



# Establishing counterpoints in the sonic framing of AI narratives

Jennifer Chubb<sup>1</sup> · David Beer<sup>1</sup>

Received: 13 September 2023 / Accepted: 17 November 2023  
© Crown 2023

## Abstract

In order to challenge dominant representations and conceptions of artificial intelligence (AI), this article explores how AI is sonically represented in documentaries. Using a corpus of documentaries alongside expert interviews with sound designers, we explore the ways in which music and sound may influence perception about AI. The notion of ‘counterpoint’ in music theory is developed as a concept to capture and explain how the integrated dynamics of human/machines are represented within these sonic framings. The concept of the counterpoint allows us to reflect on how the relations between AI and the human and how they are sonically framed in ways that separate and blend without recourse to reductive or binary futures, which potentially misrepresent AI capabilities and performance. The article identifies and develops four types of counterpoint in what we refer to as AI sonic narratives. This article provides a framework from which AI could be sonically framed responsibly, which is critical when misinformation and hype impede the public understanding of science.

**Keywords** AI · Narratives · Sound · Documentary · Culture · Public understanding · Music

## 1 Introduction

Situated at the cusp between what is here and what the future might hold, artificial intelligence (AI) is regularly portrayed through the speculative, highly visual and ocularcentric lens of science fiction. Inevitably, given that science fiction is intended to provoke and push at the extremes, these stories often reinforce dichotomies and binary positions of hope or fear [1]. These extremes are often then echoed in public discourse around AI. By drawing attention to such extremes, as Hayles [2] has argued, we may only be left with a partial or obscured picture of the integration of AI. Instead, Hayles [3] uses the term ‘cognitive assemblage’ to describe the deeply entangled forms of cognition that now combine within the everyday, including forms of technical and human agency. In doing so, Hayles [3] considers the entwining of algorithms with the human. Hayles is pushing us towards a more detailed consideration of how these varied forms of cognition combine, relate and ‘interpenetrate’.

One conclusion we can draw from this is that such a focus on reductive binaries or extremes could mislead and divert attention away from a genuine understanding of the present state of the technology [4]. Indeed, Cave et al. [5] argue that existing imaginaries frame AI ‘only in service of the dominant vision’ (p.3). These binaries are even cited in the development of policy and often dominate prominent reports and discussions about AI innovations and policies [6]. This was seen to be the case during a recent interview in which tech leader Elon Musk warned the UK Prime Minister about ‘killer robots’ such as the ‘Terminator’ who ‘could chase you everywhere’ [7], p.1) following the AI Safety Summit<sup>1</sup> at Bletchley Park, UK, 2023.

AI narratives, which are crucial to understanding the integration of AI, have become an established field of research in their own right, where the rich narrative history of AI and intelligent machines, associated imagery [8, 9], discriminatory stereotypes [10] and resulting public perceptions are explored [1, 11]. Cave et al.’s work on AI narratives provides a rich backdrop to stories about AI (2020) prompting “widespread critical consideration not of AI itself, but of the stories—particularly fictional stories—surrounding this pervasive but still largely misunderstood technology” [12].

There is an increasing focus on the way AI is presented on screen as well as across news media and fiction and a move

<sup>1</sup> <https://www.gov.uk/government/topical-events/ai-safety-summit-2023>

✉ Jennifer Chubb  
jennifer.chubb@york.ac.uk

David Beer  
david.beer@york.ac.uk

<sup>1</sup> Department of Sociology, University of York, York, UK

towards narrative responsibility and sense-making [13, 14] where “technologies are part of our stories and even shape these stories.” [13]. As a part of attempts to bring out alternative narratives of AI [15, 16] and to think in detail about the shaping of the AI imaginary, we ask how AI is sonically framed within documentary using expert interviews and a corpus of documentary films. This includes the soundtrack, diegetic and nondiegetic sound and the music which has only been peripherally explored in the context of AI [15–18]. The soundtrack—including how AI is represented in audio media, such as through the musical or audio accompaniment to the narrative—presents an important lens through which we can interrogate public perception of AI. Chubb & Maloney [18], reflecting on binary sonic framings state that “the sonic framing of AI, presented via eerie music in film, reinforces a view of an AI ‘uprising’ or some form of subtle manipulation by AI agents. The sonic framing of AI, then, combines with and reinforces stories of malevolence and danger.” (p.1). We argue that when deciding how to use music and sound to frame a story—and, therefore, the perception of the subject of the story—attention should be paid to the resonances and emotional overtones the sonic framing brings to narratives. This is particularly important with respect to AI documentaries and non-fiction which attempt to express factual information and which viewers may draw upon as a source of ‘reliable’ information.

This article aims to focus on this less explored sonic framing of what have been described as ‘AI Narratives’ [1] with the aim of moving away from binary sonic framings of AI. In doing so, we propose a field of inquiry into the concept of AI sonic narratives—as relating to how we audibly communicate and conceptualise AI through sounds, soundtracks, music, and the energy, resonances, noises and voices that accompany narratives about AI.

Cave et al. [1] suggest that socio-technical imaginaries, of the type discussed by Jasanoff and Kim [19], should include an explicit account of the important role that narratives play and assert the need for further investigation into the ways narratives impact upon the public. This is supported by a growing body of research into the critical value and persuasive power of AI narratives [12, 13, 20] and their ethical dimensions [21]. We add to the corpus of research into AI narratives (2020) by focussing their sonic dimensions. Along with an analysis of a set of 10 documentary films dealing with AI, we also discuss a set of five expert interviews conducted with established artists in sound design.

We argue that the sonic framing of AI narratives is crucial to the perception and ethics of AI [18]. To do this, we use the concept of musical counterpoint. A counterpoint accounts for the texture and interplay of music by emphasising the relationship between two or more musical lines. We repurpose and develop this concept to describe how AI is sonically framed and represented by audio in four particular

ways. We find that there is a tendency for some sonic representations of AI to reinforce dominant narratives of AI. How AI ‘sounds’ can be used to disrupt or act as a discordant presence that unsettles or provides alternative perspectives on AI. This article takes the concept of the counterpoint and adapts it in order to reveal both the sonic dimensions of AI narratives and the meshing of different forms of human/machine agency depicted in those narratives. We will focus first on the sonic framing of documentary narratives and then on the project that informs this article before defining this concept of the counterpoint further.

