the temporal unfolding of individual phonemes and producing shifts in rhythm and prosody. For soft-alignment-based models such as Tacotron2, the equations instead modulated the amount of noise injected into alignment vectors during inference, a form of network bending [7] that, at higher intensities, produced stuttering, repetition, and the dissolution of speech into abstract voice-like textures.

At the level of orchestration, the LV simulation governed two further compositional dimensions: the density of simultaneous arranged voices (ranging from dense polyphony to sparse, solitary utterance) and the type and intensity of voice conversion applied to each synthesized voice before it entered the broadcast mix. In this way, the simulation functioned as a kind of invisible ecology whose shifting dynamics left their mark on every layer of the instrument simultaneously, binding the otherwise loosely coupled modules into a single, slowly breathing system.

Another key dimension of the long-form musical structure of the work was the consolidation periods following each workshop. Four of these periods, spaced throughout the broadcast year, involved the integration of new data, model fine-tuning, and additional mediating layers such as the magic words system and the introduction of new voice synthesis architectures. Each of these consolidation periods marked a major change in what could potentially happen through the broadcast's unfolding.

3.3 Aesthetics and Listener Experience

For listeners encountering the broadcast, the most immediate aesthetic register was one of uncanny dialogues morphing in and out of an almost Dada-like unintelligibility [3]. The BroadCaster was designed to shift fluidly through speech, song, words, and soundscapes - performing words as if slowly tracing over the letters, probing and unsettling them - producing voices that were statistically plausible yet subtly displaced from ordinary human speech. At higher temperatures, the system's voice models began to deviate from intelligible utterance, producing loops resembling singing, humming, screaming, or crying - a compositional territory opened deliberately through hacking the soft alignment mechanism of Tacotron2. Over short listening sessions, the effect was of eavesdropping on something between monologue and transmission: intimate yet bodiless, occasionally collapsing into abstract phonetic texture. As workshop interventions accumulated and new voice models were integrated, listeners described encountering what one called "a strange cloud of voices, from robotic to real," finding the dense layering of disembodied voices and the varying fidelity of different models aesthetically compelling for its polyphonic effect. Many particularly noted moments when models deviated from intelligible speech into more abstract vocal expressions, pointing toward aesthetic affordances that exceed the text-to-speech paradigms the underlying models were based upon. The continuous online stream encouraged either deep relaxed listening or a more episodic "checking in" mode of temporal engagement. Critically, no single listening position or in-person intervention could apprehend the work as a whole. The composition was, by design, distributed across timescales and modes of encounter.

4 Re-Animation Through Intervention: Workshops

Re-animation in *ISoGA* did not occur solely through the statistical reconstitutions and deviations of The BroadCaster, but through an iterative process of deliberate human interventions that reshaped the instrument over time. A series of public workshops functioned as performative events in which new material was introduced and, during a follow-up consolidation period, the BroadCaster's internal conditions were reconfigured. These workshops were not auxiliary to the broadcast; rather the broadcast was the vehicle for another type of musicking in which the archive was actively encountered, re-animated and further developed (see Figure 2).

The first workshop, *Re-Writing / Re-Reading*, focused on the initial creation of the textual archive behind the work. Participants contributed texts they considered meaningful or transmissible, such as an essay by Hannah Arendt or Chief Oren Lyons' 1992 lecture to the United Nations. This collaborative archival practice foregrounded the act of choosing what might be carried forward to future generations, to be interpreted through the broadcast system. During the workshop, these texts were collectively performed using the playful notation system of Magic Words, described in section 3.1. Following the workshop, both the textual material and the notational vocabulary were incorporated into the text-generation process; altering how the system produced scripts in subsequent broadcasting.

The second workshop, *An (In)audible Chorus*, shifted attention from text to voice. Participants recorded vocal material under speculative conditions, reflecting on what aspects of their voice they might wish to transmit into the future and how others might be permitted to use it. Such wishes for future encounter, interpretation, and use were then embedded into the archival recordings in the form of audio metadata, and as the spoken material of the voice recordings themselves. The workshop additionally deployed scores for experimental forms of voice data creation and consensual use, such as scores and warmups for consensual voice swapping (see Figure 3).

