Imploding between the facts and concerns: analysing human–AI musical interaction

Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-025-04533-4 OPEN

Imploding between the facts and concerns: analysing human–AI musical interaction

Kelsey Cotton [1✉], Anna-Kaisa Kaila[2,3], Petra Jääskeläinen[2,3], André Holzapfel [2,3] & Kıvanç Tatar [1,3]

The advancement of AI-tools for musical performance has inspired exciting opportunities for interaction with musical-AI-agents. Interactions between humans and AI-agents in musical settings entail dynamic exchanges of control and power, and framings of AI-agents' roles by human performers. We probe these framings and power-control exchanges through qualitative thematic lenses, drawing from post-phenomenology, *matters of fact and concern* and feminist science and technology studies. We contribute with a novel interdisciplinary analytical method as a tool for developers and designers of AI systems to help visibilise and examine the *implicit*, the wider connections and entangled filaments in Human–AI musical interactions.

[1] Data Science and AI Division, Computer Science and Software Engineering, Chalmers Institute of Technology and University of Gothenburg, Gothenburg, Sweden. [2] Division of Media Technology and Interaction Design, KTH Royal Institute of Technology, Stockholm, Sweden. [3] These authors contributed equally: Anna-Kaisa Kaila, Petra Jääskeläinen, André Holzapfel, Kıvanç Tatar. ✉email: kelsey@chalmers.se

Introduction

The significant technological advances in machine learning (ML) models and artificially intelligent systems have seen a rising tide of engagement with these technologies in the creation and performance of artistic work. Historically, artists have always been at the forefront of engagement with new technologies (Candy et al. 2018; Ellul and Hofstadter 1979; Girão and Céu Santos, 2019), utilising advanced tools, systems and platforms. Artistic engagement with new technologies has led to the birth of significant cultural movements and genres within the world of art including electronic art; mechanical art; computer art; digital art; computer-aided art; generative art; computer generated art; evolutionary art; robot art; interactive art; NFT-art, VR art, and AI art (Boden and Edmonds, 2009; Thomson-Jones and Moser, 2021; Wang and Wang, 2021).

Increasingly, the discourse in AI-Arts has been concerned with the implications and embedded values of AI (Bramantyo, 2021; Cotton and Tatar, 2023; Rani, 2018; Tatar et al. 2023). As discussed in Tatar et al. (2023), the availability of AI models has enabled more accessible ways for producing content, which has subsequently initiated a paradigm shift within artistic practices. Tatar et al. discuss the concerns of this shifting status quo primarily in relation to the role of human labour in AI art; the democratisation of AI tools; and considerations for improving cultural sustainability and inclusion. We see these concerns as especially pertinent to the context of music, which has seen widespread early-adoption across social media platforms, and within mainstream music media (Arunsaravanakumar, 2023; Grimes [@Grimezsz], 2023a, 2023b; McEvoy, 2023; Surbano, 2023).

The utilisation of AI technologies within music has established a new domain: musical-AI. This particular field utilises machine learning and artificial intelligence for the ideation, composition, creation, analysis and performance of music (Bretan and Weinberg, 2016; Briot and Pachet, 2020; Carnovalini and Rodà, 2020; Herremans et al. 2017; Tatar and Pasquier, 2019; Weinberg et al. 2020a). Within the context of research into human and agent interaction in musical-AI, there has not yet been much consideration into the politics at play within these interactions. Although this has previously been flagged as an area of concern within academic discourse (Tatar et al. 2023), we currently lack methodologies to examine how human and AI-agents relate to each other in musical interactions, and what is implicitly communicated through how these interactions unfold. As the capabilities of AI technologies continue to (seemingly exponentially) improve, and the barrier of access continues to lower it is imperative that we prioritise the development and implementation of analytical methodologies. Current research solely prioritising technological advancement of AI technologies in such settings neglects the complexity of how humans and AI-agents engage one another in artistic settings, and the wider sociocultural correspondences that these interactions invite in.

Our concerns for how un-reflected and un-critiqued utilisation of musical-AI technologies invisibilises societal, cultural and political implications establish our research questions. These are:

1. how AI is positioned by a human in human–AI musical interactions and what this communicates about power/control within the interaction; and
2. how sociopolitical perspectives are implicitly embedded in the interaction.

We have investigated these research questions by developing and applying a novel analytical method to four artworks that utilise or feature musical-AI-agents performing with a human performer. As the domain of musical-AI continues to grow, we had to develop practical constraints to assist in executing our analysis. We, therefore, delimited our selection based on a set of criteria, which are fully outlined in the section "Some practical constraints of our analytical method".

This paper offers a series of contributions. Firstly, we contribute with an interdisciplinary methodology—intermedial analytical method—that examines the implicit values in human–AI musical interaction. Our rationale for drawing together interdisciplinary theory is grounded in the need for a clear methodology for identifying and probing the connections between the roles an AI-agent assumes in a musical performance in relation to a human (and vice versa), and the implicit connections to wider sociopolitical implications of these roles. Secondly, we apply this method to examining four case studies of human–AI musical artworks. Thirdly, we contribute with several considerations for developers, performers, and users of musical-AI-agents and systems to examine the implication of subjective choices in how they frame AI-agents in musical performance contexts. We see an urgent need for these communities to critique and reveal dimensions of their work that may have deeper implications (such as the choice of model; the usage of particular datasets; decisions in mapping input-output). We view this as a significant area for future work, especially within the domain of musical-AI, but also as important for other creative domains of AI research.

The structure of this paper is as follows. In the "Theoretical background" section, we outline existing discourses pertinent to this paper and establish our definitions and usages of terminology and theory in line with this paper's aims. In the "Methods" section, we outline the steps of our analytical method as well as some practical constraints we had to develop for our analysis. In the "Analysis" section, we provide a brief outline of our four selected case studies and present the results of our analysis. We synthesise the emergent findings from this analysis in our "Discussion" section, and we examine the "matters of concern" emerging from our methodology.

Theoretical background

This paper engages with a cross-section of theory that ultimately informed the development of our novel analytical method in the section "Methods". This methodology—intermedial analytical method—draws together a series of theoretical perspectives. We will take a moment now to give a general outline of each of the theories that have informed our intermedial analytical method, and to rationalise *how* each theory contributes to the development of our intermedial analytical method. To give a structural outline of this section, we first discuss Latour's *matters of fact and concern*, then give a brief overview of Feminist Science and Technology Studies, then Post-phenomenology, we further establish some definitions of enacted and situated agency and introduce the artists of our four case studies (Haraway, 2018; Ihde, 1990; Latour, 2014).

Overview of Latour's matters of fact and concern. Bruno Latour was a French sociologist known for his contributions to science and technology studies. Latour is especially renowned for his development of Actor-Network Theory (Latour, 2005) and his critique of scientific 'fact' as objective. He proposed the concept of "*matters of fact*" and "*matters of concern*" as a mechanism for both critiquing the systems that form knowledge and developing an awareness of what this knowledge represents within the world (Stephan, 2015). Latour defines a *matter of concern* thus:

> A matter of concern is what happens to a matter of fact when you add to it its whole scenography, much like you would do by shifting your attention from the stage to the whole machinery of a theatre (Latour, 2014).

Latour elaborates on how shifting our focus is instrumental to our understanding of the relationships that unfold between actors/stakeholders and this scenography:

> Instead of simply being there, matters of fact begin to look different, to render a different sound, they start to move in all directions, they overflow their boundaries, they include a complete set of new actors, they reveal the fragile envelopes in which they are housed (Latour, 2014).

In this paper, we engage with Latour's concepts of *matters of fact* and *matters of concern* by examining the objective or factual components of the human–AI musical interaction, and the relation of these facts to their 'whole scenography'. To make the parallel between the "machinery of theatre" Latour speaks about, and the context of this paper: we see matters of concern as constituting far more than just the mere human–AI musical interaction that is unfolding. Here, we look at the entire network of cultural, social, political and other entanglements which are the "machinery" of the musical–AI interaction.

Below, we continue with an overview of Post-phenomenology, which enables us to investigate the role of AI technology in relation to humans in musical performance. We also discuss post-phenomenology of sound as a theoretical foundation for the musical aspects of our analysis.

Overview of post-phenomenology. Post-phenomenology is a philosophical approach to critically examining the role of technology and technological artefacts in relation to humans. It examines how technology interacts with humans and shapes our relations to the world (Ihde, 1990; Rosenberger and Verbeek, 2015a; Verbeek, 2008). It emerged as a response against traditional phenomenological thought, which emphasises direct experiences and the 'intentionality of consciousness' (Pula, 2022). Post-phenomenology expands traditional phenomenology by considering technology as a mediator of human experience with the world and external factors. The classification of particular human–technology relations has been proposed previously by Don Ihde (1990), who conceptualised human–technology interactions as fulfilling four different forms: hermeneutic, alterity, background and embodiment. This has been critiqued on its singular focus "for analysing human–technology configurations in which technologies are *used*" (Rosenberger and Verbeek, 2015b). But what happens when technologies *use*? The complexity of entanglements between humans and technology is especially clear in recent work within HCI, and its use of post-phenomenology to frame and examine human–technology relationships. This is especially clear in recent work within human–computer interaction (HCI) (Frauenberger, 2019; Frauenberger et al. 2010; le Roux et al. 2019), and the use of post-phenomenology as a reflective analytical tool (Benjamin et al. 2021; Fallman, 2011; Jensen and Aagaard, 2018; Odom et al. 2009; Ohlin and Olsson, 2015).

Post-phenomenology of sound. We further ground our method using the theoretical perspectives from sound studies, as this paper examines human and AI-agent musical interactions. Specifically, we engage with post-Schaefferian theory of sound and listening.

Post-Schaefferian criticisms comment on Schaefferian theories on sound and listening. Pierre Schaeffer was a French engineer and musician renowned for his development of *musique concrète* (translation: concrete music)—an approach to creating electro-acoustic music that utilises pre-recorded sounds, natural environments, synthesisers and digital signal processing. Schaeffer is also noted for his development of a philosophy of listening—*écoute réduite* (translation: reduced listening). Schaeffers's *écoute réduite* is proposed in his *Traité des objets musicaux* (translation: Treatise on Musical Objects) (Schaeffer et al. 2017). In this text, Schaeffer develops Edmund Husserl's phenomenological notion of reduction (Husserl, 2012) to develop an understanding of the different ways that we listen to sound. Schaeffer proposed a series of listening modes that seek to separate a sound object from its notation; creation and other contextual connections that arise when we listen to a sound (Kane, 2007; Schaeffer et al. 2017).

A *post*-Schaefferian theory of sound and listening rejects the separation of a sound object from other dimensions such as *reflective*, *denotative*, and *experiential* (Tuuri and Eerola, 2012). Instead, post-Schaefferian theory conceptualises sound as "contain[ing] references to its actual or perceived origins, to some external association, or to some combination of the two" (Demers, 2010). Demers argues that "Sound, in other words, is a sign that indicates something beyond itself and as such can never exist as a pure abstraction." (Demers, 2010). Drawing together Demers' account of post-Schaefferian views on sound as representing "something beyond itself", we can therefore understand post-Schaefferian listening as a post-phenomenology of sound.

As our paper is concerned with examining the role of AI technology in relation to humans in musical performance contexts, Post-phenomenology provides a clear framing as to how this may be approached. Ihde, Verbeek and Rosenberger (Ihde, 1990; Rosenberger and Verbeek, 2015b; Verbeek, 2008) discuss how this can be approached "in the field" by classifying the "roles" that a technology seems to be fulfilling in its 'use' by a human. In our text, we look (and listen) in a post-Schaefferian sense, to see what is being communicated "beyond itself" (Demers, 2010) within the machinery and the scenography of the artwork (Latour, 2014).

Expanding human–technology relations: Navigating situated and enacted agency. Throughout this paper, we use the terms *human–AI musical interaction, AI-agent*, and *musical-AI-agent*. Human–AI musical interaction refers to an interaction (and in our context, a performance) that takes place between a human and an AI-agent. By AI-agent, we mean a technology that has the capacity to perform as an autonomous agent and is utilising machine learning algorithms. By musical-AI-agent, we mean the previously defined AI-agent that is specifically applied to a musical context.

Our understanding of agents is rooted in the multi-agent systems (Wooldridge, 2009) and musical agents (Tatar and Pasquier, 2019) literature. In the literature, a simplified world of reality is presented as an analogy to understand agent, object and environment relations. It can be described as follows. Agents have the capability of performing actions, whereas objects have affordances that change how these actions can be performed on, to, with, or around them. Agents also have affordances in the way that they execute actions. These affordances are enacted upon, or with, other agents. Agents carry out their actions through and with an enacted agency. An environment is where agents and objects are situated. This environment further has situated agency, which is acted upon all agents and objects within the environment.

With this understanding of a simplified world of reality in which objects and agents enact and are acted upon within an environment, we can better understand how agency is enacted by a human in the composition of an artwork: the world of the performance. In our text's investigation of the interaction between agents (human and musical-AI-agents), we seek to expand Ihde's human–technology

relations. We look beyond an objective classification of how a technology is positioned in relation to a human (and vice-versa). We need to also factor in an artist's subjective framing of technology in relation to the sonic and visual dimensions of their artwork. We view this subjective framing as occurring via enacted and situated agencies. In our musical performance context, these agencies constitute the decision-making possibilities within constraints that are established either pre- or during the performance. We argue that agency in these interactions may unfold—dynamically—in differing ways.

The situated agency of an environment involves the conditioning of possibilities through design decisions (Barad, 1998, 2003; Draude, 2020; Gondomar and Mor, 2021; Limerick et al. 2014; McEneaney, 2013; Worthy et al. 2021). Enacted agency is the "put into practice" action that is a subset of the situated agency (the actions that are possible to take) within the situated agency of the world or environment (Pyysiäinen, 2021; Rietveld and Kiverstein, 2014; Withagen et al. 2017).

Overview of feminist science and technology studies. Feminist science and technology studies (feminist STS) is an interdisciplinary research field that explores the overlap of gender, science and technology (Bailey and Cuomo, 2008; Harding, 2009; Mayberry et al. 2001; Weber, 2013). Its emergence as a research field came about as a response against technological practices which excluded marginalised groups within society (Carroll et al. 2020; Garry et al. 2017; Garry and Pearsall, 1996; J. Gray and Witt, 2021; J.E. Gray, 2022; Hampton, 2021; Ramazanoğlu and Holland, 2002; Schott, 2003; Wilkinson et al. 2016). Feminist STS has strong connections with the aims and perspectives of third-wave feminism (Ahmed, 2009, 2017; Beasley, 1999; Bordo, 1994; Gamble, 2001; Mayberry et al. 2001; Schott, 2003). In brief, Feminist STS is primarily concerned with formulating an understanding of how gender, amongst other societal and cultural factors such as socio-economic status, sociocultural connections, ethnicity, race, age, etc., shape and are shaped by the influence of science and technology (Åsberg and Lykke, 2010).