## 1.1 Documentary film, sound and the sonic framing of AI

There is a long history to the study of the role of images in framing stories [22]. So too there is a kind of ocularcentrism about technology with the emergence of deepfakes and memes [23]. The prioritisation of vision over sound presents a problem when we only observe external appearance, instead of the meaning that sound portrays. There is also a long history of sound experiments in documentary film going back to at least the 1930s (as described by Cox [24]. Goldman [25] has described the pessimistic framing of technology as being linked with corporate power and anti-expertise. Weingart et al. [26] observed scientists are depicted as maniacs or unethical geniuses. These narratives are not only visual but sonic as well. Joseph Auner [27] discussed how images and imaginaries of technology have influenced music itself, including the use of old or outmoded and vintage media in the creation of new music. In particular, Auner refers to how the sounds of old machines can be made to ‘speak in a variety of interpretative frameworks, including human vs mechanical’ (In 5. p.2). Often, AI is sonically reinforced as being unlike humans—as Weber describes, ‘a lifeless machine is a reified mind’ [28]. Often electronic sounding monotone ‘voice glitches’ [29] and images of non-humanoid robots against cold, blue backgrounds dominate [8]. Work on global narratives shows us clearly an anglophone preoccupation with more negative associations, compared to elsewhere across the world [5]; AI Global Narratives—Leverhulme Centre for Future Intelligence [30]).

Given its significance to perception and audience interpretation the effect of background sound for documentaries has been relatively neglected [31]. Studies differ on the degree to which background music affects perception of the topic [32–34]. The effect of the sonic framing of the film ‘Jaws’, for instance and the impact of it on the way the public perceived sharks, is an excellent example of framing a story with fear [35]. Wingstedt et al. [33] and Nosal et al. [32] describe how background music creates meaning, suggesting, ‘though typically experienced on an unconscious and unreflected level, this kind of music actively contributes

narrative meaning in multimodal interplay with image, speech and sound effects.’ (p.193). In screen media, this power is used liberally to create the pace and the framing of narratives [32]. Bouzourou [36] refers to this too in the ‘Making of Jaws.’ Using this case, Wingstedt argues for the relationship between image and sound where ‘multimodal statements’ are made when the two are combined. It has long been clear that music has a function of processing and evoking emotion [37, 38]. Music and sound can manipulate how we feel about certain people, concepts, or things [35]. Yet, the relationship between music and storytelling is often overlooked [31] despite the potential to impact audiences’ perceptions. Schaffer [39] describes how studies have shown that speed, rhythm, pitch, pitch range and intensity of sound stimuli can carry information regarding danger [40] and urgency (Edworthy et al. 1991). Schäfer et al. [41] present further the case for attention to be paid for the role music perception has in creating a sense of safety in one’s environment. Music, they argue, functions to create safety—related information about environments, which, when reverse engineered, indicates that the experience of stress and danger can be responses to other sonic experiences. Likewise, studies have investigated the link between musical expression of emotions and movement [42].

Nichols [43, 44] describes the role of sound in documentary as critical, stating that “documentary begins with the viewer’s recognition of images that represent or refer back to the historical world. To this, filmmakers add their own voice, or perspective, by various means” (p.4). In their description of the advent of sound, in which the film industry moved away from silent documentary film making in the 30 s, they show how compelling sound can be—arguing that while music and sound add a lot to documentary, its arrival meant that documentary was “both richer (in potential) and poorer (in its prevailing practice) for it” [43]. Whilst historically this may be one view, Herget and Albrecht [34] argue that the music can be a problem for representation, that unlike the use of music in drama and fiction, “the use of background music in non-fictional media formats could be considered unnecessary or even problematic” (2022. p.509). If included, what we do understand though, is that music has a function [37]. Alongside rhythm and voice, stories and storytelling always have an audio dimension. At times, the sound might be in direct competition with the dialogue [45]. Ruoff [45] describes how some documentarians will offset narration to counterpoint image and sound, a point we will develop later in this article.

It is clear that narratives are critical to sense-making technology [14] and AI is particularly subject to hype and embellishment [46, 47]. Cox [24] laid some important work tracing the connection between documentary film and sonic exploration. On the use of technologies and its effect on creative outputs such as British documentary aesthetics. Cox

speaks of how Grierson, (1931) refers to soundtracks that can appear anecdotal when sound can act as a powerful tool to help us understand the moment. He writes:

*‘There must be a poetry of sound which none of us knows ... Meanings in footsteps, voices in trees, and woods of the day and night everywhere. There must be massed choruses of sound in the factory and in the street and among all men alive ... I know not the first thing about them, though I have, like everybody else shut my eyes ... and sat for hours trying to make something of the door-bangings and footfalls and crazy oddments of conversation that broke the plush darkness of a London night. We are the tyros, all of us, with a new world opening up on the horizon. I see no reason why anyone at the moment should envy Columbus’ [48].*

In this way, the sonic framing of AI may offer a moment to consider the new world opening upon us instead of reinforcing binary positions. Boon [49] describes the ways in which filmmakers use sound in film to represent industrial modernity, where the term ‘soundscape’ [50] is ‘both a world and a culture constructed to make sense of that world.’ It is not surprising within this framing, that sound can be a powerful tool in which to help us better make sense of AI.

The role and importance of sound in narrative has been illustrated in accounts of music as a narrative structure in Hollywood Film [51], sonic motifs and soundscapes in science fiction, such as Blade Runner and The Terminator [52, 53]. Music has also been identified as a source of narrative information in HBO’s Westworld [54]. In television accounts of techno-dystopianism, there has been reflection on representations of AI such as in Charlie Brooker’s Black Mirror [55]. Yet, in more specific terms, when it comes to the broad genre of documentary film on AI, there is less mention of the specific sound design of AI documentaries or the potential of those sonic portrayals to impact public perception.

Despite the importance of sound to narrative, we note that the sonic aspects of AI in documentaries have rarely been directly accounted for and neither have their varied forms of potential consequences. With an increased focus on AI responsible AI, explainability [56] and literacy [57], documentary is becoming an influential genre in shaping understanding and mapping AI present and futures. The sonic framings used for AI narratives will be an influential part of how AI is understood and can be approached into the future.

## 2 Methods

To explore what we will refer to as ‘Sonic AI Narratives’ we use two sources of data. The thematic analysis of ten AI documentaries along with five expert interviews with directors,

**Table 1** Documentaries about AI

| Title  | Narration           | Company                  | Documentary film type |
|--|---------------------|--------------------------|-----------------------|
| Almost Human                                   | Human               | Bacon                    | Indie/arthouse        |
| Robots: a brave new world                      | Human               | Java films               | Commissioned          |
| Hyper Evolution: the rise of the robots 1 & 2* | None                | BBC 4                    | Commissioned          |
| The Truth about Killer Robots                  | Automated           | Third party films        | Indie/arthouse        |
| AI Final Countdown                             | Human               | Reality entertainment    | Indie/arthouse        |
| We Need to Talk about AI                       | None                | Warner                   | Commissioned          |
| Coded Bias                                     | None                | Netflix                  | Mainstream            |
| Social Dilemma                                 | None                | Netflix                  | Mainstream            |
| Alpha Go                                       | Human               | Moxie pictures           | Indie/arthouse        |
| iHuman   | Some Automation VFX | Norwegian film institute | Indie/arthouse        |

\*We refer to parts 1 and 2 of Hyper Evolution as a single film

composers and artists responsible for the sound design of documentaries about AI. The documentary films included in the sample are detailed in Table 1. The project received full ethics approval. Documentaries were analysed by reflexive note taking, taking note of the sonic tropes used to highlight themes, for instance that of alienation, fear, excitement or when referring to particular objects such as robots, the brain, for instance. These themes were drawn inductively and then refined following the coding which took place during interviews with composers.

