

The Ethics of Creative AI

Catherine Flick & Kyle Worrall

Abstract Creative AI has had and will continue to have immense impact on creative communities and society more broadly. Along with the great power these techniques provide, come significant ethical responsibilities in their setup, use, and the output works themselves. This chapter sets out the key ethical issues relating to Creative AI: copyright, replacement of authors/artists, bias in datasets, artistic essence, dangerous creations, deepfakes, and physical safety, and looks toward a future where responsible use of Creative AI can help to promote human flourishing within the technosocial landscape. After Vallor (2016), it suggests key technomoral values of honesty, humility, empathy, care, civility and flexibility as those which virtuous creative practitioners will want to embed within any practice conducted using Creative AI techniques.

1. Introduction

The use of artificial intelligence (AI) techniques¹ in creative pursuits has been increasing at a pace in keeping with the improvement in the use and outputs of these methods. Alongside the ethical and social impact of AI techniques more broadly, Creative AI methods have also raised interesting and problematic ethical issues, such as the emergence of deepfakes, the bias of datasets, and the potential for copyright and other authorship issues. However, Creative AI has also shown some highly beneficial impacts for society - not only the art itself, but some applications that can significantly improve people's lives and agitate for beneficial societal, environmental, and political change. It is important, therefore, to explore some of the ethical issues present in current implementations of Creative AI, which practitioners may encounter in their use of these techniques, and to also frame a future of Creative AI practice in which the positive impacts are encouraged to promote human flourishing within the technosocial landscape and negative impacts mitigated or avoided.

In this chapter, we outline some of the existing ethical issues in Creative AI and suggest ways to approach, mitigate, or avoid negative impacts from these issues. We focus firstly on ownership and authorship of Creative AI outputs, looking specifically at copyright infringement and the potential for Creative AI techniques to facilitate replacement of authors/artists. We then look at the inputs and outputs of Creative AI where we examine the issues with datasets and the artist's essence, and what role Creative AI may have in the artistic world more generally, before focusing on the potential for dangerous creations, non-consensual deepfakes, and the importance of physical safety in a world where physical AI systems can encounter bugs in their programming. Finally, we look to the future - what a virtuous Creative AI might look like that focuses primarily on contributing to human flourishing and the pursuit of a technosocial good life. We use Vallor's (2016) technomoral virtues to frame practical suggestions for practitioners to consider as they use Creative AI techniques, and as a starting point for discussion into the future of Creative AI.

¹ In this chapter we use AI as an overarching term that includes machine learning, deep learning, and neural networks.

2. Ownership and authorship

If a Creative AI is to be accepted by potential users and the public alike, it needs to be free from the risk of copyright infringement and the design of such a tool needs to be incorporated into a workflow by the user, rather than incorporating the user (or the user's identity/essence - more on this in section 3.2). Current research is examining the different facets of Creative AI relating to the use of these tools and how they impact claims of ownership and/or authorship. Examples of such research are the design of Creative AI tools and how they fit into the human workflow (Ben-Tal, Tobias Harris and Sturm, 2020; Louie et al., 2020), how humans can collaborate with or leverage Creative AI (Collins and Laney, 2017; Frid, Gomes and Jin, 2020; Hanson *et al.*, 2021) and how to best evaluate Creative AI tools and their output (Zhou *et al.*, 2020; Yin *et al.*, 2021). Recent studies have also surveyed the use of Creative AI to establish the usefulness and acceptance of such tools (Knotts and Collins, 2020; Liebman and Stone, 2020). This forms a basis for understanding that the incorporation of Creative AI into a workflow and the ethics of ownership and authorship that surround the use of Creative AI are much more complex than simple generation of artistic content. The ethics of authorship encapsulate one of the core concerns around Creative AI; that of "stealing" ideas from authors' stimuli or replacing the humans in the loop. In this section, we will look at some of the more specific issues around ownership and authorship in Creative AI.