As our paper seeks to examine the embedded values within human–AI musical interaction, we draw from select concepts of entanglement from feminist science and technology studies. The most notable of existing entanglement narratives is perhaps Haraway's (2006) notion of the Cyborg: as a fictionalised reality of combined biological and machinistic bodies. Haraway's Cyborg communicates the idea that the boundaries we socio-culturally construct between bodies can and should be challenged. This line-of-thought is now widely accepted in science and technology studies (STS). Haraway argues that the power relations that we "arrange" are reconfigurable. This premise lays a grounding for our inspections of the interactions between AI and humans. To conduct these inspections, we engaged with a method proposed by Haraway: the Implosion (Haraway, 2018). The Implosion is a method for mapping the connections between objects, artefacts or systems to social, cultural and/or political networks. We argue that an Implosion is a suitable method for investigating the richness and complexity of the limits and intersections between human and AI bodies in musical performance (Haraway, 2018). We utilise the Implosion method to investigate the societal impact of technology facts with the aim to connect matters of facts in musical-AI to matters of societal concern.

An introduction of the artists discussed in this paper. In this paper, we will extensively discuss the works of composers, media artists and technologists renowned for their engagement of AI technologies within their respective artistic practices. We will now briefly introduce the artists behind the artworks we discuss and analyse in the forthcoming Analysis section.

Marco Donnarumma is an interdisciplinary performer–artist–technologist working with bodies, machines, sound and theatre to create performances and installations. Donnarumma is renowned for his solo performance work with self-constructed technological agents and prostheses, as well as his exploration of human–technology rituals (Donnarumma, 2019).

Jake Elwes is an artist–hacker–researcher queering and working with queer artificial intelligence. They create installations and performances, harnessing algorithms and machine learning to expose the successes and failures of these technologies to engage with diverse bodies (Elwes, 2015).

Marije Baalman is an artist-researcher-technologist known for her work in interactive art. As a researcher technologist, she is especially concerned with human–technology entanglements and real-time compositional processes when working with AI technologies. Baalman largely works with open source software and hardware, often constructing her own systems which she then open-sources (Baalman, n.d.)

Shimon is a robotic marimba player developed by Georgia Tech's Robotic Musicianship group. Shimon was developed as part of ongoing research at Georgia Tech, in which the Centre for Music Technology (CMT) develops and implements robotic and AI technologies to facilitate opportunities for music making between humans and machines. Georgia Tech's CMT (Shimon Robot & Friends-About, n.d.) is guided by their advocacy of how:

> "Real-time collaboration between human and robotic players can capitalise on the combination of their unique strengths to produce new and compelling music".

Methods

In this section, we outline the steps of our methodology and the selection criteria for the chosen case studies in the section "Some practical constraints of our analytical method". We then discuss how our engagement with these interdisciplinary theories helps to inform the questions we use at each step of our methodology. Our research methodology consists of several stages in which we bring together core theoretical positions and perspectives from disciplines such as science and technology studies, feminist science and technology studies and post-phenomenology.

The steps we took in our analysis, and the specific questions we used as a result of our interdisciplinary engagement with theory are presented in Fig. 1. We establish the usage of the phases and questions, detailed in Fig. 1, as a pipeline for other researchers to implement within their own analyses.

We will now outline the progression through this methodology. There are three distinct stages to our method, with a parallel examination using Haraway's implosion and post-phenomenology.

In stage 1 (see the sections marked A1 and B1 in Fig. 1), we conduct a parallel analysis of the *facts* and *concerns* within the interaction. In stage 2 (see all sections marked A2 and B2 in Fig. 1), we deploy a number of selected post-phenomenologically informed (see the section "Post-phenomenologically informed questions") and implosion-informed questions (see the section "Implosion-informed questions"). Our post-phenomenologically informed question is

> How is the AI-agent positioned within the interaction with the human performer in regards to the sonic space and/or the physical environment?

Our Implosion-informed questions are:

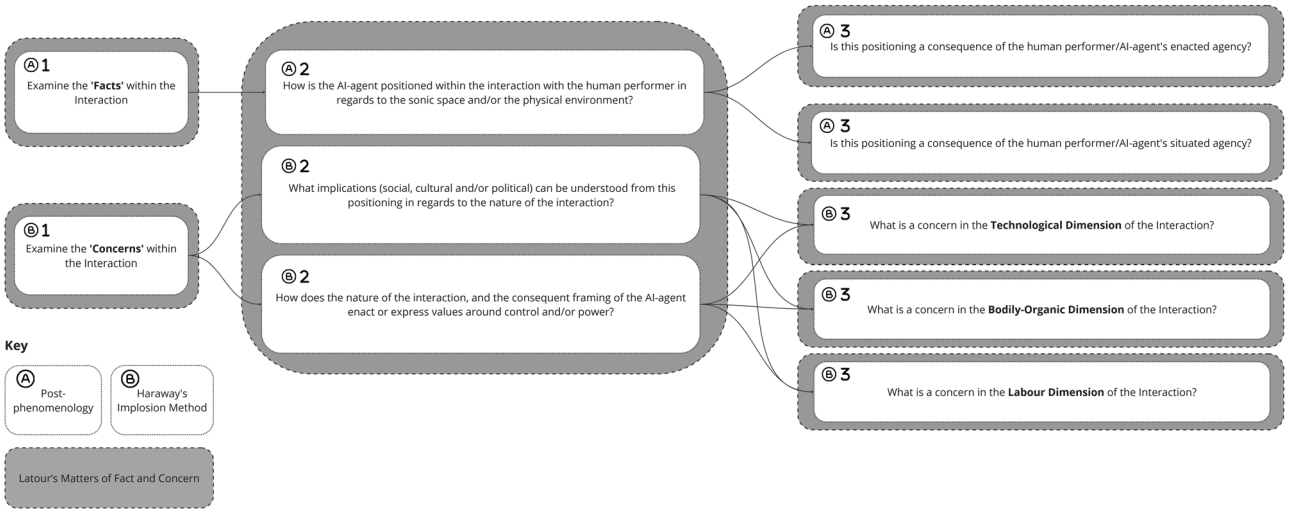

Fig. 1 The pipeline procedure outlining the methodological questions used in each phase of our analysis. The upper part (marked by A) is detailed in the subsection "Post-phenomenologically informed questions", and the lower (marked by B) is detailed in the section "Implosion-informed questions". The grey regions (marked as "Latour's Matters of Fact and Concern) are a common thread, or glue, between our engagement of post-phenomenology and Haraway's Implosion.

> What implications (social, cultural and/or political) can be understood from this positioning in regards to the nature of the interaction?; and
>
> How does the nature of the interaction, and the consequent framing of the AI-agent enact or express values around control and/or power?

The questions in this stage lead us to our third stage. In this stage (see all sections marked A3 in Fig. 1), we identify whether—and how—the identified phenomena and observations resulting from stage 2 influence how agency is communicated or navigated in the interaction or concerns in relation to other domains (see all sections marked B3 in Fig. 1).

Our method was conducted as a parallel process, in line with Latour's proposition to 'make parallel' the machinery and scenography of interaction (see Fig. 1) To do so, we establish what the 'machinery' and 'scenography' of the human–AI musical interaction is in each of our case studies. We position the 'machinery' of the interaction as the *matter of fact* of the interaction: how the AI-agent is positioned by the artist with regards to the sonic space and/or the physical environment. We positioned the 'scenography' as *matter of concern* for the interaction.

As we established in the section "Overview of post-phenomenology", post-phenomenology is an immensely helpful perspective for identifying human relations with technology. As we further established in the section "Overview of feminist science and technology studies", our usage of Haraway's Implosion assists in identifying connections between an artefact and the various dimensions that are connected to it. The combination of post-phenomenology and a feminist STS analytical approach therefore enables us to examine the connections between the 'facts' of the interaction and the 'concerns' we identified as emerging from the human–AI musical interaction. This reflects our engagement with Latour's call for examining the "matters of fact" in relation to the "matters of concern", as we established in the section "Overview of Latour's matters of fact and concern".

Post-phenomenologically informed questions. Our engagement with Post-phenomenology builds on Ihde's four human–technology relations (Ihde, 1990) and Rosenberger and Verbeek's argument on the subject (or the human) and the object (the technology) as existing in a connected manner (Rosenberger and Verbeek, 2015b). Specifically, we take our cue from Rosenberger and Verbeek's comment on how post-phenomenologists "*[analyse] the character of the relation human beings have with [a] technology and the ways in which it (the technology) organises relations between human beings and the world.*" (Rosenberger and Verbeek, 2015b) In our context, we understand this organisation as taking place through an *action* or series of actions- such as when humans use, touch, wear, listen to/through, look at/through a technology (see the section "Expanding human–technology relations: Navigating situated and enacted agency"). In our musical performance context, action takes place in the sonic and spatial domains. We therefore pose the following question in our Intermedial Analytical Method: *how is the AI-agent positioned within the interaction with the human performer in regards to the sonic space and/or the physical environment?* (see the sections marked A2 in Fig. 1). Through conducting our analysis from the perspective of an observer of the performance, we intend to examine the human-to-technology relation that we are witnessing, as it unfolds within each respective case study.

Implosion-informed questions. In this section, we explain our implementation of Haraway's Implosion (Haraway, 2018). Our usage of the Implosion requires some adaptation to our selected case studies. Dumit explains how an Implosion analysis constitutes an evaluation of 14 dimensions: Labour dimensions; Professional/Epistemological dimensions; Material dimensions; Technological dimensions; Context and situatedness; Political dimensions; Economic dimensions; Textual dimensions; Bodily/organic dimensions; Historical dimensions; Particle Dimensions; Educational dimensions; Mythological dimensions; and Symbolic dimensions (Dumit, 2014). This evaluation is conducted across four levels (Dumit calls these "twists"): constructing a knowledge map; constructing an ignorance map; locating the archives of where knowledge is held; and exhaustion of the connections across the 14 dimensions (Dumit, 2014).

As our Intermedial Analytical Method in this text is concerned with the unfolding human–technology relations within the four performances, we determine that there are three dimensions most relevant to addressing the research questions we establish in the

Introduction. These Implosion dimensions are: the technological dimension; the bodily organic dimension; and the labour dimension. Having determined our dimensions that we used in our Implosion for each case study, we then formulated the following questions to pose in our analysis of each case study: (1) *what implications (social, cultural and/or political) can be understood from the artists' positioning of the AI-agent in the interaction?* and (2) *how does the interaction, and the consequent framing of the AI-agent enact or express values around control and/or power?* (see the sections marked B2 in Fig. 1).

After establishing our Implosion dimensions and methodological questions, we then engage with one of Dumit's "twists". We did this by constructing a knowledge map of all of the phenomena we identify in each case study, according to our three dimensions. We then proceeded to a second "twist" where we examine the connections between our three chosen dimensions. This approach enables us to take a progressively expanding view. That is, we first examine the interaction, and then expand this examination to view connections between the interaction to other intercontextual dimensions. These two twists also reveal other aspects of human–AI musical entanglements (Haraway, 2006) that fall somewhat outside of the immediate 'world of the performance'. As an example, one implication upon a performance may be in the inherited 'residues' from the dataset that a model has been trained on. These inherited 'residues' can influence the aesthetic outcomes of an artwork.

Some practical constraints of our analytical method. We formulated this methodology with an understanding that the facts and concerns of human engagement with a musical-AI-agent (Haraway, 2006) form a wide network of bidirectional entanglements. We see these entanglements as first emerging from *how* an artist frames and positions the *role* a musical-AI-agent within the context of a performance, and then the implications of this positioning by the artist with relation to other dimensions. Our usage of Haraway's Implosion therefore seeks to uncover the scope of these relations. We do this by viewing the 'facts' of the interaction in direct relation to our selected technological; bodily organic; and labour dimensions.

It is important to note here that artists make crucial decisions about how they will engage with technological tools before the performance. These include decisions such as what type of algorithm will be used; how to map physical gestures to sound; what kinds of sounds will be used; how the artist will be clothed, lit and occupy the space; and how the AI-agent will be visually or sonically presented onstage are some examples of questions that artists answer before they take to the stage. These decisions and the choices artists make all have profound implications on what we, as an audience, see and understand about a performance featuring musical-AI-agents. As we are concerned with what is visibilised and invisibilised (Hampton, 2021) from our perspective as an audience, our usage of Haraway's Implosion afforded us a means to examine how these pre-performance decisions made themselves visible in our analysis of the presented artworks.

We selected the artworks for evaluation based on a practical set of criteria. As the potential pool of works for analysis is immense, these practical constraints facilitate a deeper evaluation. Our delimitations of the works selected for analysis are outlined as follows:

(a) Is there openly accessible documentation of the work online, consisting of both a video recording and additional technical details?;

(b) Does the work constitute a human–AI musical interaction? Namely, can the musical system be classified as a machine learning system; an agent system; or an AI?; and

(c) Does the work highlight a focus upon the interaction between human and AI, either in regards to the aesthetic decisions made in sound aesthetics; the musical or performative decisions made by the human musician; or the sound-interaction affordances or performative decisions made by the musical-AI system?

We have elected to solely examine musical works to probe the particularities of the human–AI musical interaction. We note that our positionality in this analysis is a post-performance, third-person stance. To do so, this requires a video documentation of musical performance. Unless indicated otherwise, the analysis of all case studies is conducted with the video material of the performance.

Analysis

In this section, we provide an overview of our selected case studies and the results of our intermedial analytical method.