Interviewees were given a consent form to read and sign electronically agreeing to take part in the project, consenting to being audio recorded, given the right to withdraw at any time and that their participation in the project be treated anonymously. Participants were given an information sheet which detailed the purpose of the project. Interviews were conducted over an online video conferencing platform Zoom, only the audio was captured, and participants will be asked to turn off their videos. For the purposes of this article, interview verbatim quotations from sound designers and directors are given the indicator DocSound followed by a number 1, 2, 3, 4 and 5. They are not directly related to the films for anonymity purposes. Participants were broadly asked to describe what stories and tropes they associate with AI, how they approached the sonic framing of this kind of technology in documentary, what themes they wanted to address in the documentary, who decided on those themes, what informed their sound design and musical choices. Data were thematically analysed using NVivo 12 software using inductive and deductive approaches—this process was iterative.

Sampling was carried out based on desk research of documentaries about AI and web searching. Crucially, this paper does not examine the power relations, incentives or self-interest of the artists concerning the type of documentary. We of course acknowledge that this is an interesting line

of enquiry for further research but make clear that this is outside of the remit of this particular article.

Interviewees were recruited for these semi-structured interviews via email and social media request. The interviews were used to add context and to enrich the analysis of the documentaries themselves. We do not reference the name of the documentaries in the discussion as related to the verbatim quotation directly in coherence with our ethics protocol. Alongside these films, the interviews enabled us to gain insights into the role of sound in framing AI.

## 2.1 Sound designers' views of sonically framing documentary

The interviews reinforced the idea that music can shape how we feel about certain people, concepts, or things [35]. The sound designers interviewed went as far as to claim that it is amongst the most important and often underrated elements in shaping audience responses:

*I would argue that that the sound design and the score is one of the most powerful elements you have and can, I mean music can form a narration for a film... music guides the viewer through...a movie in a way that helps funnel their emotions in the direction that the filmmaker wants to go, and I never pretend that it's not a manipulative art form, it is manipulative... with numerous tools, and one of the most powerful is sound design and score. And so the responsibility to use that responsibly is important and needs to be thought about. DocSound01*

Rogers [31] describes how music and sound is an overlooked narrative feature in documentary and non-fiction—with great power to influence an audience about 'reality and fiction'. Thirty years ago, Michael Rabiger wrote a prominent work on documentary filmmaking. Now, in its 7th

edition, *Directing the Documentary* explores the role and responsibility of the filmmaker in presenting factual narratives to an audience. Crucially, Rabiger discusses the use of music in documentary film saying it should never be used to ‘inject false emotion’. In the same interview, though, our participant elaborates on the importance of the soundtrack, suggesting that it is the sonic framing that enables the audience to be guided through a ‘journey’ and that those sonic framings can be used to create a more ‘cohesive’ feeling for the film:

*I approached it the way I'd approach any film, I use music and sound to help take the audience on a journey. One of the key things for the soundtrack, for the score, is to... often films are bitsy, and this film had a lot of bitsy little elements, you try to use that to tie it all together, so that it... there's a continuity of how the piece feels as a piece of work, so you can use the music to tie the different themes together, sometimes to subliminally I guess give the audience cues as to where they are, 'Okay, now we we're in this space,' or, 'Now we're talking about this question, this is the little theme that goes with that'. Music is particularly helpful and powerful in helping to give a film a cohesiveness. DocSound01*

The indication here is that the sonic framing is a powerful presence in facilitating feel in film and in guiding the audience reception and interpretation of the content. The soundtrack can, it is suggested, bring bits together. The role of responsibility in AI documentaries takes us to Mark Coeckelbergh's [13] calling for a focus on narrative responsibility. This sense of guiding an audience through information and building up a picture of the science and innovation is one we found reiterated to us by one sound designer:

*The whole idea is to let people know what's actually going on in the world of science but also especially in kind of, yeah, not only artificial intelligence but the crossover of things with modern technology, finding more about being a human being. DocSound02*

Here, we can see there is an attempt not just to provide insights into the science but also to think about and socially contextualise the relations between technological developments and human experience. In keeping with this, we find that most documentaries use sonic textures and sound to emphasise both the ‘human’ in the story and the integration of technology. The problem then is in finding a way to separate the different depictions and sonic framings of these human and machine relations. It is in addressing this problem that we adapt the concept of the counterpoint as a metaphorical device to highlight the ways in which documentarians might soundtrack the cognitive assemblage. The equivalence posited between counterpoint as metaphor/

message and musical device is not to be reductive of the emotional and representational possibilities inherent in music/sound, and the way in which this occurs within a rich web of context and past listening experience but serves as a provocation for what AI might sound like.

## 2.2 Sound tracking the cognitive assemblage through the notion of counterpoint

As we have already mentioned, Hayles [3, 58] uses the term cognitive assemblage to describe the range of forms of cognition that now combine with the social world as well as the interactions occurring between them. This is a picture that becomes more complex with the advancement of AI within the assemblage. In doing so, Hayles considers the entangled nature of algorithms with the human where new techniques give rise to new modes of cognition. This is theorised but not considered from a sonic perspective. To explore this dimension of the cognitive assemblage more directly, we turn to the counterpoint. Counterpoint, as we have alluded, describes the texture of music emphasising the relationship and interplay between two or more musical lines. The notion of the counterpoint, we suggest, can be used to understand how the integrated dynamics of human machines—the cognitive assemblage—are represented within the sonic framings of documentary accounts of AI—to understand how the integrated dynamics of humans and machines—as two lines within a composition—are represented within the sonic framings of documentary accounts of AI and the interplay between them. The counterpoint occurs when one or more independent melodies are added above or below a given melody. When there is more than one independent melodic line happening at the same time in a piece of music, that music is contrapuntal. The independent melodic lines are contrapuntal.