The workshop, held in person at CTM Festival in Berlin, offered participants the opportunity to re-appropriate the four modules of The BroadCaster itself (made available individually via a local web server), and to experiment with voice data creation and synthesis in-situ. The result of this workshop, besides the beginning of a vocal archive for the work, was a collective

[3] Audiovisual documentation of the broadcast may be found at https://www.researchcatalogue.net/view/3866796/3866991, and more general documentation about ISoGA at https://jonathanreus.com/portfolio/in-search-of-good-ancestors/

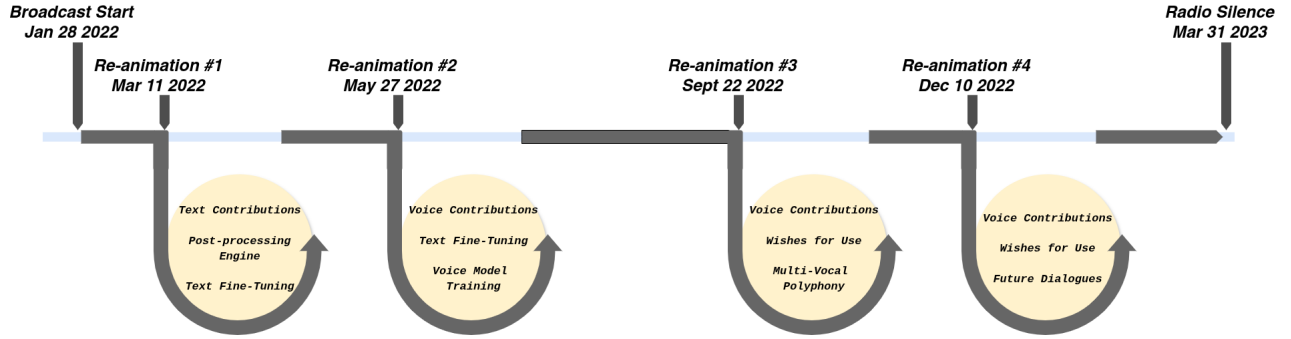

Figure 2: Temporal structure of ISoGA as a continuous year-long public broadcast punctuated by collaborative workshops. Each workshop introduces new material and constraints, followed by a system consolidation period, after which the system continues to evolve autonomously. Change becomes perceptible through drift, accumulation, and repeated listening rather than moment-to-moment interaction.

performance using a hybrid of synthetic texts, human voices and synthetic voices trained on workshop recordings. Artificial voices circulated alongside live human voices, producing dense polyphonic textures that blurred distinctions between source, speaker, and timbre.

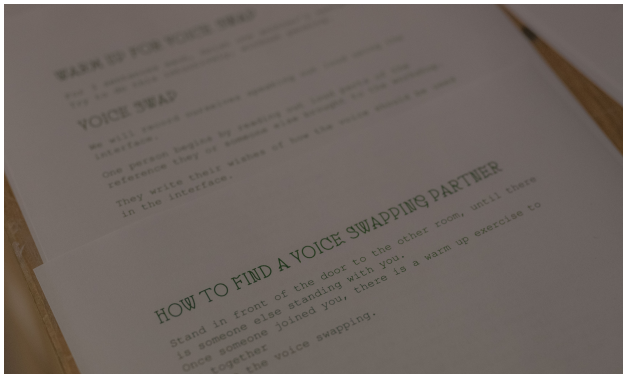

Figure 3: A photo of one of the many vocal data scores developed for An (In)audible Chorus, depicting instructions for a duet of voice gifting within the ad-hoc recording studio of the workshop at Morphine Raum, Berlin (2022).