2.1 Copyright Infringement

Lawsuits around copyright infringement are frequent within the arts, and Creative AI will add to the complexity of legal discussions around copyright. With Creative AI using databases of existing material to be trained on, both the use of that material and potentially the output of the AI could become complicated from an ownership/authorship perspective². If a Creative AI does infringe on the copyright of songs, the question will be where the blame truly falls: the user who is providing stimuli, or the creator who built the tool. For now the evaluation of the level of theft that occurs through stimuli based generation is done using evaluative algorithms (Yin *et al.*, 2021) or listening studies (Collins and Laney, 2017).

Legislation may fall to looking at prior, human-driven examples of use of copyright testing, such as sampling or remixing, which can allow an original creator the ability to claim royalties from the song that has sampled or remixed the original. However, issues of plagiarism, such as in the case between Led Zeppelin and Spirit (Carroll, 2016), could become more difficult to prove if Creative AI is involved in the process. This issue of adopting others' artistic essence is discussed in more detail in section 3.2, but Creative AI users should be aware of this potential issue in their creative process, and, where possible, allow for transparency in their algorithms to be able to prove how the art has been derived.

Finally, there might be some discussion around who owns the art if AI progresses to the point where it might be able to assert ownership. The "monkey selfie" case showed that in multiple countries,

² Indeed, one way that Creative AI has been attempted to be used to alleviate legal concerns around musical copyright and lawsuits was Riehl & Rubin's efforts to generate 'every possible melody' and release them to the public domain (Cole, 2020; Whitwam, 2020). This is an interesting use of Creative AI, although whether it will stand up in court is up for debate.

copyright is held by a “legal person”, the definition of which does not include non-human animals (Ncube and Oriakhogba, 2018). In this case, ownership-asserting AI would need to be classified as a “legal person” before it would be able to claim copyright, which would be a complex decision to make. Such a discussion about the identity of AI agents is outside the scope of this current chapter, but likely one that will be had in the coming years.

2.2 Author/Artist Replacement

Previously Creative AI tools were often left at the command line, usable only by those who understand how to access and use their terminal. Recently, these AI tools have become a lot more usable with easy to understand interfaces. In the domain of music, examples of tools that generate MIDI for use by musicians include Google’s Magenta Suite (Roberts *et al.*, 2019) and the Piano Inpainting Application for Ableton (Hadjeres, 2021). While an easy to use interface has the potential for replacing the artistic endeavour of creating music by non-musicians; both the Inpainting Application and the Magenta Suite are usable as plug-ins for Ableton, which shows a focus on incorporating the AI into the workflow of the artist (in this case musicians). This focus on incorporation of tools into the Digital Audio Workstation (DAW) shows the intention of the designers for the tool to sit within the workflow of the composer and perform an assisting role. This is important as ethically the concern would be that creatives could be replaced through clever use of Creative AI. However, researchers are designing their AI to work with practitioners in order to address these ethical considerations around their use (Roberts *et al.*, 2019; Hadjeres, 2021).

In other artistic fields, GANs are creating surprisingly good outputs, such as the Text to Image GAN by Epstein (Epstein, 2021) that can create new images “in the style of” other artists. OpenAI’s GPT-3 algorithm has written newspaper articles (GPT-3, 2020), conference talk titles (Wareing, 2021) and creative fiction (Branwen, 2020). It is quite plausible that Creative AI could be used to replace authors in some fields, for example, writing copy or other repetitive and less prestigious writing tasks. Art “in the style of” could be good enough for certain applications, and not subject to copyright or royalty claims. This use of the “essence” or “identity” of other artists is further discussed in 3.2, but it is important to point out here that there is a real possibility of the replacement of some subset of authors and artists by Creative AI applications.

3. Inputs

What goes into making Creative AI work? In this section we consider different types of inputs that might be part of the process to create Creative AI works, and the potential ethical issues with the collection and use of these data points. We explore some possibilities for ensuring responsible usage of these inputs and some scoping strategies to avoid or mitigate possible misuse or other issues.