Counterpoint, we argue, can also be used to understand how the integrated dynamics of human machines are represented within the sonic framings of documentary accounts of AI. We should emphasise that our approach here is not to directly use the concept from music theory in order to think solely about the musical scores. Instead, we are reworking, adapting and developing the concept of the counterpoint as a metaphoric tool only for thinking about how the forms of cognition within Hayles' [3] ‘cognitive assemblage’ combine and relate. Our point here is not to unpick the scores themselves, but to reflect on how the sonic framings build from counterpoints that lead towards different types of encounter with the combined aspects of cognition or agency. The counterpoint as a notion is taken and developed into a concept for the analysis of AI in particular, drawing inspiration from existing types of counterpoint whilst developing those ideas into typology of sonic framings within documentary film. This paper is not an exercise in music theory,



**Table 2** Adapted from Fux's '5 species' counterpoint [60]

|                |   |
|----------------|---|
| First species  | Note against note counterpoint—matching two voices together—e.g. 1 point in the melody is counterpointed with another point |
| Second species | Provides more of a rhythmic proportion, e.g. two minims to one note as opposed to one to one                                |
| Third species  | 4 to 1 rhythmic proportion—here the dominant line is 4 to 1   |
| Fourth species | The counterpoint is metrically offset from one another. Suspensions of one note over another                                |
| Fifth species  | The counterpoint is mixed—it sounds like actual music—mixture of the previous species in relation to the main line          |

**Table 3** Types of counterpoint in the sonic framing of AI

| AI sonic narratives framings |  |
|------------------------------|--|
| Counterpoint 1               | Matching voices: matching but maintaining a separation       |
| Counterpoint 2               | The dominant line: one line becomes prominent over the other |
| Counterpoint 3               | Being offset: several sounds against one                     |
| Counterpoint 4               | Pronounced mixing: blending and agency                       |

rather we utilise the concept means for analysing the sonic framing of AI through the development of a sonic typology of documentaries to begin to suggest what AI sounds like.

Counterpoint is the system that coordinates the voices in a polyphonic texture. The counterpoint is an expression of the how—how the listener perceives melodies and how composers attempt to approach writing the music to incorporate melodies or make elements more pronounced or mixed. Sometimes, that interwoven-ness—or in theoretical terms, the genuine polyphony—competes for a listeners' attention texturally. Importantly, not all music works with counterpoint. We use this particular term because in counterpoint you have more equal parts. With this, one can bring out important lines at any one time—in this way, interwoven concepts like 'the machine' and 'the human' can be seen as two lines which interact or even dance around each other.

It was Johann Fux who devised this method for analysing music, specifically working across '5 species' for combining two elements of voice or sound. In the case here, we use counterpoint to express the two elements at work when we consider AI narratives, that is, the human and the machine and all that comes between. Like Fux, Jeppesen and Haydon [59] describe how 'the essence of the theory of counterpoint is how two or more lines can unfold simultaneously the most unrestrained melodic development, not by means of the chords, but in spite of them' (p.402).

In adapting the concept, we have reduced this from '5 species' in music theory to 4 types of counterpoint framings. In so doing we have brought out the key features of the relations of the counterpoint to think of them in terms of AI rather than the techniques or scores themselves. In dealing with counterpoints in this way, we were able to be sensitised to the use of sound whilst focussing on narrative formations and the features of the AI being depicted. Fux's

species typology from music theory has been summarised in Table 2.

Based upon our findings, we collapse these types and create a typology of four AI sonic narratives framings in line with these broad concepts as provided by Fux in Table 3. This adapted version allows for a more specific analysis of AI framings. It shows how sound designers, composers and filmmakers are employing the capabilities afforded by music to help demonstrate complex ideas and support the experience of the viewer in a nuanced manner though many do so in service of the dominant vision, as noted by Cave et al. [1]. In the remainder, we develop this typology of counterpoints in AI documentary soundscapes, exploring each in turn and illustrating how we arrived at the typology based upon the sampled films and expert interviews. We note that we have focussed more on music and sound than on established aspects in soundtrack analysis like voiceover and silence—which could be deepened for further research.

## 2.3 Counterpoint 1: matching voices

### 2.3.1 Matching sounds but maintaining a separation

The first counterpoint is concerned with how voices match-up within the sonic framing of AI documentaries. Here, two sounds act as counterpoints to one another, *matching* but maintaining a separation. In this instance, the sonic framing of the counterpoints is used to depict a separation of the human and machine. These are matching and equivalent in presence voices that maintain a separate status and a voice distinct from the other. This type of approach to sonic framing is illustrated in one account of the sound design process in which the sound designer was working with the graphic imagery to match the voices depicted. One example is of a

stereotypical monotone robot voice with vocal glitches to match robot imagery. They explained that:

*It depended on which elements I was using the sound of the music on. I mean, there's a graphic element in the film where we've got those sort of graphic-y, which is basically me just... and that was hard work trying to come up with something that illustrated, you know, it's almost like a scented wallpaper because you're trying to show something that doesn't exist, so you have to find a way to do that, so the music and the sounds that we used. DocSound03.*

In this instance, the soundtrack was designed to reinforce the graphic imagery. The sound was used to illustrate the elements within the film, thus focussing more on the matching of voices around those depicted elements—with the elements maintaining the separation within the sonic framings that matched the images. For instance, a crescendo using electronic sound was used to match imagery of an explosion. In this case, the added problem being faced was how to depict a future through those elements and sounds.

These matching voices within the sonic framings are understood in terms of mobility and the way that movement occurs within the relations being depicted. Musical and corporeal movements are imagined together. The same composer described how:

*We work with the movement...I wanted to have a feel of flow through those graphics, you know, there's explosions which are representing things occurring at certain times, so the sound design around the imagery that was representing the AI was actually driven probably by the imagery. But I did want a sense, I wanted a sense that the AI might be communicating, because that was one of the conversations we were having in the film, the AI, 'what if the AI are talking to each other and bypassing us?' and that's something I think we need to be, not necessarily afraid of, but we need to be talking about the possibility of that happening, I think it is happening anyway already in a way, the machines do kind of talk to each other, they know that. DocSound03.*

The establishment of the connection between voices is implicit in parts of this account. The focus on the communication between AI and the 'conversations' within the film are illustrative of the role of this particular counterpoint in the sonic framing. The matching voices become a means through which this type of interactive and communicative element is expressed and reinforced. The matched voices, in this case AI voices, talking to one another become a feature of the sonic framing, in order to replicate the imagery of the film and the narrative. The counterpoint here is not just between humans and AI, but also the interaction and

relations between AI systems. The soundtrack is used to capture these separate but interacting voices. When it comes to this interaction and talking, it is explained that there:

*...was a sound element that kind of represented that... and I wanted those graphics to have a sense that there was a communication. It's amorphous, and it's representational rather than almost, you know, almost metaphoric rather than specific' DocSound04.*

The representational properties of the sonic framing are the focus in this case, capturing something of the inter-relations and communications. An amorphous sense of those voices is integrated into that framing. The composer seeks to create *"a sense of the grand...because these are the issues, you know, what it means to be human and what it means to be intelligent, you know. That's really difficult and deep."* DocSound03. The use of matching voices maintains the separation and distinctiveness in clear terms between the voices and agencies of humans on one hand and interacting AI systems on the other. Non-diegetic sound such as the subtle undercurrent of a rumbling drone allows the audience to hear what the experts speaking in the film do not.