Following this intervention, multiple new voice models were integrated into the BroadCaster, transforming it from a singular vocal presence into a heterogeneous, multi-vocal instrument. My own relationship to the act of composing also shifted dramatically following this workshop. I was no longer working exclusively with the disembodied voice of Linda Johnson, but with voices that had been offered in person, accompanied by explicit conversations about trust, use, and future circulation. This felt, in some way, like a much stronger obligation than working with textual contributions collected in the previous workshop. Voice demanded more care. And from this point onward, with voice and text contributions coming together, composing began to cohere as the re-animation of an archival corpus reflecting lived experiences. Thus, across these interventions, dataset creation functioned less as a preparatory step and more as a form of musicking. The Broadcaster was deconstructed in-situ through social-technical co-opting. Artistic and musical decisions were co-produced through collective activity; after which the system was reconfigured and released back to its radio audience. A short EP titled *Seeking Truth Seeking* was produced as a final outcome of the broadcast, produced as a quartet between vocalists Faye Houston, Jo Bramli and Kassia Zermon together with a live-coded version of The BroadCaster. The EP was played on German public radio as a farewell ritual to the broadcast, airing exactly after radio silence at midnight on March 31st, 2023 [4].

5 Discussion: Implications for NIME

5.1 Machine Learning Instruments as Performable Archives

Early NIME discourse was dominated by Perry Cook's principles, which emphasized the importance of high-resolution mapping and the "direct" control of sound by the human performer [5]. In this paradigm, liveness is a function of "physical and temporal co-presence" between the performer and the audience, as defined by Philip Auslander [2]. The goal was transparency: the audience should see a clear causal relationship between the performer's gesture and the resulting sound.

ISoGA represents a shift away from this concentrated "mastery" model toward what might be considered a temporally and spatially distributed form of liveness. Rather than a linear control chain, the BroadCaster system operates as an assemblage of interacting processes whose behavior emerges through feedback, reuse of past material, and punctuated social encounters and re-animations. The architectural looseness in particular allows for drift and forgetting, qualities that were traditionally viewed as errors in instrument design [5].

Rebecca Fiebrink's concept of ML systems as "meta-instruments" offers a way of extending techno-musical notions of where performance lies to data practices [18]. Recent research has begun to articulate this new form of liveness specifically regarding generative systems. For instance, the Hauntographic Analysis method of machine learning instruments being developed by Nicola Privato has given a thorough treatment to the temporal "out-of-jointness" of machine learning based instruments [33, 34]. There is also the voice-specific work of Kelsey Cotton on re-animation of voice through "bringing back" the hidden bodies of vocal machine learning systems [14].

This work reframes machine-learning systems not primarily as generators of new material, but as performable archives whose memory is enacted rather than stored. By treating learned

[4] A full recording of the farewell broadcast can be heard here https://www.researchcatalogue.net/view/3866796/3866991/0/6038

representations as mutable and unstable, the BroadCaster foregrounds memory as a dynamic, situated process rather than a fixed resource. For NIME, this suggests a shift away from viewing datasets as static prerequisites and toward understanding archives as ongoing sites of performance, interpretation, and change. Designing instruments that expose and work with this instability can open new aesthetic and conceptual territory.

5.2 Dataset Creation as Musicking

In Thor Magnusson's discussion of his instrument the Threnoscope, he highlights how the programming infrastructure preexists the performance, suggesting that liveness brings with it an "illusion of the clean slate" [6]. *ISoGA* begins by foregrounding the conditions of this illusion. The work plainly articulates its starting conditions, being trained on the widely used LJ Speech and The Pile datasets, making the historical inertia of these datasets part of the artistic concept [35]. Further, by allowing the system's underlying conditions and design to change publicly over a year-long duration, it in effect enacts a composition that does not only unfold over weeks and months, but fundamentally changes its underlying assumptions. This moves the locus of liveness from the "now" to a temporal distribution across dataset creation, iterative system consolidations and embodied encounters.

A central implication of this work lies in its treatment of dataset creation as a primary vehicle of musicking. The collaborative workshops functioned not only to supply training data, but were re-animating encounters themselves, ones that actively shaped the future behavior and reception of the instrument through collective action. This approach challenges common separations between data preparation, instrument design, and performance, suggesting that social and technical processes can be integrated into a single performative continuum. For NIME practitioners, this points toward instruments whose design explicitly accommodates intervention, inscription, and reconfiguration by multiple participants over time.