3.1. Data and bias

As with any AI application, Creative AI needs data to use as inputs. Creative works are necessarily products of the society they are in, so data based on these will reflect societal and institutional biases. Such biases exist at several stages: in the raw datasets that might contain previously existing creative works, such as music or art; but also at the design stage that might reflect institutional biases based on the questions the Creative AI program is asked. For example, at the design stage, a Creative AI

algorithm might be designed to create potential product designs of a chair based on existing chair designs. It may be tasked with designing one that will be extremely popular in order to sell a lot of chairs. But these AI-generated chair designs will not take into consideration edge cases of chair requirements, for example, chairs for larger people, chairs for people with disabilities, environmental sustainability of the chair, or even chairs that might just need to be for taller or shorter people. Because the parameters have been based around popularity, and “most people” are likely to buy an “average” chair, the program will naturally follow that instruction and only create chairs for “average” people.

In a world that is also shifting toward AI-driven efficiency, it’s also increasingly likely that large companies could use such programs to cheaply design potentially popular offerings without any need for an actual product designer to determine whether the design is actually any good (Martinez, 2019). Another Creative AI program tasked with producing Renaissance-style artwork based on existing artwork might create scenes that only show white people in them, because the large majority of Renaissance artwork with people in them depict only white people. Musical compositions that are used to train generative models are more often than not western classical pieces, which creates bias in the output style of generated art. Additionally, the existing code that is used to analyse MIDI files in order to build compositional analysis models is built to analyse sheet music, which is expected to fit into an even twelve-tone temperament. This means that music that does not fit into the normal or standardised sheet music needs to be parsed using a bespoke algorithm that can account for non-western tones. This is possibly the reason behind the lack of representation of eastern microtonal music or aleatoric music that makes use of timbres not definable in such analytical models.

While Creative AI applications are unlikely to be making life-or-death decisions based on the data it uses, aspects of representation, edge cases and other potential impacts of poorly handled Creative AI applications could have a real negative impact on society. In order to prevent possible problems, there is a large movement within the greater AI field to look at openness, explainability, and transparency of AI systems (Larsson and Heintz, 2020). Being able to see how an algorithm processes and weighs data can help to identify biases in datasets but does not help with solving the dataset bias itself. Debiasing datasets is very much a field in its infancy, and there is a definite place for Creative AI to potentially help with this problem. Debiasing attempts can include the removal of the biased data, for example, gender stereotypes from text (Bolukbasi *et al.*, 2016; Greenwald, 2017) and speech emotion recognition (Gorrostieta *et al.*, 2019). But it won’t change the fact that Renaissance art is largely of white people, or that chairs are largely designed for people of a certain shape, size, or ability. And simply adding more diverse data (whether contrived or real) may not be enough to redress the balance. An attempt at debiasing images of skin lesions, for example, showed that it was an extremely complex task to attempt to debias the datasets despite there being features within that dataset that could help with the process (Bissoto, Valle and Avila, 2020).

Until there are more useful methods of debiasing datasets (if any such methods exist), it is unlikely that this problem will be solved simply through changes to the dataset. Indeed, specifically biased datasets might actually be desired by the creator – in both of these cases, Creative AI practitioners have a responsibility to understand and be aware of these potential biases and address them openly as part of the process, with the possibility to end the process if such bias is problematic – particularly in live performances, shows, or engagements. If the Creative AI produces outputs that are then further acted upon by the practitioner, such as designs, it is then the practitioner’s responsibility to ensure artistic and/or societal acceptability by engaging with those who might be disadvantaged, under-

represented, or in other ways negatively affected by the process, to ensure that any outputs are sensitive to the underlying issues or are critically framed to reflect the output's origin.