## 2.4 Counterpoint 2: the dominant line

### 2.4.1 One line becomes prominent over the other

The idea of the dominant line contrasts with the relative equivalence of the matching voices counterpoint. It is not the case that this always reinforces binary positions of good or bad, it is that one line becomes prominent over the other. In these cases, the dominant line is used to highlight particular themes. As it was put in an interview: *"the way that our humanity is challenged with both the more pressing question and the more underrepresented question in terms of the way that these issues are talked about. So I feel like a lot of the aesthetic and decisions are based on that"* DocSound01. The dominant line is important for guiding attention to one aspect within the assemblage and therefore encouraging an emotional response in an audience. As it was explained: *"I feel like that's so much that fundamental feeling, either being duped by our own wellbeing or slightly seduced by or slightly empathetic or slightly reliant, that's our future"* DocSound01. The dominant counterpoint maintains the assemblage whilst drawing out one aspect amongst the parts.

As this suggests, this counterpoint describes the musical discourse in which two voices conduct each other independently combined to form a coherent whole. This device is used to frame particular themes—for instance, automation. In *Hyper Evolution*, we see robots in a factory—the sonic framing is grand and building strings are used to depict a 'robot Jurassic Park' where the robot claws are replacing human beings on the factory floor. Here, they are described

as a new species and as ‘a bit menacing’. The audio matches this menacing vibe about automation in factories by bringing a dominant line to the fore. The narrator describes how the place ‘oozes’ production and industry. The music is then a rising scale of ascending strings—an empathetic cello line but distorted. Intended to reflect an aspiration to embrace robots, not fear them, perhaps. A similar device is used in *The Truth about Killer Robots*. It was observed by one composer that “*you almost kind of sometimes try to catch yourself knowing that I can rip plausibility on the texture of the sound and how it relates to robots, but then I was kind of thinking on this and if I’m being honest like I try to untangle on a lot of it.*” *DocSound03*. Therefore, one line comes to the surface. In another, the dominant theme was used to emphasise particular emotions: “*so some places we had to kind of go for more the mystery and, of course, sometimes we go for the comical*” *DocSound02*.

In the case of the former, there are times when the view almost feels hypnotised by *Almost Human*—a magical swirling temporal sound creating a sense of confusion. In addition, to a point brings the viewer in and tells a compelling sonic story about the mystery of AI futures. In another, a scene in *Hyper Evolution* shows robots from Boston Dynamics (e.g. Atlas). Sonically framed by dark grumbling soundscape—a drone against electronic bleeping looms while the dialogue moves to talk about robots becoming their own species. Atlas is shown, accompanied by booms and bleeps, on an 80s low synth reminiscent of sci-fi. Suggesting otherworldliness as they talk about pushing the frontiers of robotic movement. Exciting and dynamic music to accompany ground-breaking robotics (robots that look like animals). In this scene, the narrators are pushing the robots, making them fall over. The music seems non-emotional, someone remarks ‘there’s no emotion here’. A cinematic dynamic brings it to a close with strings, building and soaring. The intended impact, it would seem, is emotional as they contemplate what humanity wants from robotics and whether machines will start to think like humans. The counterpoint promotes this aspect of the narrative. One participant remarks: “*this is more like maybe the sound, trying to make the sound and the music work on that deeper level because you have a narrator and the narrator actually tells the story somehow.*” *DocSound03*. The dominant line draws attention to that aspect of the narrative over the others within the film. One participant openly remarks:

*It is no secret that one of the big discussions was how dark it should be, the whole soundscape, and the atmosphere of the film, because I was a little bit worried at some point but they're getting too gloomy. So, we had to find a balance because, of course, the history of the Anthropocene is not very bright so, of course, but we had to find some kind of where it actu-*

*ally is something positive but also when it maybe has some kind of ambiguity, so it's a little bit up to you as a viewer to find out...DocSound05.*

Techno-pessimism is regularly reinforced by portrayals of AI in visual and sound media—suggestive of a dystopian future. Eerie music in film, for instance, can reinforce a view of AI uprising or express some form of subtle manipulation by AI agents.

## 2.5 Counterpoint 3: being offset

### 2.5.1 Several sounds against one

The third kind of counterpoint concerns being offset—in this case, there are several counterpoint notes on the one hand standing against one note. This device is seen in *Hyper Evolution One*, for instance, when it depicts the specific elements of humans and machines. In a scene about robotics, we are introduced to Robot Erica—who has the most human-like robot in the world. Erica is created to be the most human and most beautiful. Her voice is quite gentle. To accompany this scene, cello and lyrical strings, gentle and natural, are reflected in themes of beauty, aesthetics and robotics. The human and Erica have a conversation about their hobbies, likes and dislikes, and pizza. They discuss differences between humans and machines, he feels rude for ignoring her, despite pointing out that she cannot enjoy or feel anything. The scene describes how, when social robots become a part of our world it is not a person nor a machine, but it is this new category in between—a hybrid. It is at this point that an 80s synth sound breaks up the narrative, offsetting the harmony. We meet the Japanese innovator who believes there is no difference between humans and machines. Believing man-made objects can possess the same ‘spirit’. This sonic AI narrative is lyrical, ambient, and instrumental, reflecting the empathy expressed. The scene then sonically flicks back to a notional account of the West, and the music shifts to an eerie pounding synth. Suggesting how unnerved those in the West are by humanised robots. Words such as ‘terrifying’ and ‘scary’ litter the narration. The feel of the piece is intimidating. We meet a modern replica from the 1920s version of Eric the Robot—the Man made of Tin. More tension building soundtrack plays, with pulsing and high-pitched swelling sound. Here they talk about robots as moving from slave to uprising. We are told that there is a ‘niggling doubt that they are going to destroy us’—quite offset from Erica presented just mere moments ago. Multiple accounts are being offset as different versions of human and machine relations are depicted.

In *Almost Human*, the sound is slightly threatening to depict the omnipresence of AI—there is a drone note—some electronic cello and discordant strings when there is



talk of climate threats and the ice caps melting. This would signal that something really quite dramatic was happening and menacing music would emphasise this. The offset then comes. Following this a light-hearted, more lyrical or melodic sort of movement would split away from that. This was the case for instance when describing attitudes towards AI as culturally relative. In this sense, it was juxtaposed.