5.3 The Epistemological Perspective of the Archive

ISoGA also points to a particular ethical stance towards voice data, and indeed, digitized human traces used for machine learning more generally. Mainstream AI models rely on extremely large musical datasets often "scraped from the internet without clear consent or attribution" [8, 45]. Like any technical practice, this ideology embeds epistemological assumptions into musical systems [28]. This particular "all-you-can-scrape" mentality embeds neoliberal values of extractivism and deregulation as epistemological givens, often dismissing the contributions of musicians as fungible resources [31]. Many have argued that these epistemological givens are embedded more deeply, into the DNA of data-driven machine learning and data science writ large [16, 19], and even that a new specialization should be formed within ML that is focused on methodologies for data collection and annotation: efforts that require institutional frameworks and procedures inspired by the consent, power and ethics practices developed in archival work [25].

ISoGA attempts to stage a modest infrastructural intervention along these lines. Explicitly framing voice and textual contributions according to gift economies [1, 22], embedding consent, care, and future-oriented intent directly into the system's memory, sometimes as value-laden metadata, and sometimes in the vocal contributions themselves (voice contributors perform their future wishes and intentions in speech and song). This restorative approach aligns with "Responsible AI" principles, particularly transparency, fairness, and accountability. By holding workshops like *Re-Writing / Re-Reading* and *An (In)audible Chorus*, *ISoGA* ensures active participation in data creation and annotation labels which are context-rich. These workshops move the artistic focus of music from sonic production to critical reflection and social integration.

5.4 Designing an Open Work

Taken together, these implications suggest that machine-learning instruments can be understood not only as tools for real-time interaction, but as evolving cultural artifacts whose musical significance emerges through the archival dimensions of time, care, and collective engagement. Re-animation offers one conceptual framework for articulating this shift, positioning machine learning as a material for designing instruments that listen, remember, and change alongside their human collaborators and their cultures.

Finally, it is worth noting that the BroadCaster highlights the creative potential of designing machine-learning instruments that do not converge toward stability or optimization. By privileging openness, heterogeneity, and perpetual retraining, the system resists the impulse to refine away uncertainty and imperfection. Instead, these qualities become central to its musical aesthetic. This orientation encourages NIME designers to consider a particular way in which instruments may be left partially unresolved.

6 Conclusion

This paper has introduced *In Search of Good Ancestors / Ahnen in Arbeit* as a case study in a machine-learning system as archival reanimation. Through a year-long public broadcast, collaborative gift-based dataset creation, and an intentionally open-ended generative instrument, the work demonstrates how machine learning can be approached as a material for musical performance that unfolds through collaborative contribution and long-durational thinking. By framing re-animation as a central design orientation, the paper highlights alternative possibilities for liveness, instrumentality, mimesis and authorship within machine-learning-based music systems. Rather than prioritizing optimization or enduring presence, the work foregrounds accretion, drift, forgetting, and collective interventions as productive instrumental forces.

Viewed from this perspective, the training and curation of a machine learning model need not be understood as a technical prerequisite to musical activity, but as a continuous, performative, and social practice in its own right — one whose decisions carry aesthetic, ethical, and political weight. This reframing invites NIME practitioners to consider instruments whose musical significance emerges through duration and care rather than virtuosic immediacy: instruments that change publicly, remember imperfectly, and grow through community engagement rather than individual mastery. In this sense, community is not an afterthought to instrument design but one of its primary constitutive forces.

While *ISoGA* represents a specific artistic manifestation, the concepts and practices discussed here point toward broader implications for the field. Long-durational machine learning performance - and the questions it raises around archivality, relistening, memory, and collective inscription - remains a comparatively underexplored territory within NIME discourse. This paper offers one starting point for that conversation, and the frameworks developed here, re-animation, musicking-as-data-creation, and gift-based contribution, are intended as open resources for future work exploring how ML-based instruments might listen, remember, and change alongside their human collaborators and the communities that bring them into being.