Case Study 1: *Corpus Nil* by Marco Donnarumma. Donnarumma's work *Corpus Nil* (2016) https://vimeo.com/205899193 is a 14-min long solo performance from the *7 Configurations* cycle for a human and an AI system (see Fig. 2 for still image from performance). With regards to the technical components of *Corpus Nil*, the work utilises machine learning processes for movement data, which process and control sound and light. Movement data is collected from two sensors: an electromyography sensor to capture electrical voltages generated in muscular tissue (EMG) (Mills, 2005) and the bioacoustic sound produced by muscular tissue when it is in a state of activation/contraction (Oster and Jaffe, 1980). These sensors are placed on the upper arms. The raw data from these two biosensors is then parsed through feature extraction. The extracted features are: the muscular tension, contraction speed and recovery rate of a muscle post-contraction. These features provide information pertaining to the gesture and expressivity of Donnarumma's movements (Donnarumma, 2016). The expressivity of these movements influence three algorithms governing the sound production, mixing, lighting and sonic intensity of the piece. We will take a moment now to describe the most salient sonic and visual features of the artwork, introduced in the opening moments of the piece and sustained throughout. These aspects of the work will be analysed in the subsequent paragraphs.

The performance begins in total darkness, with a bank of clustered saw-tooth oscillators establishing a low drone. The overtones of the saw-tooth-like waves pulsate slightly as they come in and out of phase with each other. All the while, part of a heavily tattooed human figure is progressively illuminated in an aerial shot, under a diffused cold white light. It is initially difficult to determine exactly which part of the body we are viewing, as the rest of the space is shrouded in a blackout. As a darkly inked tattoo design is exposed to more light, a ribcage and shoulder blades begin to be more clearly in focus. We see skin stretched tightly over a ribcage and shoulder blades, taut tendons and straining muscles of the back. It becomes apparent that we are viewing the exposed shoulders and neck of the performer, from above. A higher-pitched, almost train whistle-like cluster suddenly joins the soundworld as the performer begins to twist, shift and twitch. A deep rumbling sub coincides with this subtle

Fig. 2 A photograph from a performance of Corpus Nil, depicting the performer in-scene. Image provided by, and approved for usage by the artist Marco Donnarumma.

choreography. The sonic effect is striking, foreboding. We watch the performer continue to twitch, roll and flex as the oscillators begin to shift more dramatically in pitch, gliding to ever increasing intervals. The rumbling sub begins to flicker and punch through the wall of oscillators more indiscriminately, seemingly coupled with the more extreme movements of the performer. The cold white light begins to flicker as the performance progresses, seemingly also connected to the performer's choreography.

How is the AI-agent positioned within the interaction with the human performer (in regards to the sonic space and/or the physical environment?)

Donnarumma's interaction with the musical-AI-agent in *Corpus Nil* is multi-faceted, dynamically evolving and complex (Grosz, 1994). The interaction between the two, is in a continual state of definition and re-definition, akin to Grosz's notion of "sense-bestowing" (Grosz, 1994). The ambiguity of how this "sense-bestowing" unfolds stems (perhaps) from the shifting agencies that we bear witness to: Donnarumma's physical body and its signal-level influence on the AI-agent's algorithms; the AI-agent's interpretation of Donnarumma's muscle signals as sound output; and how Donnarumma then responds physically to these sounds. We can clearly see that the AI-agent is specifically positioned in both the sonic space and physical environment in very particular ways. Both Donnarumma and the musical-AI-agent occupy different roles at different stages in the performance: they alternately exert control of the sonic and visual performance world at differing points; and present and enact power in differing capacities within the course of the artwork. Overall, we can understand the human–technology relation unfolding in *Corpus Nil* as dynamically shifting throughout.

We will first discuss our observations and analysis of how the musical-AI-agent is positioned in the *sonic* space. Overall, Donnarumma positions the AI-agent as incredibly prominent in manipulating and establishing the sonic space of *Corpus Nil*. Foremost, we observe this prominence in the musical-AI-agent's role in shaping the soundscape of the performance, based on Donnarumma's input biodata. Specifically, the AI-agent dynamically maps Donnarumma's muscle movements to the processing of the oscillator banks. In this regard, we can understand that the AI-agent establishes the sonic space in its processing of input to generate sound output. But this interaction is more complex than a simple input-output mapping relation. We argue that there is a dynamic shifting of agencies with regards to how the sound interaction is navigated by Donnarumma (as a performer and composer of the work). We see a navigation of enacted agency made apparent through the musical-AI-agent's mapping of Donnarumma's gestural input to provoke algorithmic reactions that govern the timbral evolution, spatialisation and musical form *Corpus Nil*. We further note that Donnarumma similarly has had a significant role pre-performance in choosing how exactly the sounds are generated. During the composition phase, he has constrained the decision-making potentials—the situated agency—of the AI-agent by choosing to use particular oscillators and choosing which biodata signals are mapped to which processing banks. Ultimately these choices establish the overall sinusoidal, glitching, rumbling sonic aesthetic of *Corpus Nil*. Further, these choices establish the situated agency of the performance world and inform the enacted agency of the musical-AI-agent engaging within Donnarumma's compositional constraints. The outcome of Donnarumma and the AI-agent interacting within the constraints of the sonic space creates an impression that this metallic, glitchy, droning soundworld emanates from the cyborgic, entangled Donnarumma-musical-AI-agent body (Grosz, 1994; Haraway, 2006). Overall, we may understand that the sonic space is dynamically constructed and manipulated by both the musical-AI-agent and Donnarumma at differing temporal stages, with each assuming different levels of control with respect to how sound is generated, manipulated and spatialised.

Considering the positioning of the AI-agent with regards to the *physical* environment, this is more ambiguous. Firstly, we must begin with noting that Donnarumma has obviously made a clear compositional choice as to what is physically present and absent on-stage. His body is the only body physically present in the performance environment. Further, Donnarumma has chosen to "blackout" the majority of the 'stage' and large portions of his own body to create a sort of 'body' of darkness. The overall effect of this 'body' of darkness is almost overwhelming, demanding a visual focus upon a relatively small area-of-attention amidst an expansive blackout. We can thus determine that Donnarumma's compositional choice on the lighting–and the aesthetic

implications this creates–constraints the enacted agency that is available to the AI-agent to interact with the physical environment. It is only able to respond to Donnarumma's biodata input to change the physical environment by plunging it into darkness, or illuminating it.

This invisibilisation of presence (both in regards to Donnarumma and the AI-agent) mainly unfolds through the visual domain via the change in light. We might draw parallels to it's invisibilised presence and the metaphorical darkness within which its functional role in the physical environment is enshrouded. Thus, although there is no physical body that the AI-agent inhabits of its own, it inhabits the body that is created through the blackness of the space. At times, the AI-agent also appears to inhabit or embody the body of Donnarumma himself (as we have discussed above).

Examining the AI-agent's physical positioning further, we view its algorithmic reactions to reshape our literal view of the performer onstage as an enactment of agency. That is, its physical absence through a defined and visible morphology communicates symbolically and literally that it is not the body we should focus our attention on during the performance. However, the metaphor of the AI-agent's body as represented by its control of the lighting and the onstage "darkness" it occupies juxtaposes this: the darkness becomes a tangible 'body' in its own right and naturally draws our attention. It should be emphasised that Donnarumma has also enacted agency here, making the compositional choice to map his biodata to the lighting output. Overall, we could view that exchange of agencies between Donnarumma and the musical-AI-agent are in a tightly coupled feedback loop: Donnarumma has constrained the environments within which the AI-agent can take an action related to the sonic space or physical environment, and Donnarumma in turn tries to provoke the AI-agent to provoke it to take additional actions.

What implications (social, cultural, and/or political) can be understood from this positioning (in regards to the nature of the interaction)?

Drawing now from the Bodily Organic Dimension from Haraway's Implosion, there is an inherent tension between the boundaries between the "synthetic" AI-agent and the "organic" Donnarumma. Indeed, as Haraway tells us, the Cyborg body present in *Corpus Nil* is continually being reconfigured by the performer and the AI-agent and further reconfigured through our understanding of what unfolds from our seat in the audience. Sonically, there is an ambiguity about what triggers the pitch shifts, what elicits the deep rumbling sub drone to overpower the soundworld, what the crackling electronic pops are triggered by. Donnarumma further reflects this sonic ambiguity in the visual domain: using angles, poses and choreography that render his body in-human and unfamiliar.

What implications (social, cultural, and/or political) can be understood from this positioning (in regards to the nature of the interaction)?

In *Corpus Nil*, we see implications as connected to the core interaction material of the performance: the physical choreography and the biodata values that are captured and used to provoke sonic and lighting changes throughout the artwork. Namely, the implications of Donnarumma's labour. Contextualising this within the technological dimension from Haraway's Implosion, *Corpus Nil* makes a very deliberate choice of technology to emphasise the artistic intention: to probe bodily and gestural interaction with an AI-agent. Foremost, we see cultural implications on labour in connection to AI, mainly through Donnarumma's pre-performance choices regarding how they utilise specific choreographic movement to generate biosignals for the input data of the musical-AI-agent.

Continuing in our observation of how pre-performance decisions implicate the final artwork: Donnarumma has deliberately invisibilised the labour of the AI-agent. There is no visualisation of the biosignals that are being input into the musical-AI-agent. The mappings between Donnarumma's actions and the sound output are not immediately clear, likewise for the correlations between the bodily movements and the lighting. It is only through the evolution of the *Corpus Nil* that we begin to understand that there is some kind of connection between Donnarumma's labour and the AI-agent's interpretation of this to mapping sound and lighting control.

The visual consequence of Donnarumma's labour is foregrounded: this choreography appears quite strenuous to perform, which we recognise in the tension visible in Donnarumma's muscles as they strain and twist into positions that seems to challenge the audience's understanding of the constraints of human bio-mechanics. This choreography—the labour—that Donnarumma is engaging with perhaps reflects on societal implication and conceptual values around the labour of human data production (Frank et al. 2019; Newlands, 2021). Indeed, Donnarumma's labour is a core material of this human–AI interaction On a fundamental level, there is a requirement for Donnarumma's human labour for the AI-agent to have material to work with. This in turn establishes a power relationship between Donnarumma and the AI-agent. We observe again, Donnarumma has made a pre-performance choice here for his biodata to be used as a core material within *Corpus Nil*. This in turn also dynamically shifts the power relationship at play within the performance, implicated by the decisions made pre-performance. Donnarumma has made a choice here for his movement and choreography to be the main way of engaging with the musical-AI-agent's computational processes.

Further connecting Haraway's labour dimension to the interaction we see unfolding in *Corpus Nil*, we draw attention to Donnarumma's selective use of visibilisation and invisibilisation. Specifically, we draw attention to a pre-performance choice that has been made by Donnarumma: we can only witness what the artist wishes us to see. We can only see Donnarumma when the "right" combination of biosignals are parsed to and reacted upon by the AI-agent. Throughout the performance the visbilisation of human labour is selective, mediated by computational processes.

How does the nature of the interaction, and the consequent framing of the AI-agent enact or expresses values around control and/or power?

On a micro-level, the interaction between Donnarumma and the musical-AI-agent in *Corpus Nil* is all centred on the ambiguity of which party is in control of what at any given moment in time. As we have discussed earlier, Donnarumma has exerted a degree of control in deciding the fundamental procedure for generating sound and the input mechanisms for how his biosignals are to be fed into the algorithmic processes of the AI-agent. This, as we have previously discussed, constrains the enacted agency of the musical-AI-agent with regards to the kinds of actions they can take. We may interpret Donnarumma's choice here as an exertion of power, in that Donnarumma has established the confines of the potential environment–the situated agency–that may exist within the immediate world of the performance. He has established the conventions governing the actions that the musical-AI-agent may take within the performance. However, this in turn constrains Donnarumma as a performer, in that his pre-performance choices establish how he may engage, provoke and interact with the AI-agent to create, manipulate and spatialise sound. Further, the pre-defined control that the musical-AI-agent has within the immediate performance context imbues it with a power: its responses to Donnarumma ultimately defines what we hear and

what we see in *Corpus Nil*. On a macro-level *Corpus Nil* frames power and control as a process of iterative negotiation: if Donnarumma wishes to manipulate the frequencies of the oscillator clusters then he must physically provoke the AI-agent to trigger its algorithmic responses.

Case Study 2: *The Zizi Show* by Jake Elwes. *The Zizi Show* (2020) (https://zizi.ai/)[1] is a web-based performance installation by Jake Elwes (2020) (see Fig. 3 for still image from performance). It is an interactive Cabaret show, where the audience curates their own performance experience. The audience selects the music to be performed, and the drag artist they wish to perform the song in question.

To provide a sonic and visual explanation for this artwork, the Zizi Show launches in the browser in front of a red theatre curtain. The opening act is *Me the Drag Queen* performing 'Wilkommen' (voiced by Joel Grey), which is also the opening song from 1966 musical 'Cabaret'. *Me* enters the stage, lip-syncing and dancing as they morph and shift continuously between the cast of drag artists available for selection within the show. After the entrance music, the voice of Elwes welcomes the audience and provides a brief introduction to the show. The entr'acte concludes with the ensemble reprise of 'Wilkommen', as *Me* is presented in triplicate (there are three instances of *Me* onstage) who dance and lip sync, morphing once more between the drag artist cast. The main 'hub' of the performance is akin to a waiting room. *Me* the Drag Queen occupies the centre pane of the browser window, dynamically posing and shifting their weight as they 'wait' for the audience member to select a song and for the act to begin. On the leftmost pane of the browser window are a series of clickable controls: pause, rewind, skip ahead and full-screen mode. The rightmost pane presents the cast of available drag artists to select for performance. The eight musical acts available for selection include: 'Raise Your Glass' sung by Pink and movements by Lilly Snatchdragon; 'Nancy Boy' sung by Placebo and movements by Ruby Wednesday; 'This Is My Life' sung by Shirley Bassey and movements by Me; 'You Make Me Feel' by Sylvester and movements by Chiyo; 'Five Years' sung by David Bowie and movements by Ruby Wednesday; 'Sweet Dreams' (the live recording) sung by Beyoncé and movements by Cara Melle; 'Freedom!'90' sung by George Michael and movements by Mark Anthony; and 'I Am What I Am' from the musical La Cage Aux Folles and movements by Me. Once a song and drag artist is selected for performance, the act begins. During the act, there is additional customisability in that the viewer can switch between performers mid-song (either randomly or through manually re-selecting their chosen performer during the song), being able to zoom-in or zoom-out on the performer, as well as to prematurely end the song.