## 2.6 Counterpoint 4: pronounced mixing

### 2.6.1 Maintaining the separation of types of agency whilst also representing a blending of distinct sounds

The fourth kind of counterpoint is based around the way that the sonic framing situates the human and machine within a more pronounced form of mixing. In this, the sonics are used to maintain the separation of types of agency whilst also representing a blending of distinct sounds. As one documentarian/composer put it, it is *'more like a suggestive or like the ties that binds it together, so the whole idea was that every theme of the film has its own sound'* (DocSound03). The sound is used to make themes within the documentary distinctive whilst also connecting or mixing them into the narrative structures. As was then explained further, *'So when we are in the brain, the brain has one sound and is coming back. The internet has a sound. The robot has a sound, etc. And that's how we kind of approach the sound for this film'* (DocSound03). The components or elements within the narrative are allocated sounds, including the brain, the internet, the robot. These do not though maintain the same type of separation, with heightened mixing becoming means for narrating how these elements deeply relate.

One key aspect of this is how the organic is used to represent the inorganic, meaning that a pronounced mixing occurs within the sounds themselves and therefore within the depiction of human and machine relations. One sound designer explains that they *'wanted to make organic sounds from electronic sources'* (DocSound02). This integrates the living into the sonic framing of technology. They go on to explain further that:

*Some of the more electronic stuff is made of organic stuff. That means that when you see kind of the neurons of the brain, and it's measured with some machines, it's actually insects, the sound of it, and the same when there is some part where we see some planets and stars and then I turn it around and then I use what we call a B-format microphone which picks up electromagnetic signals. But the sounds of electromagnetic signals when you hear them, I don't know what they are, they're the sound like cicadas or crickets or things like that. DocSound02.*

The separation of the electronic and the organic becomes a boundary for questioning and exploration. In this instance, the sound of insects is used to soundtrack the technological visualisation of brain function. It was added that they *'combined electronic organic sounds and organic electronic sounds together'* (DocSound02).

In this counterpoint, a pronounced mixing of the organic and technological is occurring, this is achieved through the combination of sounds and the mixing of organic sounds with inorganic imagery. The mix of sounds is actively being used to try to emphasise the integration of different aspects of the assemblage being depicted:

*So they are mixed sometimes with the more electronic stuff and I also use a little program called Biotech which is a program where we kind of put organic stuff in it and then you can manipulate it so you get more like a musical tone. And again it's a little bit blurred, the source, but it makes that really nice connection between, yeah, the technology and the biology. Doc-Sound02.*

The use of organic sounds is part of how this shaping is approached, with boundaries then being reshaped through pronounced mixing.

The sounds of the different elements mix whilst also having this advanced sonic mixing of the organic framing and technological imagery. The pronounced mixing of sounds becomes a means to achieve a sense of both contrast and meshing. The importance of mixing is directly acknowledged in the account, with the suggestion that:

*I thought we could work with sound that we actually mix things together so you don't actually know where the machine starts and the human or the more organic thing ends. DocSound04.*

Here, we see directly how the counterpoint and the use of pronounced mixing of sounds are used in AI sonic narratives to make it hard to see where the divide between human and machine is located. This tangling of forms of agency or cognition with the soundtracking of the cognitive assemblage is in part achieved through the sonic framing. The mixing is used to make it hard to see where the boundaries around the human and machine are to be placed. This is understood to be about using the soundtrack to create a feel rather than directly to reflect the narrative. They explained that *'we don't actually go into so much of sound as a storyteller, but more like that sound is the flow in the themes'* DocSound05. The sounds are used to promote themes, including the themes of the entanglement of agency and the breakdown of the simple divide between the human and machine or between the organic and the technological.

In an open access interview on the sound design of The Social Dilemma, the composers talk about how this

presented ‘really cool opportunities for sound design’ to enter an AI world within the devices—the drama of it had to convey characters and spaces and high concepts—the composers describe this as a ‘sort of sonic playground’ [61]. It is referred to as quasi-science fiction with a ‘buffet of ideas’. The interviewer describes how ‘*it never felt like we were coming back to the same sense of dread—there was always a new sense of dread in a new kind of frame*’. This framing and the ‘sonic playground’, as it is called, is captured in *The Social Dilemma*’s featured interview:

*A simplistic piece played in different evolutions. At the beginning it is more human sounding, it is more sweeping and emotional and cinematic—as you go along it starts to become more cold—all the humanity starts to fade from it and it becomes more electronic—I was fascinated by the idea of one piece to go through the whole film (The Sound and Music of The Social Dilemma, 2021).*

Here, the music transforms into a sonic representation of the processes we are witnessing being played out on the screen. This also illustrates how the pronounced mixing type of counterpoint takes different forms and leads to sonic depictions of the complex dynamics of human/machine relations—two lines that are distinct but interweave—this is also known as duophony in musical terms.

### 3 Conclusions

Sound is a powerful shaper in the documentary experience—all the more powerful because that power is poorly understood by audiences and often goes unnoticed. Nonfiction video is also a formative agent of public socio-technical imaginaries. In this exploratory article, we have asked how documentary soundtracks work to construct public understandings of artificial intelligence, and then how we might conceptualise these different sonic angles or framings. Ocularcentricism may dominate the popular discourse but AI narratives are not solely visual in form. We have explored how the relations that make up Hayles’ cognitive assemblage are soundtracked in ways that both separate and blend forms of thinking without recourse to reductive, misleading or binary stories of utopian and dystopian futures. This article begins to think about the different ways that such combinations and blending might be mapped out and explored. By focussing upon documentary film and the soundtracks provided we have posed one potential angle that would open up a wider field of analysis into the sonic framing of representations of AI, from fiction to the news media, tech promotional materials, marketing and advertising, instructional and training videos, through to art, music videos or other forms of documentary than those covered in this article. The same might

be said of the integration of AI into political communications too. The notion of the counterpoint provides a way into understanding how the sonic framings of AI operate and how the interplay of agents can be captured in relation to one another. To capture the interweaving and entanglements AI should sound contrapuntal—blending the human and natural world with the sound of machines and reflecting a more balanced and nuanced view of how we are living with AI.

The typology we have offered in this article explores four counterpoints within the sonic framing of AI narratives. Each of these counterpoints is concerned with the combination of features or agents rather than following a myopic narrative account. The counterpoint approach enables differing and even competing narratives to occur together, offering possibilities for thinking in terms of the relations between the human and machine and the active meshing of agencies that occurs. Not all sonic framings of AI are based on counterpoints, many are likely to aim to reduce this meshing and to reinforce dominant ideas and tropes of AI, which significantly impact public understanding of AI. At the same time, an exploration of where counterpoints occur and the various forms those counterpoints can take across or even within documentary film can then be used as a point of comparison when understanding the wide-scale sonic framings of AI narratives for more responsible storytelling.