7 Ethical Standards

This research was conducted as part of an artistic research project within a university doctoral program. The workshops described in this paper involved voluntary participation by adult individuals. Participants were informed in advance about the nature of the project, the use of voice recordings and textual contributions within a generative machine-learning system, and the public broadcast context in which outputs would appear. Voice and text contributions were gathered under explicit conditions of consent. Participants were informed that their contributions would be used to train or fine-tune generative models incorporated into The Broadcaster system, and that resulting synthetic outputs might be publicly broadcast. Workshop participants retained the right to withdraw their contributions prior to model integration. No biometric identification or sensitive personal data were collected beyond voluntary voice recordings and textual materials provided for artistic purposes.

No commercial sponsorship or corporate funding influenced the design, execution, or interpretation of the research. The author declares no financial or non-financial conflicts of interest related to this work.

Acknowledgments

In Search of Good Ancestors was originally commissioned by CTM Festival, Deutschlandfunk Kultur, and ORF Ö1 Kunstradio. The broadcast ran continuously from January 2022 to February 2023, with fragments aired on German Public Radio. The project was supported by the Leverhulme Trust as part of the Interdisciplinary Doctoral Scholarship Programme From Sensation and Perception to Awareness at the University of Sussex.

Model training resources for the neural audio synthesis system were provided through the University of Sussex and the EMUTE Lab. The development of the neural synthesis system and associated training approaches benefited from consultation with researchers at Coqui.ai.

Workshop activities exploring voice data ownership, voice sharing, cloning, and intergenerational transmission were conceived in collaboration with Angeliki Diakrousi, Joana Chicau, amy pickles, and Cristina Cochior of VARIA, with additional inspirational discussions and writing contributed by Eleni Ikoniadou. The author gratefully acknowledges all workshop participants and voice contributors whose recordings and reflections form the living archive of the project, as well as visitors to the web portal who contributed textual materials to the evolving broadcast. The farewell radio album Seeking Truth Seeking featured improvisations by Faye Houston, Kassia Zermon, and Johanna Bramli in duet with the system in its final broadcast phase.

References

[1] Frank Adloff and Björn Bosserhoff. 2022. *Politics of the gift: towards a convivial society.* Bristol University press, Bristol.

[2] Philip Auslander. 2022. *Liveness: Performance in a Mediatized Culture* (3 ed.). Routledge, London. doi:10.4324/9781003031314

[3] Jeronimo Barbosa, Joseph Malloch, Marcelo M Wanderley, and Stéphane Huot. 2015. What does "Evaluation" mean for the NIME community?. In *International Conference on New Interfaces for Musical Expression.*

[4] Frederic Charles Bartlett. 1932. *Remembering: A study in experimental and social psychology.* Cambridge University Press, New York, NY, US. xix, 317 pages.

[5] S. M. Astrid Bin. 2018. *The Show Must Go Wrong: Towards an understanding of audience perception of error in digital musical instrument performance.* Ph. D. Dissertation.

[6] Alan F. Blackwell, Emma Cocker, Geoff Cox, Alex McLean, and Thor Magnusson. 2022. *Live Coding: A User's Manual.* MIT Press.

[7] Terence Broad, Frederic Fol Leymarie, and Mick Grierson. 2021. Network Bending: Expressive Manipulation of Deep Generative Models. In *Artificial Intelligence in Music, Sound, Art and Design.* Springer, Cham, 20–36. doi:10.1007/978-3-030-72914-1_2

[8] 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 (AM '25).* Association for Computing Machinery, New York, NY, USA, 70–81. doi:10.1145/3771594.3771601

[9] Antoine Caillon and Philippe Esling. 2021. RAVE: A variational autoencoder for fast and high-quality neural audio synthesis. arXiv:2111.05011 (Dec. 2021). doi:10.48550/arXiv.2111.05011 arXiv:2111.05011 [cs, eess].