There are multiple neural networks in the AI architecture layers involved in the execution of *the Zizi Show*. For the generation of the drag artists presented throughout the Zizi Show, Elwes has trained a neural network (NN) on a corpus of video frames (static images) which have been collated by Elwes themself. The frame input into the NN is provided with supplementary data, which is primarily skeletal (or pose) tracking, silhouette estimation and optical flow prediction. These are used as conditioning input in a generative adversarial network (GAN) which constructs the drag artists' bodies. Elwes utilises the OpenPose system (Cao et al. 2019) to train conditioning data for

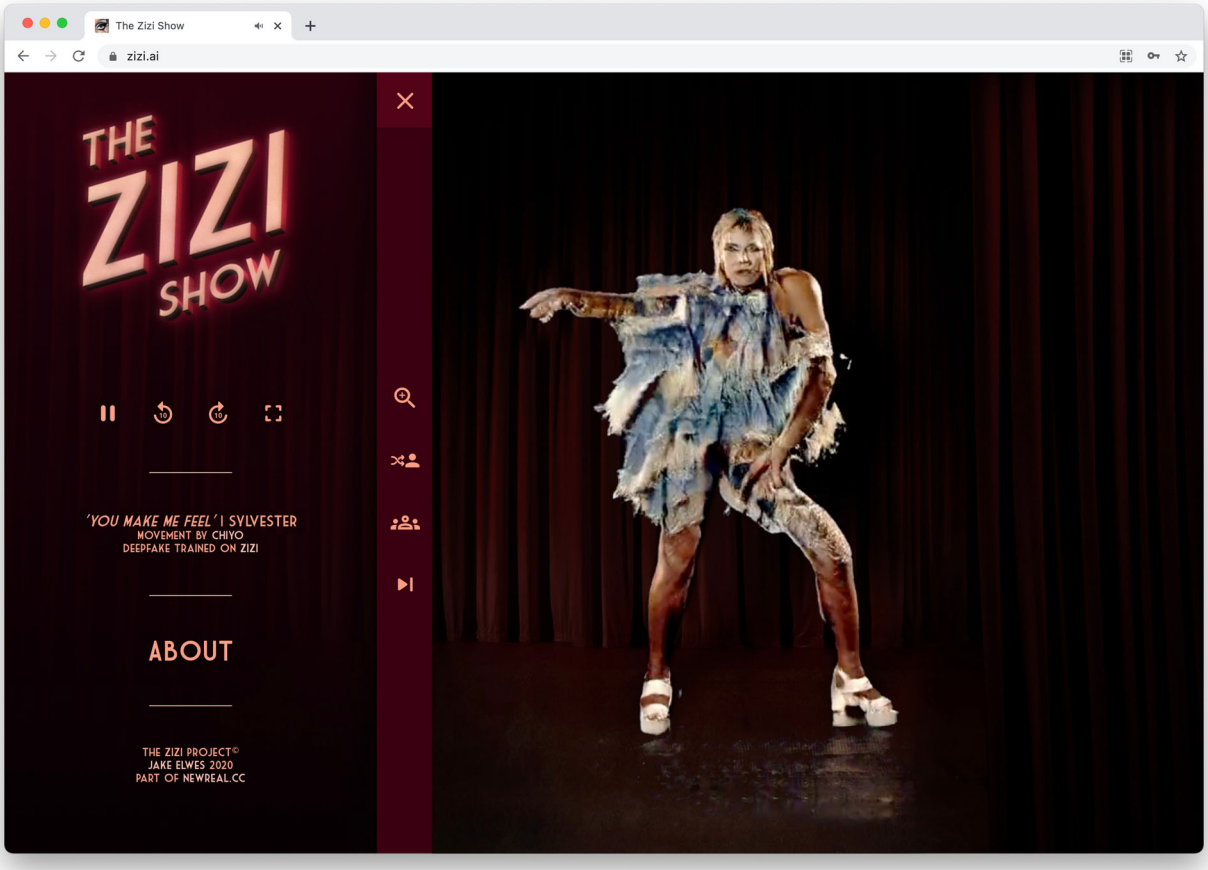

Fig. 3 A screen capture from the Zizi Show, depicting the GUI of the performance. Image provided by, and approved for usage by the artist Jake Elwes.

detection of poses. This incorporates body, hand, foot and facial points in the video frame data corpus. Further usage of silhouette estimation, via the DensePose Recurrent Convolutional Neural Network (DensePose-RCNN) system (Güler et al. 2018), provides additional data. Finally, the usage of the FlowNet architecture (Ilg et al. 2017) predicts the optical flow of objects within static image data: primarily the transition or transformation of the drag artists' movements between frames. The combination of the video image data and the conditioning input of skeletal, silhouette and optical flow, is treated as input data for the GAN.

Each song available for selection by the audience in *The Zizi Show* has a base artist. When the audience selects an artist to perform, the model generates the actions of the audience-chosen artist based on the model's generation of the base artists movement. In the instance of one of the performers—Zizi—their performance within the show is achieved through Video-to-Video synthesis (T.-C. Wang et al. 2018) of the other 12 artists 'cast' in *The Zizi Show*. 13 performers in total are available for selection. One of the performers, Zizi, is a Deepfake artist, trained on the accumulative data of the other 12 drag artists. As described by Elwes (2020), this project deliberately engages with notions of bias and the queering of datasets (Fosch Villaronga and Malgieri, 2024; Parslow, 2023).

How is the AI-agent positioned within the interaction with the human performer (in regards to the sonic space and/or the physical environment?)

The unique setting of *the Zizi Show* warps the interaction experience between AI-agent and performer: the human performers (the real-life drag artists) are absent in the real-time experience both sonically and in regards to the physical environment. The web-browser setting of *the Zizi Show* stands alone as a performance experience. As an audience, we experience it in a similar fashion as a film, albeit with the capacity to choose what we want to experience within the larger structure of the show. In that respect, we could understand the framing of *the Zizi Show* as establishing a constraint for the musical-AI-agent. Its enacted agency is constrained by the situated agency of the web-based performance format.

When we further examine the browser setting as a boundary between human and AI-agent, we can see that Elwes' compositional choice renders the human performer(s) a historical contributor of the data that the architecture has been trained on. This consequently forms a *matter of concern*: the AI-agents are virtually positioned to serve as a vehicle through which the digital memory (data) of the human performer is re-contextualised, reconfigured and re-realised. Here, we might contextualise this positioning of the AI-agents as a flux between an enacted agency and a situated agency. Here, we emphasise the situated agency as a consequence of Elwes' compositional choice on the web-browser format. We acknowledge that this choice was motivated by the immediate socio-cultural context of the artwork, which was created and published during the Covid-19 pandemic. Nevertheless, this compositional choice positions the AI-agents both sonically and physically to become morphological mediators of the drag artists. We see this reflected directly in the functional role of the AI-agent in their dancing and lip-syncing.

The glitching body of the AI-agent (quite literally) re-interprets the physical world and actions of the human performer(s). This glitch is also significant. Although a practical outcome of the Video-to-Video synthesis, the glitchy aesthetic of the AI-agent's generation of performing bodies is understood in two lights. Firstly, that it is reflective of Elwes' intention to "disrupt the dataset" (Elwes, 2020). Secondly, that it is representative of the AI-agent's capacity to re-interpret human bodies. This framing of the AI-agent as a flawed, glitchy mediator of human drag performance is notably contrasted with 'perfect' audio. The audio tracks featured in the performance do not share the glitching aesthetic of the visuals.

What implications (social, cultural, and/or political) can be understood from this positioning (in regards to the nature of the interaction)?

Considering the implications of the positioning of the AI-agent in *the Zizi Show*, we observe politically oriented commentary being made in regard to the failings of AI. The socio-cultural and political implications of this are plainly apparent here, in the failings of digital systems to re-present fleshy human bodies. We can understand Elwes utilises the failures of the technology to make a political statement on how AI-agents such as deepfakes approximate queer and non-binary bodies (Blas and Cárdenas, 2013; De Gregorio, 2023; Greene, 2016; Shah, 2023; Turtle, 2022). Despite the human performers' temporal absence from the interaction, we do experience a degree of the presence, style, aesthetic and 'spirit' of the human performer through their AI-agents' reincarnation. This therefore positions the AI-agents in quite a novel position when it comes to the narrative that the performance event as a whole takes. The drag artists' (AI) bodies are the ones we watch; their movements and lip-syncing is what we see; their re-realisation of an organic form is the lens through which we look. We consider this choice of Elwes as establishing politically coloured commentary, namely in relation to the issue of deepfakes. Within *the Zizi Show*, deepfakes are engaged within a playfully provocative manner: the technological failings of the AI-agents to reconstruct the drag artists' bodies are embraced by *Me* as emblematic of the human artists' "disrupting the dataset" (Elwes, 2020). We note here the tongue-in-cheek irony of Lilly Snatchdragon's dancing body pixelate and morph into physiologically impossible human movements as Pink's lyrics for 'Raise Your Glass' implore:

So raise your glass if you are wrong
In all the right ways
All my underdogs
We will never be, never be, anything but loud
And nitty gritty dirty little freaks

The poetics of AI-agent failure in *the Zizi Show* illustrate the shortcomings of models when they are accustomed to hetero-normative or cisnormative data (Ansara and Berger, 2016; Bauer et al., 2009; McAra-Hunter, 2024), further connecting to Haraway's bodily organic and labour dimensions (Dumit, 2014; Haraway, 2018). With regards to the bodily-organic, Elwes' "queering the dataset" through *the Zizi Show* consciously celebrates the deepfake failures, reclaiming them as reflective of the fluid identities of drag artists (Fosch Villaronga and Malgieri, 2024; Haraway, 2006). We, as an audience, experience the human performers' continual states of 'becoming' through watching the Deepfake artists continual bodily reformations (Grosz, 1994). The AI-agents' glitch is especially pertinent within the context of physical presentations of queer bodies, which do not conform to narrow and hetero-normative conceptions of what a body is (Grosz, 1994). We can plainly see the AI-agents struggling to generate a performer, mashing together movements and expressions into a pixelated human-esque assemblage (Haraway, 2006). Whilst this could be construed as more-than-human (Giaccardi and Redström, 2020; Nicenboim et al. 2020) drag, Elwes plainly embraces the 'misalignment' or 'misunderstanding' that is occurring here at the entanglement between synthetic and organic forms (Haraway, 2006). Essentially, Elwes harnesses and 'troubles' the AI glitch by transforming it into a medium of political and socio-cultural commentary, by challenging the socio-cultural assumptions of AI as "all-powerful". Indeed, Elwes prompts us to question how powerful it can be if it cannot understand a non-normative, queer, body.

How does the nature of the interaction, and the consequent framing of the AI-agent enact or expresses values around control and/or power?

Examining the framing of the interaction between the AI-agent and (temporally displaced) human performer within *the Zizi Show*, the AI-agents are entirely responsible for the construction of the performance world of *the Zizi Show*, and largely operate somewhat independently (barring the audience interaction component). In this regard, the AI-agents are endowed with significant control over the narrative of *the Zizi Show*, which we position as a *matter of concern*. This AI-agent controlled narrative is a marked departure from conventions of drag performance, which tread a fine-line between imitation (lip-syncing as a chief example) and live-ness (through real-time response and engagement with the audience and other performers). The 'packaging' of the Zizi Show as a choose-your-own-AI-drag-adventure situates the AI-agent's capacity to enact agency within a situated context that Elwes has determined ahead of time.

There is an inherent notion of Elwes "taking back power" in their playful exposure of glitching deepfakes. The visually dominant errors of the AI-agents morphing the drag artists' bodies into strange, pixelated hybrid morphologies and even the underlying skeletal tracking algorithms powering the deepfakes' movements. This is juxtaposed by the jarring audio clarity of the songs available for selection, which are unmarred by the visual glitch of the performer's lip-syncing to them. This embracing of technological failure is notably at odds with theatrical conventions of theatre and Cabaret, which emphasise a sort of performative 'perfection': "the show must go on". The sound throughout the performance experience emphasises this performative perfection, remaining solid, intact and devoid of glitches in juxtaposition to the visuals.

Turning now to the labour dimension, the engagement of deepfakes of drag artists and the unique performance-world setting unfolds a variety of possible perspectives towards labour. Firstly, this takes place in the agency of the audience as a curator of their own show. From the sumptuous stage of the waiting room, *Me the Drag Queen* advises the digital audience of their role in selecting the songs they wish to listen to. Here, there is a subversion of the expectations of performer-labour versus audience-labour. This breaks the Fourth Wall (Stichter, 2016) and places an increased responsibility upon the audience to 'programme' their own experience. The audience is "in control" of the digital drag bodies, they decide who performs, what is performed, and when to give performers the digital "hook" (Slide, 2012). Yet this "control" is failing upon multiple levels: our experience in 'choosing our own AI drag-adventure' is at the mercy of glitchy AI-agents.

Additionally, the significant labour of the human cast behind the novel digital experience of the Zizi Show cast must be acknowledged (Elwes, 2020). Taking into account the various layers of labour with *the Zizi Show*, there are multiple intersecting layers of 'failing labour'. The audience is not in control of the performance and how it will look to them, but the AI-agents are not in total control either. They continually fail and glitches in their reconstructing of the bodies of the drag artists. Neither is the Deepfake artist in control, in that they can be swapped in or out by the audience. Similarly, the original artist is not in control, emphasised by the impossibility of having them physically present in a browser-based experience. These accumulative layers of failure that are constructed and presented within *the Zizi Show* illustrate a nuanced commentary on the AI-agents' understandings of fleshy human bodies.

Case Study 3: *"the machine is learning"* by Marije Baalman. *The machine is learning* (2020) is an artwork that exposes the process of training a machine learning model within the context of a musical performance (see Fig. 4 for still image from performance). During the 25-min long work (https://www.researchcatalogue.net/view/908792/908793[2]), the performer

Fig. 4 A photo of Baalman performing the machine is learning, depicting the various terminals in which the gestural recognition and gestural coding language windows are prominently visible. Image provided by, and approved for usage by the artist Marije Baalman.

(Baalman) performs a series of repetitive arm movements and gestures whilst directed by a synthetic voice.