3.2. Artistic identity and essence

Another type of input to consider when working with Creative AI is that of artistic identity and essence, both of which come with their own ethical and legal considerations. When we speak about artistic identity, we mean the style of artistic works that arises from the individual who is creating the art. The artist/musician has developed a style of producing their art or music over the course of a career, which encapsulates their identity through the output of their chosen medium. By essence, we mean the way in which they perform their art. An example of this in musical terms is that if virtuosic violinist Niccolò Paganini were to perform a violin rendition of Johann Sebastian Bach's Suite No. 1: Prelude, the artistic identity belongs to Bach, who composed the piece of music. However, the performance of western classical music is often left to the interpretation of the performer (in this case Paganini), who has to interpret information that is represented on the sheet music with some vagueness. This interpretation and the performance it leads to is the essence of the creative practice; another performer might interpret Bach's instructions in a completely different way that encapsulates their own essence. Yet Paganini is also recognisable as having a particular style of interpretation when it comes to the performance of music - this is his identity. Another example of this in terms of artistic practice is that the essence of Peter Paul Rubens demonstrates the 'dynamism, vitality and sensuous exuberance' of the Baroque painting style (Scribner, 2000). However, when compared to other Baroque-era painters such as Caravaggio and Rembrandt, the artistic identity of each is discernible to the trained eye, although in essence they are all in the Baroque style. Much of this is likely due to the way in which they created their art which captures their own essence. Indeed, "Rembrandt" was "brought back to create one more painting" in the "Next Rembrandt" project (J Walter Thompson Amsterdam, 2017), in which "the computer learned how to create a Rembrandt face based on [...] 'typicalities' [of a Rembrandt portrait]" (Dutch Digital Design, 2018).

Relating this back to Creative AI, it is important to consider the role of both essence and identity as they apply to training or using Artificial Intelligence in the pursuit of artistic or creative endeavours. An example of this within the realm of music is that expressive rendering. Expressive rendering is the study of improving the mechanical performance of music (specifically MIDI) by training algorithms to apply temporal, dynamic or performative elements to the musical output (Widmer, 2002; Widmer and Goebel, 2004; Flossmann, Grachten and Widmer, 2009; Grachten and Widmer, 2011; Grachten and Krebs, 2014). This practice has arisen as 'a mechanical performance of a score is perceived as lacking musical meaning and is considered dull and inexpressive' (Canazza, De Poli and Rodà, 2015). This approach has been one of the first steps towards the development of an essence of performative measures within Creative AI. The data used within expressive rendering comes from datasets such as the Aligned Scores And Performances (ASAP) dataset, which is a collection of over two hundred distinct musical scores and over one thousand performances of classical piano pieces from fifteen western classical composers (Foscarin *et al.*, 2020). These performances were captured from various performers and amalgamated elements of all of their performances. However, if you were to create a dataset from a single musician and use said data as the inputs for modelling your expressive rendering, then you could be capturing, and ultimately imitating, their creative essence with the output. This raises the question of what reparations such a performance would be worth, considering that the output of the Creative AI can be used and applied to multiple projects, whereas a musician would be paid well for each recording of a performance were it to be traditionally recorded.

In recent years, Artificial Intelligence has also been leveraged in order to perform more generative tasks. Recent examples of generative models used for art can be found in music (Collins and Laney, 2017; Huang *et al.*, 2018; Collins, 2020) or the recent phenomena of blockchain or CryptoArt (Finucane, 2019; Franceschet *et al.*, 2019). These generative models capture more of the artistic identity present in musical composition or in artworks in a similar manner to the expressive rendering example above, and raise similar questions of ethical use of such technology (which could be used to create “deepfakes” of artistic works or performances), what rights artists have to safeguard their artistic identity and essence, and how artists might be compensated if these rights are violated.

An example of how artistic identity and essence can be used controversially in Creative AI is seen in the ‘Lost Tapes of the 27 Club’ (2021). This is an initiative whereby Creative AI has been used to create new songs for a variety of artists who all died at the age of 27. This was done in order to raise awareness regarding mental health issues within the music industry by Over The Bridge, a non-profit organisation that is focused on tackling mental illness within the music industry. These songs were all modelled on existing music by these artists and deepfakes were created for the original vocal performance. Although for a good cause, these tracks have received a slightly controversial reception online, but to much acclaim from fans (Brodsky, 2021; Grow and Grow, 2021). This shows how Creative AI can be leveraged in order to replace artists and create music, even for those that are deceased. The ethical considerations of deepfakes like these are discussed in depth later in section 5.1.