[10] Edresson Casanova, Julian Weber, Christopher Shulby, Arnaldo Candido Junior, Eren Gölge, and Moacir Antonelli Ponti. 2022. YourTTS: Towards Zero-Shot Multi-Speaker TTS and Zero-Shot Voice Conversion for everyone. arXiv:2112.02418 (Feb. 2022). http://arxiv.org/abs/2112.02418 arXiv:2112.02418 [cs, eess].

[11] Paul Clarke, Simon Jones, Nick Kaye, and Johanna Linsley. 2018. *Artists in the Archive: Creative and Curatorial Engagements with Documents of Art and Performance.* Routledge.

[12] Sasha Colby. 2015. Live from the Archive: Film, Folders, and Mina Loy. *Performance Matters* 1, 1–2 (May 2015), 70–75.

[13] Paul Connerton. 1989. *How Societies Remember.* Cambridge University Press, Cambridge. doi:10.1017/CBO9780511628061

[14] Kelsey Cotton, Katja De Vries, and Kıvanç Tatar. 2024. Singing for the Missing: Bringing the Body Back to AI Voice and Speech Technologies. In *Proceedings of the 9th International Conference on Movement and Computing.* ACM, Utrecht Netherlands, 1–12. doi:10.1145/3658852.3659065

[15] Jacques Derrida. 1996. *Archive Fever: A Freudian Impression.* University of Chicago Press.

[16] Catherine D'Ignazio and Lauren F. Klein. 2020. *Data Feminism.* MIT Press.

[17] Alice Eldridge. 2007. *Collaborating with the behaving machine: simple adaptive dynamical systems for generative and interactive music.* PhD Thesis. University of Sussex. https://archive.ecila.org/ecila_files/content/papers/thesis/behavingMachines_frontMatter.pdf

[18] Rebecca Fiebrink, Dan Trueman, and Perry R. Cook. 2009. A meta-instrument for interactive, on-the-fly machine learning. In *Proceedings of the International Conference on New Interfaces for Musical Expression.* International Conference on New Interfaces for Musical Expression, 280–285. https://collaborate.princeton.edu/en/publications/a-meta-instrument-for-interactive-on-the-fly-machine-learning/

[19] Marion Fourcade and Daniel N Kluttz. 2020. A Maussian bargain: Accumulation by gift in the digital economy. *Big Data Society* 7, 1 (Jan. 2020), 2053951719897092. doi:10.1177/2053951719897092

[20] Leo Gao, Stella Biderman, Sid Black, Laurence Golding, Travis Hoppe, Charles Foster, Jason Phang, Horace He, Anish Thite, Noa Nabeshima, Shawn Presser, and Connor Leahy. 2020. The Pile: An 800GB Dataset of Diverse Text for Language Modeling. arXiv:2101.00027 (Dec. 2020). http://arxiv.org/abs/2101.00027 arXiv:2101.00027 [cs].

[21] Owen Green. 2013. *User serviceable parts: Practice, technology, sociality and method in live electronic musicking.* PhD Thesis. City University London. https://openaccess.city.ac.uk/id/eprint/2730/

[22] Mary Jo Hinsdale. 2022. Of Gifts, Reciprocity, and Community. *Philosophy of Education* 78, 1 (2022), 38–51. doi:10.47925/78.1.038

[23] Keith Ito and Linda Johnson. 2017. The LJ Speech Dataset. https://keithito.com/LJ-Speech-Dataset

[24] Robert H. Jack, Tony Stockman, and Andrew McPherson. 2016. Effect of latency on performer interaction and subjective quality assessment of a digital musical instrument. In *Proceedings of the Audio Mostly 2016 (AM '16).* Association for Computing Machinery, New York, NY, USA, 116–123. doi:10.1145/2986416.2986428

[25] Eun Seo Jo and Timnit Gebru. 2020. Lessons from archives: strategies for collecting sociocultural data in machine learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20).* Association for Computing Machinery, New York, NY, USA, 306–316. doi:10.1145/3351095.3372829

[26] Théo Jourdan and Baptiste Caramiaux. 2023. Machine Learning for Musical Expression: A Systematic Literature Review. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Zenodo, 319–331. doi:10.5281/zenodo.11189198

[27] Jaehyeon Kim, Jungil Kong, and Juhee Son. 2021. Conditional Variational Autoencoder with Adversarial Learning for End-to-End Text-to-Speech. arXiv:2106.06103 (2021). doi:10.48550/arXiv.2106.06103 arXiv:2106.06103 [cs].