We will take a moment now to discuss how this artwork is sonically and visually presented. The performance begins with Baalman standing centred in front of a projector screen, which presents a series of windows comprising a control system. Baalman appears to press something on a wrist-mounted sensor, triggering a welcome message from a disembodied, robotic voice. The voice of *Visi* is somewhat masculine coded (denoted by a lower pitched voice), rather monotonous and produces certain phoneme sounds in a way that sounds incredibly unnatural and synthetic. It welcomes us: "Hello I am Visi, I will help you to learn the machine learning using GeCoLa" and gives a small monologue explaining some of the system components and programming languages that are used in *the machine is learning*. Baalman stands motionless, expressionless, during this monologue and only responds when *Visi* announces: "The performer will now start the performance". When Baalman begins to move, the control system projected behind here is updated with terminal commands. We begin to see that the control system is comprised of a series of interconnected software modules. Along the top of the projection is the first window—a gestural recognition window. This contains a visualisation of the recorded arm gestures that Baalman performs, populated in real-time. In this window, the learning, recording and storage status of the agent is presented alongside the gestures' identification number. The gestural data: accelerometer, gyroscope and magnetometer data are plotted linearly in the second window in the top right panel. This SenseStage terminal window presents the input commands from the wireless sensors Baalman wears on both hands. The third window, centred in the bottom of the projection, is the GeCoLa terminal. GeCoLa is the gestural coding language that Baalman utilises to programme their sound (M. Baalman, 2019). This GeCoLa terminal is populated with the text-to-speech (TTS) transcript of *espeak*. *espeak* is a voice synthesis programme that generates the 'voice' of the musical-AI-agent, which identifies itself as *Visi*. The fourth window is an additional terminal window, which presents the dynamic time-warping training and recording commands.

As for the sound aesthetics of the performance, *espeak* provides the sole sonic output for the first 8 min of the work. As Baalman begins to methodically perform a series of gestures, *Visi* narrates this entire process, questioning whether to store certain samples and confirming that they have been stored. During this spoken sequence, *Visi's* voice remains perfunctory and robotic. It is only when Baalman completes the gestural input for gesture 1 that we hear *Visi* introduce some sonic variance, in which *Visi* repeats the word "new" with a progressively pitched down effect. This sonic pattern: *Visi* narrating the training process, announcing the model training and then performing a small sonic gesture (mainly the introduction of a new vocabulary word and a sound effect) forms the main sonic organisational structure of the piece. Regarding the more "musical" sounds, which we separate momentarily here from *Visi's* narration, there is a relatively small bank–three in total that we were able to determine–of triggered sounds that are utilised in the performance. One is a hollow, low-toned percussive hit with a slight reverb applied. The tone is reminiscent of a low-pitched gong. Another is a higher-toned percussive hit without reverb applied, sounding similar to a high-hat with added white noise. The third is a higher-pitched, synthetic and square-wave style beep, sounding similar to a digital metronome sound. There is also a bank of gestures that are more control-oriented. These add new gestures add additional beats to the currently playing phrases, remove beats, remove phrases, as well as speed up and slow down phrases.

We will take a moment here to explain the technical components utilised within this artwork. Baalman wears a wrist-mounted 9 degrees of freedom (DOF) wireless motion sensor (3-axis accelerometer, 3-axis gyroscope, 3-axis magnetometer) to track their gestures. The accelerometer, gyroscope and magnetometer data is sent into SenseStage software (infrastructure for wireless sensing) (Baalman et al. 2010), which then translates this gestural data into OpenSoundControl (OSC) commands (Fraietta, 2008; Wright, 2005). These OSC commands are then sent into SuperCollider (Wilson et al. 2011), where they are further formatted for use within a C++ programme, which is based on Nick Gillian's Gesture Recognition Toolkit (Gillian and Paradiso, 2017). During the process of building a dataset for a dynamic time-warping (DTW) algorithm, Baalman periodically tests the model by performing the gestures she has trained. Depending on the accuracy of the gesture recognition, this triggers a collection of sonic and structural events throughout the course of the performance.

How is the AI-agent positioned within the interaction with the human performer (in regards to the sonic space and/or the physical environment?)

The human–technology relation between Baalman and the AI-agent in *the machine is learning* is quite literal in its presentation. Baalman's interaction with the model is deeply physical and highly performative, whilst the model's response is highly formulaic, and sound-centric. There is a tight coupling—and the occasional system failure—between human physical input and the sonic space that is generated by the AI-agent. We understand that this artwork example provides a subversion of an expected usage of an AI-agent within a musical performance: instead Baalman's enacted agency becomes constrained by the recognition capacities of the AI-agent she has developed for the performance. (Grosz, 1994) This subversion is emphasised in Baalman's compositional and artistic choices. We note here that this artwork is thoroughly composed in terms of its structure, the technological engagement and the choices made with regard to the sounds that are featured and the visualisation of Baalman's software pipeline.

On the positioning of the AI-agent to Baalman (and vice-versa) in the sonic space, we note that Baalman has deliberately chosen to remain silent during the entire performance. Only her heavier, presumably frustrated, exhalations contribute to the soundworld that the AI-agent triggers in response to her actions. Essentially, the musical-AI-agent's voice is the only one the audience hears, and as a speaker it dominates the entire sonic space alongside the triggered sonic events. As we become progressively attuned to the sonic results of Baalman's actions, we become less interested in her physical actions themselves. Instead, we become more preoccupied with the various text updates, the refreshing terminals and the addition of the waveforms displaying the accelerometer, gyroscope and magnetometer data appearing on the projection. Thus, the sonic dominance of the AI-agent further established its dominance in the physical performance environment. This is achieved by Baalman foregrounding the voice of the AI-agent—and indeed the visible scope of its computational processes—as the more intended object for our attention. This is periodically disrupted when the training is completed and Baalman tests the AI-agent. When these moments occur, our attention shifts back to her choreography.

Through the framing of the AI-agent as the more sonically and visually dominant point-of-focus, we can identify a similar negotiation of situated and enacted agencies as in Donnarumma's case study. Baalman's input is crucial to the training of the model, and in that respect we can understand that she has intentionally constrained the potential actions that the AI-agent can take in mapping her identified gestures to sonic events. She also has a

more clearly defined stake in determining what gestures are stored (she presses buttons on her sensors to confirm or reject the storage of samples), and the progression of the performance is entirely at her own physical pace. Yet there is somehow a dominance to *Visi's* disembodied presence, which is implied visually through the prominent visuals of the terminal windows and its sonic prominence as the only sound 'source' in this artwork. Baalman is thus progressively positioned in a more peripheral role in terms of the physical environment, even though she is spatially placed closer to the audience than the software interface is.

What implications (social, cultural, and/or political) can be understood from this positioning (in regards to the nature of the interaction)?

Within the curious and shifting positioning of the AI-agent (and Baalman herself) in *the machine is learning*, we see a subversion of typical power dynamics that we might assume to exist between a human musician and AI-agent. This is immensely curious, as we note that Baalman's physical actions are crucial to the successful training of the AI-agent and the generation of new sound material. Connecting the implications of these complex hybrid relations to Haraway, we see commentary made regarding the technological dimensions. Within *the machine is learning*, this is concentrated on notions of machine error and recognition of human bodies. Within the performance, there are multiple instances of failure in the recognition of Baalman's gestures which are indicated through Baalman's visible frustration at the inconsistencies of *Visi* (the musical-AI-agent) in recognising particular gestures. The inability of *Visi* to reliably identify Baalman's movements is clearly evident in Baalman's frustrated expressions and huffing breaths.

From an alternative perspective, the mise-en-scène of the performance also emphasises the visualised presence of the technology. Its 'body' is immense, continually updating, and visually demanding of our attention. Such an emphasis on the visualisation of the technological processes involved in this artwork illustrates a sort of dominance of technology over the human body performing. Considering this in relation to Haraway's labour dimension, we may understand that Baalman is making commentary on the nature of labour and dataset construction processes for machine learning. Although choreographically simple, Baalman's physical labour in this artwork is substantial in terms of its physicality: she utilises wide arm movements, with a large range of motion. This exaggerated physicality serves both a functional purpose, in that the more expansive and clearly identifiable the gesture is- the easier it is for *Visi's* recognition. We have noted previously that Baalman frequently 'resets' her body (shaking and rolling out her shoulders), which implies that her precise execution of these gestures is somewhat physically demanding. We might understand this as either a commentary upon the need for homogeneity when it comes to data for machine learning, or that Baalman is exercising physical restraint to ensure that her movements conform to the input needs and competencies of *Visi*. We acknowledge here that Baalman has herself had some degree of control over the type of sensor and thus, the type of data that *Visi* receives as input.

How does the nature of the interaction, and the consequent framing of the AI-agent enact or expresses values around control and/or power?

Although there is an obvious correlation in regards to the individual roles occupied by both Baalman and the AI-agent, there is an unevenly distributed component of power and control within the interaction. We see this most saliently expressed through the physical use of and dependency upon labour as a medium for control. There is a clear and fundamental necessity for Baalman's real-time and situated engagement with the AI-agent. Specifically, Baalman's performative labour of training the AI-agent through building the dataset of gestures ties into the bodily-organic and labour dimensions from Haraway (2018).

In regards to the bodily-organic, Baalman's methodical and repetitive routine positions the AI-agent as an entity that is being provided *with* information, at a physical cost to the provider of information (the performer). We see and hear this cost in the micro-interactions between Baalman and the AI-agent: Baalman's heavy exhales and sighs when the recognition of particular gestures fails; how she shakes out her body in preparation for precisely repeating a gesture. We might consider that an imbalance is being illustrated here: the infallibility of the human body compromises the competency of the AI-agent's gestural recognition (Grosz, 1994). Or conversely, the AI-agent is unable to adequately interpret and map the human body's input. Through this error-success interaction loop, Baalman becomes somewhat demarcated as a performing body: her body is a working body. We do not see it as coincidental that Baalman is clothed in trades-person overalls in a highly visible orange colour.

Continuing on the bodily-organic implications, we should also consider the physicality or visual style of Baalman's movements: they are stylistically neutral yet highly specific and easily reproduced arm gestures. We cannot draw any discernible connotations to the sonic functionality of the gestures themselves. Although the narration of the model ("plus", "minus") gives us some clues as to the influence of the gestural recognition upon the sound outcome, Baalman has not chosen a more literal physical gesture (such as raising or lowering her arms to indicate "plus", "minus" respectively). This renders the physical action onstage somewhat mundane. This "mundane-ing" positions our focus on the performer as an 'object', to be recognised by the model based on the consistency of the performed actions. Indirectly, this communicates a perspective on the physical navigation of the interaction loop: that the labour of the performer, although crucial for the triggering of sound events, is otherwise a peripheral or background 'item' and therefore should not be afforded significant audience attention. The mundane-ness of movement is especially noticeable when we also consider the sonic space of the performance. In this regard, we clearly see the framing of the AI-agent in *the machine is learning* as enacting agency as to what gestures are initially of *use* or recognisable-enough. Baalman must then agree or disagree with the model's assessment. The particularities of this mundane-ing and recognition of gesture–of labour—by the AI-agent enacts understanding as to the power of the AI-agent in this interaction. This in turn creates a *matter of concern*: what happens if this mundane-ing of labour by AI-agents is permitted to continue without human oversight? (Burgess, 2021; Eslami et al. 2017; Hampton, 2021; Nowotny, 2021).

Case Study 4: Shimon the musical robot. The specific case study we utilise for the analysis of Shimon is a performance at TivoliVredenburg in Utrecht, Belgium on September 26, 2018 (see Fig. 5 for still image from performance). The documentation of the performance, broadcast by VPRO Vrije Geluiden, is available at https://youtu.be/kG16-yIE8Sw. Shimon is a robotic marimba player, performing in a jazz concert alongside 6 human performers. To introduce how this performance unfolds, the 28 min long concert takes place in a theatre with the band onstage. The concert features a typical jazz line-up with a rhythm section and horns. The ensemble onstage is comprised of alto saxophone (Richard Savery); electronic bass (Gil Weinberg); guitar (Zach Kondak); drums (player's name not provided on performance documentation); keys (player's name not provided on

Fig. 5 A photo of Shimon performing in the Shimon and Friends Concert at TivoliVredenburg Utrecht. Image provided by and approved for usage by VPRO Vrije Geluiden.

performance documentation); trombone (Jeroen Verberne); trumpet/vocals (Dash Smith); and trumpet (Gerben Klein Willink). The setlist features Gil Weinberg's 'Bafana' and 'Iltur 3'; 'Blue Monk' by Thelonious Monk; 'Oh Oh' by Zachary Kondak; 'AI' by Gil Weinberg and Dashille 'Dash' Smith; and a demonstration of Shimon performing a newly composed Hiphop song with Dash Smith. Stylistically, the setlist encompasses bebop, hip-hop, and free.

We will now briefly discuss the software system that facilitates Shimon's performance. The control system behind Shimon's musical engagement is built from a series of interaction modules. These facilitate Shimon's ability to listen, keep time, move their head and limbs, and generate musical material in response to human musical partner(s) (Weinberg et al. 2020b). A central infrastructure interfacing with the interaction modules facilitates Shimon's receiving of audio input from fellow players, and manages Shimon's musical actions. As reported in Hoffman and Weinberg (2010), a MIDI input is utilised to receive note input from a MIDI-enabled keyboard. A universal *Beat Keeper* module dynamically adjusts the metronome coordinating Shimon's musical responses. Finally, Shimon's system utilises three differing chord representations: a fixed set within Shimon's playable range (presumably, as determined by its actuated percussion strikes and limb coordination); an octave-agnostic set comprising a set of notes with octave harmonics; and representative base notes with set-octave or octave-agnostic harmonics. Utilising the central infrastructure outlined above, there are three key modules that mediate Shimon's interactions in musical settings: a call-and-response interaction module; an opportunistic overlay improvisation module; and a rhythmic phrase matching improvisation module (Hoffman and Weinberg, 2010).

Module 1—the call-and-response interaction—utilises beat tracking (typically of the bass line performed by a human keyboard player) and gesture anticipation. Module 2—the opportunistic overlay improvisation module—is largely responsible for Shimon's gesticulations during performance. The choreographed actions of Shimon (head nodding, as an example) are actuated based on a beat detection percept, which tracks the beat of the music based on a bass line performed by a human musician on the MIDI keyboard. Within Module 3—rhythmic phrase matching improvisation module—Markov Chain analysis is utilised to generate Shimon's musical responses, based on the human keyboardist's note input. The improvisation module is trained on performances by John Coltrane. Shimon's pitch and durational material is modelled in second-order Markov chains (Bretan and Weinberg, 2016). This Markov Chain analysis generates Shimon's musical material so that it is harmonically coherent with the keyboardist's provided material. Within this generation process, operations such as transposition, inversion of material and delays are utilised to introduce musical novelty.

How is the AI-agent positioned within the interaction with the human performer (in regards to the sonic space and/or the physical environment?))