One lesser explored lens of this enquiry which we were unable to cover here but which is nonetheless relevant to mention is the ways in which documentary makers with differing incentives may choose to frame AI. For instance, we note a prevalence within indie and arthouse filmmakers to consider in a more nuanced way of sonically framing AI as both relative to national and global contexts, as well as avoiding binary tropes of ‘good’ or ‘bad’ reflected in loud/discordance or more gentle fluent lines. So too we note their tendency to avoid binary analogue/non-analogue sound when framing AI unless it is to make transparent differences in capabilities, as opposed to inflating them. Conversely, we note the tendency of documentarians with large investment, such as Netflix to use attention grabbing sonic framings in which to excite an audience. These dimensions must not be ignored in the sonic framing of AI and ought to be the subject for future research.

As we have shown, bringing in the sonic dimensions can show how alternative and varied narrative structures may be accompanying AI and illustrate how the sonic framings may complicate or create questions for a public audience. A key aspect of analysing the sonic framings of AI is to look not just at where they simply replicate narrative in audio form but also where they might unsettle and provoke different readings of the AI undermining public understanding. The sonic framing can create counterpoints in which binaries are problematised through the blending of sonic lines—a combination that we have begun to reflect upon as being

part of an attempt to think about alternative ways of blending distinctions between human and machine that go beyond separation and conflict on one hand and simple integration and sleek future-making on the other.

It is crucial to consider how a failure to consider the sonic framings of AI may influence or undermine attempts to broaden public understanding. Based on our preliminary impressions in a field that will need development, we argue that the sonic framing of AI is just as important as other narrative features and propose that creatives, practitioners consider the impact of the sonic framing of media in a responsible way that considers the publics' resources for interrogating not just what is seen, but what is heard. Further research needs to show how sonic framing connects to wider linguistic, ideological, discursive, visual, and other forms of framing in documentaries and the implications more generally for the movement towards responsible AI.

This can extend to ethical reflections on the development of voice agents as well as representations of AI in the media and in culture [15, 16]. How AI is responsibly sonically framed is then a critical question at a time when misinformation and varied interests make it hard for the public to gain a balanced view of the current state of the technology through the lens of documentary. AI sonic narratives can be contrapuntal, expressing both the independent nature of humanity from machines as well as the 'intimate entanglements' [62] of living with AI. The concept of the counterpoint is one means by which we may both analyse and think creatively about the sonic framing of AI.

**Acknowledgements** This research was undertaken while the authors were working on XR Stories funded by the AHRC (AH/S002839/1). We would like to thank Dr Liam Maloney for his contributions to the AI, what is that sound project and his thoughts on this article. We would also like to thank the participants who shared their insights with us during the research for this article.

**Funding** This article is funded by Arts and Humanities Research Council, AHRC (AH/S002839/1), Jennifer Alison Chubb.

**Data availability** The data analysis can be made available on request in accordance with ethics protocol.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Cave, S., Dihal, K., Dillon, S. (eds.): *AI Narratives: A History of Imaginative Thinking About Intelligent Machines*. Oxford University Press (2020)
2. Hayles, N.K.: *My Mother Was a Computer: Digital Subjects and Literary Texts*. Chicago University Press, Chicago (2005)
3. Hayles, N.K.: *Unthought: The Power of the Cognitive Nonconscious*. University of Chicago Press, Chicago (2017)
4. Cave, S., Dihal, K.: Hopes and fears for intelligent machines in fiction and reality. *Nat Mach Intell* **1**(2), 74–78 (2019)
5. Cave, S., Dihal, K., Hollanek, T., Katsuno, H., Liu, Y., Tailandier, A., White, D.: The meanings of AI: a cross-cultural comparison. In: *Imagining AI: How the World Sees Intelligent Machines*, 16 (2023)
6. Cave, S., Craig, C., Dihal, K., Dillon, S., Montgomery, J., Singler, B., Taylor, L.: *Portrayals and perceptions of AI and why they matter*. (2018)
7. O'Grady, S.: Excited schoolboy Sunak gushes as mentor Musk warns of humanoid robot catastrophe. *The Independent*. <https://www.independent.co.uk/voices/elon-musk-sunak-ai-summit-b2440728.html> (2023). Accessed 3 Nov 2023
8. Dihal, K., Duarte, T.: *Better images of AI: a guide for users and creators*. Cambridge and London: The Leverhulme Centre for the Future of Intelligence and We and AI (2023)
9. Singler, B.: The AI creation meme: a case study of the new visibility of religion in artificial intelligence discourse. *Religions* **11**(5), 253 (2020)
10. Cave, S., Dihal, K.: The whiteness of AI. *Philos. Technol.* **33**(4), 685–703 (2020)
11. The Royal Society. *Portrayals and perceptions of AI and why they matter*. <https://royalsociety.org/topics-policy/projects/ai-narratives> (2018). Accessed 3 Nov 2023
12. Henrickson, L.: Review of AI narratives: a history of imaginative thinking about intelligent machines, (review no. 2433) DOI: <https://doi.org/10.14296/RiH/2014/2433> Date accessed: 3 November, 2023 (2023)
13. Coeckelbergh, M.: Narrative responsibility and artificial intelligence: how AI challenges human responsibility and sense-making. *AI Soc.* (2021). <https://doi.org/10.1007/s00146-021-01375-x>
14. Felt, U.: "Response-able Practices" or "New Bureaucracies of Virtue": The Challenges of Making RRI Work in Academic Environments. In: *Responsible Innovation 3*, pp. 49–68. Springer, Cham (2017)
15. Chubb, J., Reed, D., Cowling, P.: Expert views about missing AI narratives: is there an AI story crisis? *AI Soc.* (2022). <https://doi.org/10.1007/s00146-022-01548-2>
16. Chubb, J., Missaoui, S., Concannon, S., Maloney, L., Walker, J.A.: Interactive storytelling for children: a case-study of design and development considerations for ethical conversational AI. *Int. J. Child-Comput. Interact.* **32**, 100403 (2022)
17. Chubb, J.: AI Narratives, sound and equality, <https://blog.betterimagesofai.org/ai-narratives-sound-and-equality/> (2022). Accessed 1 Nov 2023
18. Chubb, J., Maloney, L.: AI what's that sound? Stories and Sonic framing of AI, <https://blog.betterimagesofai.org/ai-whats-that-sound-stories-and-sonic-framing-of-ai/>, (2021). Accessed 11 Feb 2022
19. Jasanoff, S., Kim, S.H. (eds.): *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. University of Chicago Press (2015)
20. Kim, M.S.: Meta-narratives on machinic otherness: beyond anthropocentrism and exoticism. *AI & Soc.* **38**(4), 1763–1770 (2023)