Finally, the shift of essence or identity to compensate for AI is likely to happen if Creative AI becomes part of the process of art creation. For example, Alex Kiessling talks about how he changed his artistic style to compensate for the range of movements the robot arms are able to make in *Long Distance Art* (The Method Case, 2013). While this is not a new thing - humans have compensated for the limitations of their tools since the dawn of art - AI-based tools can be less predictable in their requirements and could have a significant impact on the essence/identity of the artist.

Those developing or using AI that encapsulates the identity or essence of a performer or artist should take into consideration the impact on the artist/performer and ensure that their AI-generated work is not misrepresented as being that of the artist or performer or replace that artist or performer unless with their agreement.

3.3. Creative AI’s Ship of Theseus

Similarly to the issues presented above, there are deep philosophical considerations around replacing parts of a creative exercise with artificial intelligence. To slightly extend the analogy of the Ship of Theseus (Wikipedia, 2021), if the human made “planks” used to create the artistic work are exchanged with artificially made “planks”, such as methods for piecing together existing works or generating designs, at what point is the art more machine than human created? If a composer, for example, were to use a tool to generate ideas that use their own music as a stimuli, then is it still part of their original work? If they were to create a new piece, using the last generated piece as a stimuli, how long before the music looks nothing like the original piece, yet might still be considered the same piece? When does it stop being that artist’s work? What if they were a still life artist, and never painted a landscape, and the AI painted a landscape in their style? Or, if highly successful in replicating an artist’s style, does a piece of work inherit some of the value of the original artist’s work? And does it matter?

Certainly, as discussed in section 4, if the outputs are honest as to their inception, this should not matter. But there are concerns around the replacement of human creativity with machine creativity to, for example, fill gaps in larger projects that humans might otherwise fill. Video games, for example, might use Creative AI to generate music or design levels, which could remove job opportunities for composers or level designers. Would it be the same video game if the artificial components replace what might have otherwise been human-made components? In games like *Candy Crush Saga*, AI is being used to test levels, removing some of the human input that might otherwise go into the game. King argues that this actually makes the game better, as it allows for faster testing turnaround (King, 2019), but there is an argument to be made that it could potentially limit the game as the testing AI is unlikely to think “outside the box” like humans will. Although this is not traditional Creative AI work, other companies are using procedural generation in other aspects of games, such as the world building in *Minecraft* and *Valheim*; with sophisticated-enough generators it could be that this kind of integration of Creative AI follows King’s testing regime and removes human input altogether because it is faster or easier to make (and perhaps also less expensive).

While this replacement of human input may not necessarily be a problem, in the bigger picture of Creative AI integration into art it is important to recognise that this could potentially mean that the future of art inherently has Creative AI embedded within it. While there might be a push for more “traditional” art (much like digital photography has not fully replaced film photography (Keats, 2020)), it is likely that there is no going back to a time before Creative AI was introduced to the art world.

4. Outputs

This section highlights some of the potential ethical issues with the outputs of Creative AI. Recent forays into deployment of AI entities have shown that, without human intervention, AI can quickly produce outputs that can potentially be dangerous, be used to defame or cause social anxiety, or have issues with expectation management as to their capabilities (Wolf, Miller and Grodzinsky, 2017; Toews, 2020). We capture these concerns in terms of dangerous creations, deepfakes and similar problematic uses, and issues of safety in physical performance with Creative AI and humans working side-by-side.

4.1. *Dangerous creations*

It’s well known that certain types of artistic outputs can have adverse effects, for example certain types of light flashing can trigger seizures, loud noises can cause hearing loss, and certain kinds of movement in video games can cause motion sickness (Stoffregen *et al.*, 2008). It is plausible that within the AI’s creative process there could be these kinds of outputs that could cause harm. Monitoring the output and potentially putting in checks for the elements that could trigger physical harm such as these is therefore very important. We cover physical safety of artists co-working with AI driven tools (such as robots) in section 4.3.