Within the interaction, Shimon is prominently featured within the physical environment and occupies a central position in the ensemble formation onstage. Shimon is actively contributing to the sonic space: it is collaborative in its musical involvement throughout the entirety of the performance, 'takes' a number of solos and interacts with the melodic material presented by the other musicians. As Shimon is programmed to specifically utilise the material from the keyboardist, we hear it quoting, slightly transforming and occasionally extending their melodic components and phrasings. Throughout the first piece–Bafana–Shimon's musical contribution continually focuses on a 'lick' (a main melodic phrase or motif) that the keyboardist establishes. There are moments where Shimon does begin to become more rhythmically adventurous with the lick, but it largely appears to stay within the particular register, rhythmical structure and keeps the same melodic pattern as established by the key player. However, this focus on the keyboard material notably shifts in the second piece on the setlist, where Shimon takes on a more 'independent' role when it establishes the tempo and groove of Iltur 3. Here, Shimon initiates the transition into Iltur 3 with a slightly syncopated descending bass line figure. The horns are more sonically prominent, and the solos appear more spread around to the other band members. Shimon takes a number of small solos in Iltur 3, and gives the impression of cuing solos from the other players by swivelling its head towards individual players. This combination of the physical and sonic positioning of Shimon establishes that it is contributing to the scene on a fundamental

level and that it—as a musical entity—is actively engaged in the act of performance with the human musicians. We can see obvious physical signifiers of Shimon's engagement illustrated through the programmed behavioural responses. It bops its head to the beat and the circular head rolls and swivels to 'direct attention' to signify listening to the human musicians.

What implications (social, cultural, and/or political) can be understood from this positioning (in regards to the nature of the interaction)?

The interaction on-stage indicates a number of implications pertaining to (the music) performance as a social and deeply collaborative activity. Shimon's central physical and musical presence in the interaction presents us with a utopic view of human–machine musicianship. Human musicians navigate a sonic space alongside a machine, collaboratively. However, there is an obvious juxtaposition in Shimon's form and physicality in comparison to their fellow musicians. It is obviously not designed to resemble a human morphology, as is evident by its multiple limbs. This simultaneously Other-s Shimon as it foregrounds it. The visual aesthetic of Shimon establishes a highly industrialised visual connotation: it is not difficult to imagine Shimon within a manufacturing setting. Its more-than-human form, with multiple limbs, further emphasises an impression of it as a high-precision artefact. The visual connotations implied by its highly industrial aesthetic are further emphasised by the anthropomorphic coding of its gestural behaviours. Specifically, its head movements are smoother, continuous and at a slower pace than might be expected from something resembling an industrial robot. One could even describe it as elegant. But this industrialised aesthetic invites socio-cultural implications, mainly in regard to Haraway's labour dimension. Shimon is not visually dissimilar to the same robots utilised within industrial manufacturing contexts which perform sometimes hazardous labour tasks. Such a reading of Shimon therefore creates an interesting tension in its presence in a jazz performance—with the historical roots of jazz music (and its early connections to creative expression and culture from victims of the Atlantic slave trade). One reading of this is the juxtaposition of the 'freedom' of expression (and the artistic labour) that Shimon is demonstrating as a robot playing jazz, despite the robot's highly industrialised appearance.

Continuing further in the implications of Shimon's morphology, although this concert setting appears concentrated towards showcasing Shimon's capacity in performing at a "human-like", or indeed "more-than-human"—like levels of competency, the Otherness of Shimon and the subsequent pulling of focus admittedly carries deeper cultural and political implications around inclusion and access. We could understand this in one of two ways: either, that inclusion of robotic AI-musicianship within a human music-making context necessitates 'more-than-human' morphology, or alternatively, that anthropomorphism of peripheral music-making behaviours is crucial to convincingly situate AI-musicianship on an 'equal' level with human musicianship. Naturally, this invites a hierarchy: for inclusion to occur, there must be a surpassing of human ability or a human-esque behavioural coding.

How does the nature of the interaction, and the consequent framing of the AI-agent enact or expresses values around control and/or power?

The presentation of Shimon, its presence and physical interaction as well as sonic engagement with the performance all clearly point towards it having a significant degree of control or power in the performance. The musical role it has throughout the setlist is plainly obvious. Shimon is a musical centre-point-of the line-up and takes quite a leading role in prompting the band members to take solos; it takes a great many solos throughout the concert and it even establishes the tempo and groove of a number of pieces. We note that this is clearly motivated by the context of this performance, which is plainly to showcase the technical sophistication of Shimon as a robotic—and musical—artefact.

Discussion

Our engagement with Post-phenomenology and Haraway's Implosion wrought extensive and rich insights into the implicit in the interaction between the human performers and musical-AI-agents across our four case studies, demonstrating the importance of looking beyond how a technology *is used* in an interaction. In this section, we discuss the results from our analysis, and their relation to our research questions posed in our Introduction section.

Across all four of our case studies, we have seen the artists' alternative approaches to framing the interaction between themselves and the AI-agent. One salient finding is that these framings have not been static, and have continually been reconfigured throughout the course of each artwork's performance. We see this in Donnarumma's *Corpus Nil* engagement with the AI-agent's decision-making to provoke and elicit different sound interactions and trigger different lighting events. Donnarumma himself describes this: *"As the body parts shift to a different position and acquire a new function, another kind of reorganisation takes place: the reorganisation of the instrument's algorithms"* (Donnarumma, 2016). Further, Donnarumma's pre-performance choices around the sound aesthetic and the sound mapping constraints establish a continual negotiation between himself and the musical-AI-agent. Donnarumma has constrained the available actions that the AI-agent is able to take within the various dimensions of the interaction, and yet he is in a continual state of trying to provoke the AI-agent to respond in different ways. Likewise in *the Zizi Show*, the framing of the organic and the synthetic as temporally disconnected achieves a reorganisation of human bodies with the AI-agent as a mediator. We see this most potently in the juxtaposition of the dynamically re-forming and glitching drag artists' 'bodies' against the 'untouched' audio. In *the machine is learning*, we bear witness to Baalman training a model–a process that is normally completely invisibilised to audience. We watch the physical negotiations that Baalman must engage with to ensure that *Visi* has adequate and appropriate training data. In Shimon's performance with its human peers, we see the harnessing of its presence as a performer, and the impact of its presentation as a performing Object to navigate real-time interaction in a jazz performance.

Dynamics of (in)visibilisation. In the 'how' of interaction between organic and synthetic matter, we observe that visibilisation (or invisibilisation) is used as a mechanism to 'shape' the organic bodies in relation to the synthetic, and vice-versa. By this, we refer to the (in)visibilisation of certain processes which are integral to the audience's understanding of what is actually happening in the technical aspects in the artwork. Some examples revealed in our case studies include: the invisibility of mappings between organic and synthetic bodies (as in Donnarumma); the visibility of the failure of synthetic bodies to reconstruct organic bodies (as in Elwes); the visibility of the organic labour of training synthetic algorithms (as in Baalman); and the (hyper)visibilisation of an anthropomorphically coded synthetic body that inadvertently invisibilises organic bodies (as in Shimon).

This leveraging of invisibilisation and visibilisation is significant in establishing how boundaries–sometimes fuzzy ones–are actively drawn between the organic and synthetic. As we have understood from our analysis, these fuzzy boundaries are influenced or established through processes of

(a) invisibilising and visibilising the organic and/or synthetic material in the interaction; and/or

(b) visibilising and invisibilising labour and failure.

Significantly, these processes of (in)visibilisation have wider implications. The decisions that artists make to visibilise and invisible certain aspects or components within their artworks establish constraints for both themselves as a performer, and the AI-agents performing with them. As we have illustrated in our analysis of *Corpus Nil*, Donnarumma imposed constraints on how he may interact with the AI-agents as a performer which in turn established a cyborgic entanglement Haraway (2006) between his fleshy human body and the sonic 'body' of darkness of the AI-agent. In *the Zizi Show*, the constraints of the web-browser format created a "choose your own adventure" experience for the audience. Elwes' use of deepfake drag artists was further constrained by their capacity to generate the non-binary bodies of the human drag artists. In *the machine is learning*, Baalman was constrained by *Visi's* ability to accurately identify her physical gestures after being trained upon them. In Shimon's jazz performance, the visibility of the highly industrialised robotic figure emphasised a boundary between Shimon itself and the human jazz musicians it shared a stage with.

These processes of (in)visibilisation ultimately establish dynamic distributions and exchanges of power and control within human–AI musical interactions, but they do not exist in a vacuum. As we have discussed in each of our case studies, the compositional decisions of artists have significantly informed how we—as an audience—understand the unfolding interaction between human and musical-AI-agent. In the case of Donnarumma, his compositional choices regarding the generation of the oscillators and the design of the space and lighting shapes our reception of Donnarumma and the musical-AI-agent as an entangled, cyborgic body (Haraway, 2006). Similarly, in *the Zizi Show*, Elwes' engagement with deepfake models that consistently fail to generate drag artists' movements and gestures without dissolving into extreme visual glitches positions us to view a particular commentary on AI-agents' capacity to synthesise, comprehend and construct non-binary human forms. In Baalman's case, her choice of sensor, the design of the performance space and her choice of choreography positions us to view the interaction as a stilted (and sometimes comedic) exchange between the human and the digital. Whereas in Shimon, we see the impact of using anthropomorphic coding of an industrialised robot to make commentary on collaborative human and machine music-making–as harnessing differing morphologies and physical affordances to create novelty.

Across each of our case studies, we identify that these dynamic compositional choices carry residues. These residues are not just the literal, artistic or musical consequences of how an AI-agent is involved in a musical performance, but the cumulative impact that these constraints have in speaking to implications from the wider world. The way that artists intentionally and unintentionally position the technology in their art is significant. Merely classifying human-AI-agent relationships ignores the intentionality that has been constructed in the relationship and imbues it with additional layers of meaning. As designers, artists, audiences and users of these technologies seeking to accrue and generate knowledge, we have a responsibility to cultivate awareness of how our positioning of an AI-agent carries residues that extend far beyond the roles we give to musical-AI-agents.

Returning to our Research Question 1: how AI is positioned by a human in human–AI musical interactions and what this communicate about power/control within the interaction. Power and control lie in choice, and in the consequences of choice. Throughout our analysis of each of the case studies we have identified that the pre-performance choices made by the artists have significantly shaped and informed how visibly or invisibly the values around power and control are embedded within their artworks. In Donnarumma's example, the ambiguous mapping between human input and musical-AI-agent-mapped output establishes implicit values around the physical ways humans communicate with synthetic AI bodies through exchanges of control. In Elwes' example, glitching deepfakes of drag artists' bodies communicate cultural and political values around the mis-recognition of queer and non-normative bodies by AI technologies—which speaks to the power these models may wield. In Baalman's example, the positioning of human data input against machine error to identify this data implicitly communicates cultural values pertaining to the role of human labour in the data pipeline of AI systems—speaking to notions around control. In the example of Shimon, its industrialised morphology and anthropomorphic coding of movement and behaviours (head bobbing, 'looking' at performers, etc.) implicitly embed cultural values around—and provocations about—music-making between human and robotic agents. As our analysis of each case study further revealed, there can sometimes be significant and evolving differences as to how control and power are enacted, and how they are portrayed. Here, the distinction may reside in the subjective pre-performance choices that artists make when they commence an artistic engagement with a technology. As we have discussed, these pre-performance choices largely concern the relationality between the human and the AI-agent, namely through the juxtaposition of enacted and portrayed power and control (through the control, manipulation or generation of sound, lighting, visual events).

Returning to our Research Question 2: how sociopolitical perspectives are implicitly embedded in the interaction. Within the more general setting of human–AI musical interaction, we see power and control explicitly presented in the artists' subjective choices around what is (in)visibilised in the interaction, and the consequences of this (in)visibilisation. As we have established, these choices are not exclusive to the immediate world of the performance and also constitute choices made in the composition stages of an artwork. This has explicit and implicit connotations about how power and control are presented in two key ways. Firstly, there is an explicit choice made by the artist in positioning the AI-agent in the interaction. This assigns a level of control or power that the AI-agent has within the interaction. Consequently, this also influences the level of power and control that the artist can have upon the AI-agent's actions. We have also observed that this positioning-in-performance is not static and is liable to dynamically change as the performance unfolds. Secondly, there are the implicit consequences of positioning the AI-agent within a particular role in a musical interaction. These consequences encompass the intercontextual connections—or residues—to domains such as the technological, the bodily-organic and labour. We have seen this demonstrated in the implicit communication of social, cultural, and/or political values across our four case studies.

Expanding upon our 'residues' metaphor, we see human–AI musical interaction as further containing implicitly embedded social, cultural and political values. These values are implicitly embedded through the artist's choices about the presentation of the role of the AI-agent through performance, but also in the artistic choices made in the materiality of the artworks. By materiality, we refer to the sound aesthetic, the visuals, and the artwork format and presentation. As an example: if the AI-agent is wholly responsible for the sound or visual aesthetic of the artwork, what does this communicate to us as an audience? In this example, an implicitly embedded value may reflect an assumption on the technological 'perfection' of the AI-agent against an 'imperfect' human. Aesthetics carry residues or

allusions to social, cultural and political phenomena, and the artist's implicit referencing to these phenomena through their aesthetic choices conveys a particular and implicit stance (or stances).

Approaching from a top-down perspective, we now establish the importance of parallel thinking when it comes to the development of, interaction with and composition for artworks utilising musical-AI-agents and systems. Here, we draw parallels to theoretical concepts that we have already grounded this exploration within: Latour's matters of fact and concern.

Matters of fact and concern. As a first principle, *matters of fact* attempt to contextualise an objective fact of a human–technology relation, within the confines of the interaction itself. The lens and scope of the examination is constrained and specific. Although we have established that such an examination craves wider contextualisation, matters of fact attempt to deal solely with fact: what *is*?; what *unfolds*?; and how do we classify this *is-ness*? Latour's call for this objectivity to be "appreciated,…experimented upon, mounted, prepared, put to the test" therefore establishes the first principle of that objectivity as it emerges from the interaction. It is fragile, messy, and invites scrutiny.

This therefore justifies our need to relate the factual to the implications of fact: *matters of concern*. In this study, the movement from 'word to world' took place by applying an iteratively expansive analytical view. Through this lens, we reveal that power and control form a kind of filament that accompanies and is entangled within the sociodigital and its material. These filaments may be unearthed through an artist's particular framing of an AI technology in relation to herself, and its connections to wider sociopolitical contexts.