[28] Thor Magnusson. 2009. Of Epistemic Tools: musical instruments as cognitive extensions. *Organised Sound* 14, 2 (Aug. 2009), 168–176. doi:10.1017/S1355771809000272

[29] Thor Magnusson. 2019. *Sonic Writing: Technologies of Material, Symbolic, and Signal Inscriptions*. Bloomsbury Academic.

[30] Niamh Moore, Rachel Thomson, and Ester McGeeney. 2023. *Putting place back into the patriarchy through rematriating feminist research: The WRAP project, feminist webs and reanimating data*. Routledge.

[31] Fabio Morreale, Megha Sharma, and I.-Chieh Wei. 2023. Data Collection in Music Generation Training Sets: A Critical Analysis. 37–46 pages. doi:10.5281/zenodo.10265217

[32] Alessandro Portelli. 2009. *What Makes Oral History Different*. Palgrave Macmillan US, New York, 21–30. doi:10.1057/9780230101395_2

[33] Nicola Privato. 2026. *Hauntography. Algorithmic Presence and the Out of Joint in Music AI*. PhD Thesis. University of Iceland.

[34] Nicola Privato and Thor Magnusson. 2024. Querying the Ghost: AI Hauntography in NIME. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Zenodo, 432–438. doi:10.5281/zenodo.13904901

[35] Jonathan Chaim Reus. 2022. In Search of Good Ancestors / Ahnen in Arbeit. In *Nordic Human-Computer Interaction Conference (NordiCHI '22)*. Association for Computing Machinery, New York, NY, USA, 1. doi:10.1145/3546155.3547294

[36] Claudia Schmitz and Nicola Leonard Hein. 2021. UnStumm - Artificial Liveness. doi:10.21428/92fbeb44.a1d9cdc3

[37] Diemo Schwarz. 2007. Corpus-Based Concatenative Synthesis. *IEEE Signal Processing Magazine* 24, 2 (March 2007), 92–104. doi:10.1109/MSP.2007.323274

[38] Christopher Small. 1998. *Musicking: The Meanings of Performing and Listening*. Wesleyan University Press.

[39] Eleanor Spens and Neil Burgess. 2024. A generative model of memory construction and consolidation. *Nature Human Behaviour* 8, 3 (March 2024), 526–543. doi:10.1038/s41562-023-01799-z

[40] Diana Taylor. 2003. *The Archive and the Repertoire: Performing Cultural Memory in the Americas*. Duke University Press.

[41] Rachel Thomson and Liam Berriman. 2023. Starting with the archive: principles for prospective collaborative research. *Qualitative Research* 23, 2 (April 2023), 234–251. doi:10.1177/14687941211023037

[42] Rachel Thomson and Ester McGeeney. 2024. Reanimating data: Working with archives to revitalise young sexualities, past and present. *Health Education Journal* (Dec. 2024), 00178969241304725. doi:10.1177/00178969241304725

[43] Rachel Thomson, Alex Peverett, and Janet Holland. 2024. Revisiting young masculinities through a sound art installation: What really counts? *The Sociological Review* 72, 6 (Nov. 2024), 1259–1276. doi:10.1177/00380261231222726

[44] Varia. 2020. Read & Repair feat. Minimal Viable Learning. https://varia.zone/en/2020/rr-minimal-viable-learning/

[45] Elizabeth Wilson, Anna Wszeborowska, and Nick Bryan-Kinns. 2025. A Short Review of Responsible AI Music Generation. In *Proceedings of the 6th AI Music Creativity Conference (AIMC 2025), Brussels, Belgium, September 10-12, 2025*. doi:10.5281/ZENODO.16946342