Other kinds of harm could be socio-cultural, for example, racist, sexist, or other offensive or hate speech. These kinds of outputs have already been seen in existing AI (natural language processing) projects such as the Microsoft TayBot (Wolf, Miller and Grodzinsky, 2017), and in the bias shown in language outputs (Dinan *et al.*, 2020), for example. Much of this bias is due to the inputs described in section 3.1, and can be mitigated through thoughtful data collection and usage. Monitoring these kinds

of usages of natural language processing is a core recommendation of the more generally applied ACM Code of Ethics, which, in Principle 2.5 states: “A system for which future risks cannot be reliably predicted requires frequent reassessment of risk as the system evolves in use, or it should not be deployed” (Gotterbarn *et al.*, 2018). For Creative AI, we also recommend such vigilance.

4.2. Deepfakes

Deepfakes are the result of using AI process (specifically deep learning techniques) to simulate a person in audio, video, or still imagery. Some techniques require the use of prior work, e.g. audio or video, to create a “persona” that then renders a realistic approximation of the original person in that medium. Other techniques include “face swapping” that takes images of the person and places it over another actor (Meskys *et al.*, 2019). Some famous examples include the Nicolas Cage deepfakes where fans of the actor inserted him into a number of classic films (Neilan, 2018), or more recently (and commercially), the use of AI to generate Anthony Bourdain’s voice in a documentary about him (Tangcay, 2021). While recent *Star Wars* shows and films have also re-created actors who have passed away, they have been created without AI techniques, instead using more classic CGI techniques alongside motion capture, and were considered “eerie” (Dockterman, 2016), veering sharply into the “uncanny valley” (Hsu, 2012). Fans have since made these “performances” more realistic using the deep learning methods that are used in deepfakes, such as in *Rogue One: A Star Wars Story* (Suciu, 2020) and *The Mandalorian* (Kain, 2020).

While these might seem trivial or even useful, deepfakes have a more problematic history in that they were originally primarily used in pornographic videos to replace the actors with celebrities, or to engage in “revenge porn” - the sharing of sex videos in order to “humiliate, threaten, or mark other harm to a person who has broken off the relationship” (Meskys *et al.*, 2019). Legislation has been updated in many jurisdictions to include deepfakes within revenge porn laws (*ibid.*). Deepfakes could also be used in other ways to harm, for example, deepfakes of politicians could be used to influence elections (Diakopoulos and Johnson, 2020), or fake news made to look credible by impersonating trustworthy sources (Ajder, 2019).

Outside of these kinds of uses, deepfakes can be used in more meaningful and helpful ways as well. For example, the Deep Empathy project (MIT Media Lab, 2017) aims to induce empathy for disaster victims by using deep learning techniques to simulate disasters in cities around the world, essentially a deepfake of the cityscape. Deep learning techniques can also recreate how a motor neuron disease (MND) patient speaks from voice banks, which MND patients add clips of their voice to use to communicate when they can no longer speak (Bonifacic, 2019). It will attempt to recreate the tone, accents and colloquialisms that the person would have used when they spoke in almost real time, allowing for a greater quality of life for MND patients. These “greyfakes” also encompass the possibility to realistically represent yourself in virtual reality environments, or to create more realistic AI voice assistants (Ajder, 2019). They could also be used to create digital likenesses of you for your family and friends to engage with after you pass away. Microsoft, for example, patented a chatbot that did this, but has no plans to actually make it (as yet) due to concerns about social acceptability (Harbinja, Edwards and McVey, 2021), though it has been argued that these chatbots “can offer an important source of support to mourners” (Elder, 2020).