Conclusion

Within this paper, we have concentrated on interaction with musical-AI-agents in an artwork performance context. As we have established, the engagement and interaction with these agents have wider cultural, social and political implications. We have argued that these cultural, social and political aspects have further implications for power and control, pertaining to how humans and musical-AI-agents mutually interact in this specific setting. As we have delimited the scope of our study to a specific performance context, we have not examined what wider implications there may be outside of the world of the interaction. We instead note this as a critical area for future research.

Discussions around power and control are extensive in the context of algorithms in re-forming musical and social culture (Amershi et al. 2019; Bown, 2011; Bown and Brown, 2018; DeVito et al. 2017; Martin et al. 2020; Siles et al. 2020; Werner, 2020). We see similar challenges within Human–AI musical interaction, and share research concerns about shifting values around human–technology interaction (Birhane et al. 2022; Tatar et al. 2023); the making of meaning (Bourne, 2015; Dahlstedt, 2018); and situated knowledge (Haraway, 1988) and digital pervasiveness in both everyday life and creative practices.

Of additional and critical importance to our future work, is greater diversity in regards to geography and musical cultures of musical-AI research. We see a need for methodologies for how we conduct our research in a manner that is ethically sound, and does not inadvertently embed the values, biases and assumptions of the researcher (Abdelnour-Nocera et al. 2012; McKay et al. 2022; Offenwanger et al. 2021).

We further acknowledge that a concentrated focus on work produced by a certain demographic of those already in privileged positions as producers of culture inadvertently overlook the richness of sub-cultural movements occurring at a grassroots level. We see the mindful and participatory navigation of barriers to access and collaboration, such as linguistic barriers, as an important area for future attention in this field. We acknowledge that there is important work to be done in ensuring that our whole community is represented in the research work we do and that smaller communities and minorities are afforded just and equitable access and opportunity.

As summary, this paper presented a multidisciplinary analysis of four case studies of artistic works featuring human-AI musical interactions. Our approach to analysis drew from theoretical perspectives and methods from Post-phenomenology and Feminist Science and Technology Studies. Our Intermedial Analytical Method was presented in the "Methods" section as a novel method to simultaneously examine the *matters of fact* and the *matters of concern* of human–AI musical interaction. We propose this Intermedial Analytical Method—presented in Fig. 1—as a phase-by-phase procedure for other researchers to implement within their own analyses. This analysis was motivated by an investigation into (1) implicit or explicit presentations of power and control in human-AI musical interaction, and (2) how sociopolitical perspectives are implicitly embedded into the interaction. In the "Analysis" section, we presented our observations and analytical findings derived from our intermedial analytical method. In the "Discussion" section, we examine the emergent meta-findings from our intermedial analytical method. In the "Discussion" section, we further identified processes of visibilisation and invisibilisation to be central mechanisms for shaping:

a) the interaction of organic (humans) and synthetic (AI-agents);

b) the drawing of boundaries between the organic and synthetic in practice and knowledge-making;

c) how are sociopolitical values implicitly embedded in the interaction to present notions of power and control

These three mechanisms may contribute invitations for artists, developers and designers of musical-AI agents and systems to more critically consider their utilisation and development of these technologies for creative AI research.

To conclude, our utilisation of multidisciplinary perspectives afforded the examination of the facts and concerns of human–AI musical interaction. It further enabled an examination of the connections and entangled filaments between them and made visible previously invisibilised aspects of human–AI musical interaction. This process of visualisation reveals a diversity of value positionings related to the negotiation of power and control in human–AI musical interactions. The recognition and analysis of these interconnected facts and concerns reveals the layered embeddedness AI technology has within sociopolitical contexts. As the tide of novel engagement with creative AI technologies continues to rise, it becomes imperative for us as creative AI developers and users to critically examine how we engage with these technologies in our art and the contributions our engagement may make to the birth or advancement of new creative AI art movements.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 12 October 2023; Accepted: 7 February 2025;
Published online: 03 June 2025

Notes

1 Please note that, at the time of writing, this artwork was openly and publicly accessible. Presently, however, it is not possible to access the artwork due to music licensing. It is possible to email: zizidragshow@gmail.com to request access.

2 Select the 'Performances' tab in the browser and scroll to Baalman.

References

Abdelnour-Nocera J, Michaelides M, Austin A, Modi S (2012) An intercultural study of HCI education experience and representation. In: Proceedings of the 4th International Conference on Intercultural Collaboration, ICIC'12. Association for Computing Machinery, New York, NY, USA, pp. 157–160

Ahmed S (2009) Embodying diversity: problems and paradoxes for black feminists. Race Ethn Educ 12(1):41–52. https://doi.org/10.1080/13613320802650931

Ahmed S (2017) No. Feministkilljoys. https://feministkilljoys.com/2017/06/30/no/

Amershi S et al. (2019) Guidelines for human–AI Interaction. In: Proceedings of the 2019 CHI conference on human factors in computing systems. Association for Computing Machinery, New York, NY, USA, pp. 1–13

Ansara YG, Berger I (2016) Cisgenderism. In: The Wiley Blackwell encyclopedia of gender and sexuality studies. John Wiley & Sons, Ltd., pp. 1–3

Arunsaravanakumar S (2023) Deepfake music sends ripples across the music industry. The Wildcat Tribune. https://thewildcattribune.com/17528/ae/deepfake-music-sends-ripples-across-the-music-industry/

Åsberg C, Lykke N (2010) Feminist technoscience studies. Eur J Women's Stud 17(4):299–305. https://doi.org/10.1177/1350506810377692

Baalman M (2019) GeCoLa. Personal website https://marijebaalman.eu/projects/gecola.html

Baalman M (n.d.) Personal website. https://marijebaalman.eu/biography.html. Accessed 13 Mar 2024

Baalman M, De Belleval V, Malloch JW, Thibodeau J (2010) Sense/stage—low cost, open source wireless sensor infrastructure for live performance and interactive, real-time environments. ICMC

Bailey A, Cuomo CJ (2008) The feminist philosophy reader. McGraw-Hill, Boston

Barad K (1998) Getting real: technoscientific practices and the materialization of reality. Differences 10(2):87–128. https://doi.org/10.1215/10407391-10-2-87

Barad (2003) Posthumanist performativity: toward an understanding of how matter comes to matter. Signs Gender Sci: New Issues 28(3). https://www.academia.edu/1857574/Posthumanist_performativity_Toward_an_understanding_of_how_matter_comes_to_matter

Bauer GR, Hammond R, Travers R, Kaay M, Hohenadel KM, Boyce M (2009) 'I don't think this is theoretical; this is our lives': how erasure impacts health care for transgender people. J Assoc Nurses AIDS Care Transgend Health HIV Care 20(5):348–361. https://doi.org/10.1016/j.jana.2009.07.004

Beasley C (1999) What is feminism?: An introduction to feminist theory. SAGE Publications

Benjamin JJ, Berger A, Merrill N, Pierce J (2021) Machine learning uncertainty as a design material: a post-phenomenological inquiry. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21. Association for Computing Machinery, New York, NY, USA

Birhane A, Kalluri P, Card D, Agnew W, Dotan R, Bao M (2022) The Values Encoded in Machine Learning Research. In 2022 ACM Conference on Fairness, Accountability, and Transparency, ACM, Seoul Republic of Korea, pp 173–84. https://dl.acm.org/doi/10.1145/3531146.3533083

Blas Z, Cárdenas M (2013) Imaginary computational systems: queer technologies and transreal aesthetics. AI Soc 28:559–566. https://doi.org/10.1007/s00146-013-0502-y

Boden MA, Edmonds EA (2009) What is generative art? Digit Creat 20(1–2):21–46. https://doi.org/10.1080/14626260902867915

Bordo S (1994) Feminism, Foucault and the politics of the body. In: Reconstructing Foucault: essays in the wake of the 80s. Routledge

Bourne J (2015) A theory of analogy for musical sense-making and categorization: understanding musical Jabberwocky. PhD thesis, Northwestern University

Bown O (2011) Experiments in modular design for the creative composition of live algorithms. Comput Music J. 35(3):73–85

Bown O, Brown AR (2018) Interaction design for metacreative systems. In: Filimowicz M, Tzankova V (eds) New directions in third wave human–computer interaction: vol 1—Technologies. Human–computer interaction series. Springer International Publishing, Cham, pp. 67–87

Bramantyo T (2021) Digital art and the future of traditional arts. J Musicol Res 1:96–110. https://doi.org/10.33779/2587-6341.2021.1.096-110

Bretan M, Weinberg G (2016) A survey of robotic musicianship. Commun ACM 59(5):100–109. https://doi.org/10.1145/2818994

Briot JP, Pachet F (2020) Deep learning for music generation: challenges and directions. Neural Comput Appl 32(4):981–993. https://doi.org/10.1007/s00521-018-3813-6

Burgess M (2021) Getting AI to work in a fleshy, messy world is harder than you think. Wired UK. https://www.wired.co.uk/article/artificial-intelligence-future

Candy L, Edmonds E, Poltronieri F (2018) Explorations in art and technology, 2nd edn. Springer series on cultural computing. Springer, London

Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y (2019) OpenPose: realtime multi-person 2D pose estimation using part affinity fields. In: IEEE Transactions on Pattern Analysis & Machine Intelligence 43(01):172–86. https://doi.org/10.1109/TPAMI.2019.2929257

Carnovalini F, Rodà A (2020) Computational creativity and music generation systems: an introduction to the state of the art. Front Artif Intell 3:14. https://doi.org/10.3389/frai.2020.00014

Carroll S, Garba I, Figueroa-Rodríguez O, Holbrook J, Lovett R, Materechera S, Parsons M et al. (2020) The CARE principles for indigenous data governance. Data Sci J 19:1–12. https://doi.org/10.5334/dsj-2020-042

Cotton K, Tatar K (2023) Caring trouble and musical AI: considerations towards a feminist musical AI. AIMC

Dahlstedt P (2018) Action and perception: embodying algorithms and the extended mind. In: Dean RT, McLean A (eds) The Oxford handbook of algorithmic music. Oxford University Press, United Kingdom, pp. 41–66

De Gregorio G (2023) The normative power of artificial intelligence. Universidade Católica Portuguesa, Rochester, NY

Demers JT (2010) Listening through the noise: the aesthetics of experimental electronic music. Oxford University Press, Oxford, New York

DeVito MA, Gergle D, Birnholtz J (2017) 'Algorithms ruin everything': #RIPTwitter, folk theories, and resistance to algorithmic change in Social Media. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, CHI '17. Association for Computing Machinery, New York, NY, USA, pp. 3163–3174

Donnarumma M (2016) Configuring corporeality: performing bodies, vibrations and new musical instruments. Doctoral, Goldsmiths, University of London

Donnarumma M (2019) Biography. Personal website. https://marcodonnarumma.com/about/biography/

Draude C (2020) 'Boundaries do not sit still' from interaction to agential intra-action in HCI. In: Kurosu M (ed) Human–computer interaction. Design and user experience. Springer International Publishing, Cham, pp. 20–32

Dumit J (2014) Writing the implosion: teaching the world one thing at a time. Cult Anthropol 29(2):344–362. https://doi.org/10.14506/ca29.2.09

Ellul J, Hofstadter D (1979) Remarks on technology and art. Soc Res 46(4):805–833

Elwes J (2015) Jake Elwes—about. Personal website https://www.jakeelwes.com/about.html

Elwes J (2020) The Zizi show—a deepfake drag cabaret. Web-based artwork. https://zizi.ai/

Eslami M, Vaccaro K, Karahalios K, Hamilton K (2017) 'Be careful; things can be worse than they appear': understanding biased algorithms and users' behavior around them in rating platforms. In Proceedings of the international AAAI conference on web and social media, vol 11(1). pp 62–71

Fallman D (2011) The new good: exploring the potential of philosophy of technology to contribute to human–computer interaction. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI '11. Association for Computing Machinery, New York, NY, USA, pp. 1051–1060

Fosch Villaronga E, Malgieri G (2024) Queering the ethics of AI. In: Gunkel DJ (ed) Handbook on the Ethics of Artificial Intelligence, Edward Elgar Publishing, p 336

Fraietta A (2008) Open Sound Control : Constraints And Limitations. In Proceedings of the International Co nference on New Interfaces for Musical Expression. Genoa, Italy, Zenodo. https://doi.org/10.5281/ZENODO.1179537

Frank MR et al. (2019) Toward understanding the impact of artificial intelligence on labor. Proc Natl Acad Sci USA 116 (14): 6531–6539. https://doi.org/10.1073/pnas.1900949116

Frauenberger C (2019) Entanglement HCI the next wave? ACM Trans Comput–Hum Interact 27(1). https://doi.org/10.1145/3364998

Frauenberger C, Good J, Keay-Bright W (2010) Phenomenology, a framework for participatory design. In: Proceedings of the 11th biennial participatory design conference, PDC'10. Association for Computing Machinery, New York, NY, USA, pp. 187–190

Gamble S (2001) The Routledge companion to feminism and postfeminism. Routledge

Garry A, Khader SJ, Stone A (2017) The Routledge companion to feminist philosophy. Routledge philosophy companions. Routledge, New York, NY

Garry A, Pearsall M (eds) (1996) Women, knowledge, and reality: explorations in feminist philosophy, 2nd edn. Routledge, New York

Giaccardi E, Redström J (2020) Technology and more-than-human design. Des Issues 36(4):33–44. https://doi.org/10.1162/desi_a_00612

Gillian N, Paradiso JA (2017) The gesture recognition toolkit. In: Escalera S, Guyon I, Athitsos V (eds) Gesture recognition. Springer International Publishing, Cham, pp. 497–502

Girão LM, Céu Santos M (2019) The historical relationship between artistic activities and technology development. Publications Office, LU

Gondomar R, Mor E (2021) Understanding agency in human–computer interaction design. In: Kurosu M (ed) Human–computer interaction. Theory, methods and tools. Springer International Publishing, Cham, pp. 137–149

Gray JE (2022) What can feminism do for AI ethics? MLearning.Ai (blog). https://medium.com/mlearning-ai/what-can-feminism-do-for-ai-ethics-b7e401889441

Gray JE, Witt A (2021) A feminist data ethics of care for machine learning: the what, why, who and how. First Monday. https://doi.org/10.5210/fm.v26i12.11833