21. Sartori, L., Theodorou, A.: A sociotechnical perspective for the future of AI: narratives, inequalities, and human control. *Ethics Inf. Technol.* **24**(1), 4 (2022)
22. Messaris, P., Abraham, L.: The Role of Images in Framing News Stories. In: *Framing Public Life*, pp. 231–242. Routledge (2001)
23. Geddes, K.: Ocularcentrism and deepfakes: should seeing be believing? *Fordham Intell. Prop. Media Ent. LJ* **31**, 1042 (2021)
24. Cox, G.: ‘There must be a poetry of sound that none of us knows...’: early British documentary film and the prefiguring of musique concrète. *Organ. Sound* **22**(2), 172–186 (2017)
25. Goldman, S.L.: Images of technology in popular films: Discussion and filmography. *Sci. Technol. Human Values* **14**(3), 275–301 (1989)
26. Weingart, P., Muhl, C., Pansegrau, P.: Of power maniacs and unethical geniuses: science and scientists in fiction film. *Public Underst. Sci.* **12**(3), 279–287 (2003)
27. Auner, J.: Making old machines speak: images of technology in recent music. *ECHO: Music-Center. J.* **2**(2), 2 (2000)
28. Weber, M., Kalberg, S.: *The Protestant Ethic and the Spirit of Capitalism*. Routledge (2013)
29. Borkowski: Vocal aesthetics, AI imaginaries: reconfiguring smart interfaces. *Afterimage* **50**(2), 129–149 (2023)
30. The Leverhulme Centre for Future Intelligence (2022). *Global AI Narratives*. Accessed 15 Dec 2022 from <https://www.ainarratives.com/>
31. Rogers, H. (ed.): *Music and sound in documentary film*, p. 1. Routledge, New York (2015)
32. Nosal, A.P., Keenan, E.A., Hastings, P.A., Gneezy, A.: The effect of background music in shark documentaries on viewers’ perceptions of sharks. *PLoS ONE* **11**(8), e0159279 (2016)
33. Wingstedt, J., Brändström, S., Berg, J.: Narrative music, visuals and meaning in film. *Vis. Commun.* **9**(2), 193–210 (2010)
34. Herget, A.K., Albrecht, J.: Soundtrack for reality? How to use music effectively in non-fictional media formats. *Psychol. Music* **50**(2), 508–529 (2022)
35. Brown, S., Volgsten, U. (eds.): *Music and Manipulation: On the Social Uses and Social Control of Music*. Berghahn Books (2005)
36. Bouzourou, L. (Writer/Director/Producer). ‘The Making of Jaws’ [DVD Documentary]. On Jaws: Anniversary Collector’s Edition [DVD Motion Picture]. United States: Universal Studios Home Video (2000).
37. Maloney, L.: Music as water: The functions of music from a utilitarian perspective. *AVANT. Pismo Awangardy Filozoficzno-Naukowej*, (S), 57–67 (2017)
38. Shepherd, J., Devine, K.: Introduction: Music and the Sociological Imagination—Pasts and Prospects. In: *The Routledge Reader on the Sociology of Music*, pp. 17–38. Routledge (2015)
39. Schaffer, F.C.: *Elucidating Social Science Concepts: An Interpretivist Guide*. Routledge (2015)
40. Rosenblum, L.D., Wuestefeld, A.P., Saldana, H.M.: Auditory looming perception: influences on anticipatory judgments. *Perception* **22**(12), 1467–1482 (1993)
41. Schäfer, T., Huron, D., Shanahan, D., Sedlmeier, P.: The sounds of safety: stress and danger in music perception. *Front. Psychol.* **6**, 1140 (2015)
42. Giordano, B.L., Egermann, H., Bresin, R.: The production and perception of emotionally expressive walking sounds: similarities between musical performance and everyday motor activity. *PLoS One* **9**(12), e115587 (2014)
43. Nichols, B.: *Representing reality: issues and concepts in documentary* (1991)
44. Nichols, B.: Documentary and the coming of sound. *Documentary Box* 6, 1–8 (1995)
45. Ruoff, J.: Conventions of sound in documentary. *Cine. J.* **32**(3), 24–40 (1993)
46. Lauer, D.: Facebook’s ethical failures are not accidental; they are part of the business model. *AI Eth.* **1**(4), 395–403 (2021)
47. Milne, G.: *Smoke & Mirrors: How Hype Obscures the Future and How to See Past It*. Robinson (2020)
48. Grierson, J.: *On the Talkie Horizon*. The Clarion, January, 13. Google Scholar (1930)
49. Boon, T.: The cinematic sound of industrial modernity: first notes. In: *Being Modern: Science and Culture in the Early 20th Century* (2018)
50. Thompson, E.: *The Soundscape of Modernity: Architectural Acoustics and the Culture of Listening in America, 1900–1933*. MIT press (2004)
51. Kalinak, K.M.: *Music as narrative structure in Hollywood film* (Doctoral dissertation, University of Illinois at Urbana-Champaign) (1982)
52. Leydon, R., Hosokawa, S., Zuberi, N., Lerner, N., Coyle, R., Théberge, P., Evans, M.: *Off the Planet: Music, Sound and Science Fiction Cinema*. Indiana University Press (2004)
53. Bartkowiak MJ (2010) The intergalactic lounge: Barbarella and hearing the future. In: *Sounds of the Future: Essays on Music in Science Fiction Film*, 134
54. Marshall, K.: Music as a Source of Narrative Information in HBO’s *Westworld*. In: *Reading Westworld*, pp. 97–118. Palgrave Macmillan, Cham (2019)
55. Shaw, D., Marshall, K., Rocha, J. (eds.): *Philosophical Reflections on Black Mirror*. Bloomsbury Publishing (2022)
56. Coeckelbergh, M.: Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Sci. Eng. Eth.* **26**(4), 2051–2068 (2020)
57. Long, D., Magerko, B.: What is AI literacy? Competencies and design considerations. In: *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–16) (2020)
58. Hayles, N.K.: Cognitive assemblages: technical agency and human interactions. *Crit. Inq.* **43**(1), 32–55 (2016)
59. Jeppesen, K., Haydon, G.: On counterpoint. *Music. Q.* **21**(4), 401–407 (1935)
60. Fux, J.J., Wollenberg, S.: ‘Gradus ad Parnassum’ (1725): concluding chapters. *Music Anal.* **11**(2/3), 209–243 (1992)
61. The SoundWorks Collection. *The Sound and music of the social dilemma*. Available at <https://www.youtube.com/watch?v=7HBY51yt120> (2021). Accessed 14 Dec 2022. University Press.
62. Latimer, J., López Gómez, D.: Intimate entanglements: affects, more-than-human intimacies and the politics of relations in science and technology. *Sociol. Rev.* **67**(2), 247–263 (2019)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.