Throughout all of these uses of deep learning technologies runs the issue of informed consent. Deepfake technologies can be used in positive ways when accompanied by informed consent by the person impersonated and the viewer, in that they know they are watching an impersonation. When this

does not happen, situations like the backlash against Anthony Bourdain’s AI created voiceover arise, whereby viewers were horrified by the convincing impersonation that the documentary-makers created. This is especially likely with famous people that viewers feel they “know” (through the effect of parasocial relationship) (Rosner, 2021), if the person is not able to sign off on the process (in this case, because Bourdain had passed away). Another issue is disclosure of a deepfake: even though Bourdain wrote the words the AI generated voice read, the lack of indication to the audience that the voiceover was faked is a breach of trust between viewer and documentary-maker. “Creative signalling” is a way that this could be mitigated - some documentaries do this with reconstructions or other contextual signals that the audio/visuals might be indicative rather than real (ibid.). And while people may well be more comfortable with this use of technology in future, right now it warrants sensitive and context-aware use in Creative AI.

4.3. Physical safety

In Alex Kiessling’s “Long Distance Art” a set of industrial robot arms in other countries recreated an artwork in real time as Kiessling drew the original (The Method Case, 2013; Kiessling, 2021). Kiessling regularly uses industrial robot arms within his artwork, which have a potential to be quite dangerous if programmed incorrectly. Industrial robot arms have a number of recognised hazards, including impact or collision accidents, crushing and trapping accidents, mechanical part accidents, and other accidents that can result from working in the vicinity of a robotic arm such as if hydraulic lines rupture (United States Department of Labor, 2020). According to the OSHA Technical Manual for Industrial Robots and Robot System Safety (ibid.), “*the greatest problem, however, is over familiarity with the robot’s redundant motions so that an individual places himself in a hazardous position*” (emphasis sic). An artist working with a robot they have programmed to perform certain tasks, whether AI powered or not, would therefore need to be particularly vigilant, especially if they are working within range of the robot’s physical movement.

Creative AI in music has a similar potential for the use of robotics and AI to control them in the process of making music. For example, Vear’s “Embodied Musicking Robots” is a performative use of Creative AI that specifically aims to engage humans and robots together in the music-making experience (Vear, 2020). Physical safety is one of the principles in Vear’s work, though it is interesting that the emphasis is on the robot’s safety (through “self-preservation”) rather than that of any human performer that might share the robot’s stage or audience member in the vicinity. It is likely that this is because the technology is in its infancy; further development should take into account the possibility of trip hazards and over-enthusiastic manipulation of instruments by the robots involved that might potentially cause harm to humans involved as well.

5. The Future: Virtuous Creative AI

In previous sections, we have looked at specific ethical issues in the area of Creative AI and suggested mitigations or solutions for these. In this section we look more at the possibilities for ethical Creative AI. We frame this within a “virtuous Creative AI” perspective in the manner of Vallor (2016, p. 28). Virtue ethics is, as Vallor notes, “a uniquely attractive candidate for framing many of the broader normative implications of emerging technologies in a way that can motivate constructive proposals for improving technosocial systems and human participation in them” (ibid., p. 33). To do this, we will examine some recommendations for practitioners and developers that work with Creative AI from the perspective of Vallor’s technomoral virtues (ibid., p. 120). We have focused on six of these,

grouped together, which are the most applicable to the application of Creative AI in a technosocial environment.

5.1. Honesty and Humility

Vallor describes the technomoral virtue of honesty as “an exemplary respect for truth, along with the practical expertise to express that respect appropriately in technosocial contexts” (2016, p. 122). In Creative AI, this is best applied to the “creative signalling” that might be required to ensure that audiences know that what they are interacting with is AI-generated. Of course, this is highly contextual, and we are not suggesting that Creative AI practitioners need to fully disclose all methods all of the time, but more that the question needs to be asked about audience reaction to the use of AI methods within the setting and whether and what kind of disclosure is appropriate. Similarly, the use of someone’s data to create a likeness of them needs to be properly disclosed and consented to by the appropriate people, for example, in Anthony Bourdain’s deepfakes case discussed above, his family and other important people such as his manager, since Bourdain is not able to consent himself.