Greene S (2016) Bina48: gender, race, and queer artificial life. https://scholarsbank.uoregon.edu/xmlui/handle/1794/26811

Grimes [@Grimezsz] (2023a) I think it's cool to be fused w a machine and I like the idea of open sourcing all art and killing copyright. X, Formerly Twitter. https://twitter.com/Grimezsz/status/1650304205981089793

Grimes [@Grimezsz] (2023b) I'll split 50% royalties on any successful AI generated song that uses my voice. same deal as I would with any artist i collab with. Feel free to use my voice without penalty. I have no label and no legal bindings. https://T.Co/KIY60B5uqt. X, Formerly Twitter. https://twitter.com/Grimezsz/status/1650304051718791170

Grosz E (1994) Volatile bodies: toward a corporeal feminism. Indiana University Press

Güler RA, Neverova N, Kokkinos I (2018) DensePose: dense human pose estimation in the wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 7297–7306. https://openaccess.thecvf.com/content_cvpr_2018/html/Guler_DensePose_Dense_Human_CVPR_2018_paper.html

Hampton LM (2021) Black feminist musings on algorithmic oppression. In: Proceedings of the 2021 ACM conference on fairness, accountability, and transparency, vol 1. FAccT '21. Association for Computing Machinery, New York, NY, USA

Haraway D (1988) Situated knowledges: the science question in feminism and the privilege of partial perspective. Fem Stud 14(3):575–599. https://doi.org/10.2307/3178066

Haraway (2006) A Cyborg Manifesto: science, technology, and socialist-feminism in the late 20th century. In: Weiss J, Nolan J, Hunsinger J, Trifonas P (eds) The international handbook of virtual learning environments. Springer Netherlands, Dordrecht, pp. 117–158

Haraway (2016) Staying with the Trouble: Making Kin in the Chthulucene. Duke University Press, Durham

Haraway (2018) Modest\-Witness@Second\-Millennium. In: FemaleMan\-Meets \-OncoMouse: feminism and technoscience, 2nd edn. Routledge, New York, NY

Harding S (2009) Postcolonial and feminist philosophies of science and technology: convergences and dissonances. Postcolon Stud 12(4):401–421. https://doi.org/10.1080/13688790903350658

Herremans D, Chuan CH, Chew E (2017) A functional taxonomy of music generation systems. ACM Comput Surv 50(5):1–30. https://doi.org/10.1145/3108242

Hoffman G, Weinberg G (2010) Gesture-based human–robot jazz improvisation. In: 2010 IEEE international conference on robotics and automation. pp. 582–587

Husserl E (2012) Ideas: general introduction to pure phenomenology. Routledge

Ihde D (1990) Technology and the lifeworld: from garden to earth. Indiana University Press

Ilg E, Mayer N, Saikia T, Keuper M, Dosovitskiy A, Brox T (2017) FlowNet 2.0: evolution of optical flow estimation with deep networks. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Honolulu, HI, 1647–1655. https://doi.org/10.1109/CVPR.2017.179

Jensen MM, Aagaard J (2018) A postphenomenological method for HCI research. In: Proceedings of the 30th Australian conference on computer–human interaction, OzCHI '18. Association for Computing Machinery, New York, NY, USA, pp. 242–251

Kane B (2007) L'Objet sonore maintenant. Organ Sound 12(1):15–24. https://doi.org/10.1017/S135577180700163X

Latour B (2005) Reassembling the social: an introduction to actor-network-theory. Oxford University Press

Latour B (2014) What is the style of matters of concern? In: Gaskill N, Nocek AJ (eds) The lure of whitehead. University of Minnesota Press, pp. 92–126

le Roux P, van Staden C, Kraus K (2019) Phenomenology: excavating contextual practices from open distance learning tutoring experiences. In: Proceedings of the South African Institute of Computer Scientists and Information Technologists 2019, SAICSIT '19. Association for Computing Machinery, New York, NY, USA

Limerick H, Coyle D, Moore JW (2014) The experience of agency in human–computer interactions: a review. Front Hum Neurosci 8. https://www.frontiersin.org/articles/10.3389/fnhum.2014.00643

Martin CP, Glette K, Nygaard TF, Torresen J (2020) Understanding musical predictions with an embodied interface for musical machine learning. Front Artif Intell 3. https://www.frontiersin.org/articles/10.3389/frai.2020.00006

Mayberry M, Subramaniam B, Weasel LH (2001) Feminist science studies: a new generation. Psychology Press

McAra-Hunter D (2024) How AI hype impacts the LGBTQ+Community. AI Eth. https://doi.org/10.1007/s43681-024-00423-8

McEneaney J (2013) Agency effects in human–computer interaction. Int J Hum–Comput Interact 29:798–813. https://doi.org/10.1080/10447318.2013.777826

McEvoy C (2023) Vocal AI deepfakes of major artists are cropping up everywhere —should artists be worried? MusicTech. https://musictech.com/features/music-deepfakes-ai-drake-grimes-weeknd/

McKay D, Zhang H, Buchanan G (2022) Who am I, and who are you, and who are we? A scientometric analysis of gender and geography in HCI. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing systems, CHI '22. Association for Computing Machinery, New York, NY, USA

Mills KR (2005) The basics of electromyography. J Neurol Neurosurg Psychiatry 76, ii32–ii35. https://doi.org/10.1136/jnnp.2005.069211

Newlands G (2021) Lifting the curtain: strategic visibility of human labour in AI-as-a-service. Big Data Soc 8:20539517211016026. https://doi.org/10.1177/20539517211016026

Nicenboim I, Giaccardi E, Juul Søndergaard ML, Venugopal Reddy A, Strengers Y, Pierce J, Redström J (2020) More-than-human design and AI: in conversation with agents. In: Companion publication of the 2020 ACM Designing Interactive Systems conference, DIS' 20 Companion. Association for Computing Machinery, New York, NY, USA, pp. 397–400

Nowotny H (2021) In AI we trust: power, illusion and control of predictive algorithms. Wiley, Polity

Odom W, Pierce J, Stolterman E, Blevis E (2009) Understanding why we preserve some things and discard others in the context of interaction design. In: Proceedings of the SIGCHI Conference on Human Factors in computing systems, CHI '09. Association for Computing Machinery, New York, NY, USA, pp. 1053–1062

Offenwanger A, Milligan AJ, Chang M, Bullard J, Yoon D (2021) Diagnosing bias in the gender representation of HCI research participants: how it happens and where we are. In: Proceedings of the 2021 CHI conference on human factors in computing systems. CHI'21. Association for Computing Machinery, New York, NY, USA

Ohlin F, Olsson CM (2015) Beyond a utility view of personal informatics: a postphenomenological framework. In: Adjunct proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2015 ACM international symposium on wearable computers, UbiComp/ISWC'15 Adjunct. Association for Computing Machinery, New York, NY, USA, pp. 1087–1092

Oster G, Jaffe JS (1980) Low frequency sounds from sustained contraction of human skeletal muscle. Biophys J 30(1):119–127. https://doi.org/10.1016/S0006-3495(80)85080-6

Parslow J (2023) Kings, queens, monsters, and things: digital drag performance and queer moves in artificial intelligence (AI). Contemp Theatre Rev 33(1–2):128–148. https://doi.org/10.1080/10486801.2023.2173592

Pula B (2022) Does phenomenology (still) matter? Three phenomenological traditions and sociological theory. Int J Politics Cult Soc 35(3):411–431. https://doi.org/10.1007/s10767-021-09404-9

Pyysiäinen J (2021) Sociocultural affordances and enactment of agency: a transactional view. Theory Psychol 31(4):491–512. https://doi.org/10.1177/0959354321989431

Ramazanoğlu C, Holland J (2002) Feminist methodology: challenges and choices, vol 1. Sage, London Thousand Oaks, CA

Rani A (2018) Digital technology: it's role in art creativity. J Commer `I' Trade 13:61. https://doi.org/10.26703/JCT.v13i2-9

Rietveld E, Kiverstein J (2014) A rich landscape of affordances. Ecol Psychol. https://doi.org/10.1080/10407413.2014.958035

Rosenberger R, Verbeek PP (2015a) A field guide to postphenomenology. postphenomenological investigations: essays on human–technology relations. Lexington Books, pp. 9–41

Rosenberger R, Verbeek PP (2015b) Postphenomenological investigations: essays on human–technology relations. Lexington Books

Schaeffer P, North C, Dack J (2017) Treatise on musical objects: an essay across disciplines, 1st edn. University of California Press

Schott RM (2003) Discovering feminist philosophy: knowledge, ethics, politics. Feminist constructions. Rowman & Littlefield, Lanham, MD

Shah N (2023) I spy, with my little AI: how queer bodies are made dirty for digital technologies to claim cleanness. In: Klipphahn-Karge M, Koster AK, Morais dos Santos Bruss S (eds) Queer reflections on AI London, Routledge

Shimon Robot & Friends-About (n.d.) Personal website. Shimonrobot. https://www.shimonrobot.com/info. Accessed 13 Mar 2024

Siles I, Segura-Castillo A, Solís R, Sancho M (2020) Folk theories of algorithmic recommendations on spotify: enacting data assemblages in the global south. Big Data Soc 7:2053951720923377. https://doi.org/10.1177/2053951720923377

Slide A (2012) The encyclopedia of vaudeville. University Press of Mississippi

Stephan P (2015) Designing 'matters of concern' (Latour): a future design task? In: Jonas W, Zerwas S, von Anshelm K (eds) Transformation design—perspectives on a new design attitude, Birkhäuser, pp 202–226

Stichter A (2016) The role of breaking the fourth wall in audience ego identification. Aletheia 1. https://doi.org/10.21081/ax011

Surbano EE (2023) VAVA, Thai Pop's first AI artist, has dropped her first single. Lifestyle Asia Bangkok
Tatar K et al. (2023) A shift in artistic practices through artificial intelligence. Leonardo 57(3):4. https://doi.org/10.1162/leon_a_025
Tatar K, Pasquier P (2019) Musical agents: a typology and state of the art towards musical metacreation. J N Music Res 48(1):56–105. https://doi.org/10.1080/09298215.2018.1511736
Thomson-Jones K, Moser S (2021) The philosophy of digital art. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, Spring 2021. Metaphysics Research Lab, Stanford University
Turtle G (2022) Mutant in the mirror: queer becomings with AI. In: Lockton Det al (eds) DRS2022: Bilbao, DRS biennial conference series, Bilbao, Spain. https://doi.org/10.21606/drs.2022.782
Tuuri K, Eerola T (2012) Formulating a revised taxonomy for modes of listening. J N Music Res 41(2):137–152. https://doi.org/10.1080/09298215.2011.614951
Verbeek PP (2008) Cyborg intentionality: rethinking the phenomenology of human–technology relations. Phenomenol Cogn Sci 7(3):387–395. https://doi.org/10.1007/s11097-008-9099-x
Wang, TC et al. (2018) Video-to-video synthesis. In Advances in Neural Information Processing Systems. vol 31. Curran Associates, Inc. https://papers.nips.cc/paper_files/paper/2018/hash/d86ea612dec96096c5e0fcc8dd42ab6d-Abstract.html
Wang V, Wang D (2021) The impact of the increasing popularity of digital art on the current job market for artists. Art Des Rev 9(3):242–253. https://doi.org/10.4236/adr.2021.93019
Weber J (2013) From science and technology to feminist technoscience. In: Wyer M, Barbercheck M, Cookmeyer D, Ozturk H, Wayne M, (eds) Women, science, and technology, 3rd edn. New York, Routledge
Weinberg G, Bretan M, Hoffman G, Driscoll S (2020a) 'Play like a machine'—generative musical models for robots. In: Weinberg G, Bretan M, Hoffman G, Driscoll S (eds) Robotic musicianship: embodied artificial creativity and mechatronic musical expression. Automation, collaboration, & E-services. Springer International Publishing, Cham, pp. 95–142
Weinberg G, Bretan M, Hoffman G, Driscoll S (2020b) Robotic musicianship: embodied artificial creativity and mechatronic musical expression. https://doi.org/10.1007/978-3-030-38930-7
Werner A (2020) Organizing music, organizing gender: algorithmic culture and spotify recommendations. Pop Commun 18(1):78–90. https://doi.org/10.1080/15405702.2020.1715980
Wilkinson MD et al. (2016) The FAIR guiding principles for scientific data management and stewardship. Sci Data 3(1):160018. https://doi.org/10.1038/sdata.2016.18
Wilson S, Cottle D, Collins N (2011) The supercollider book. The MIT Press
Withagen R, Araújo D, de Poel HJ (2017) Inviting affordances and agency. N Ideas Psychol 45:11–18. https://doi.org/10.1016/j.newideapsych.2016.12.002
Wooldridge M (2009) An introduction to multiagent systems, 2nd edn. Wiley, Chichester, UK
Worthy P, Hunter T, Matthews B, Viller S (2021) Musical agency and an ecological perspective of DMIs: collective embodiment in third wave HCI. Personal Ubiquitous Comput 25(4):797–807. https://doi.org/10.1007/s00779-020-01429-9
Wright M (2005) Open sound control: an enabling technology for musical networking. Organ Sound 10(3):193–200. https://doi.org/10.1017/S1355771805000932

Acknowledgements

This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Programme—Humanities and Society (WASP-HS) funded by the Marianne and Marcus Wallenberg Foundation (Grant 2020.0102), by the Swedish Research Council (2019-03694), and the Marcus and Amalia Wallenberg Foundation.

Author contributions

Cotton and Tatar conceived the central idea of the paper, inviting Kaila, Jääskeläinen and Holzapfel to contribute. Jääskeläinen introduced the Implosion Method and connected theory. All authors developed the theory informing the construction of the analytical method. Cotton and Tatar formulated the interaction of the differing theoretical perspectives into a multi-stage parallel process. Cotton and Kaila conducted the analysis on the case studies, and established the necessary constraints for the implementation of the analytical method. Holzapfel introduced Latour's theory, and informed the implementation of these theories into the analysis. All authors discussed the results and contributed to the writing of the final manuscript, with Cotton taking a leading role.

Funding

Open access funding provided by University of Gothenburg.

Competing interests

The authors declare no competing interests.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors, and thus ethical approval was not required.

Informed consent

This article does not contain any studies with human participants performed by any of the authors, and thus does ethical approval was not required. All figures depicting human figures have been provided by, and approved for usage by the artists depicted therein.

Additional information

Correspondence and requests for materials should be addressed to Kelsey Cotton.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.