This virtue also ties into the general use of Creative AI: that of an augmentation of creative processes. Creative AI practitioners need to be honest with themselves about how these techniques can complement or augment their own abilities. Whether it’s through using GPT-3 to encourage creativity, e.g. as a muse for creating new characters (Smith, 2021) or to help to physically create the work, such as Vear’s musical robots (2020) or Kiessling’s artistic robotic arm (The Method Case, 2013), the focus should remain on the artist producing the work. Honesty with the audience as to the creation process is essential, however, even when that process acquires or uses another artist’s essence (as discussed in 3.2), or is used as a refinement technique for original human-created art, such as in filters for visual arts that use AI.

Honesty in Creative AI also links in with the virtue of humility, defined by Vallor as “a recognition of the real limits of our technosocial knowledge and ability; reverence and wonder at the universe’s retained power to surprise and confound us; and renunciation of the blind faith that new technologies inevitably lead to human mastery and control of our environment” (Vallor, 2016, p. 127). This should not be a surprising virtue to the Creative AI practitioner, who likely encounters it, and perhaps even embraces it, within their practice. However, humility also requires ensuring that the human context is profoundly centred within the use of Creative AI techniques - the impacts of and on humans in, for example, the inputs used or the outputs created. It is an obligation for Creative AI practitioners to be critical of the use of AI techniques in their work, rather than follow blind techno-optimism or -pessimism, to be hopeful about the possibilities of Creative AI but understanding that we do not always know what the outputs might be, and that they might, in fact, cause harm. Examples here include potential impacts on humans as a result of bias, or physical harms from unexpected movements of robots or dangerous outputs of audio or visual AI creativity as discussed in sections 4.1 and 4.3. Practitioners need to be as humble as to continue to value the human touch in the creative arts and the value of “traditional” creativity (see section 3.3), even if use of Creative AI becomes the norm. Hopefulness in humility’s sense may be in terms of looking toward future use of Creative AI, such as for non-specialists to become creative, or to use it in order to make creation of art more accessible to those who would not normally be able to create (demonstrated by Louie *et al.* (2020)).

5.2. Empathy and Care

The technomoral virtue of empathy is defined as “a cultivated openness to being morally moved to caring action by the emotions of other members of our technosocial world” (Vallor, 2016, p. 133). Vallor makes a specific distinction between empathy, “a form of co-feeling, or feeling *with* another”, and sympathy, “a form of benevolent concern *for* another’s suffering” (ibid.). Given that the outputs of creativity are frequently expressions aimed at affecting the audience’s emotions, helping the audience to understand another’s perspective, or similar, Creative AI has a lot of potential to positively affect the audience and increase their empathetic concern for a subject. Projects such as Deep Empathy (discussed in section 4.2) can show the power this might have. It is, however, important that, when affective Creative AI is implemented, that this is done carefully in order to respect the audience as well. Psychological manipulation and deception can undermine the desired outputs, with honesty, as mentioned before, an important value to hold here as well. Hence the pairing of empathy with the technomoral virtue of care here, defined as “a skillful, attentive, responsible, and emotionally responsive disposition to personally meet the needs of those with whom we share our technosocial environment” (Vallor, 2016, p. 138). Creative AI can help to promote human flourishing by incorporating the virtue of care. While regulation, such as the upcoming EU regulation of AI, is likely to limit some of the more risky uses and abuses of machine learning, such as the use of subliminal techniques (MacCarthy and Propp, 2021), there is always the potential for creative uses of AI technologies that (intentionally or unintentionally) go beyond these regulations and pose a potential for harm.

The responsible and attentive practitioner should be able to shut down any harmful AI applications and allow for reflective interrogation of the approach used in order for auditing so that others might learn from it. Techniques such as explainable AI or other audit processes for machine-made decisions are key here, with an example being the issue of bias, which was not foreseen in the implementation of AI techniques at the beginning (and the harm of which is discussed in section 3.1). The responsibility for any Creative AI’s output must always rest with a human individual such that in deploying it, there is someone that is able to halt the application if needed. The Taybot discussed in section 4.1 is a good example of where this should have happened much quicker than it did. Additionally, shifting moral responsibility onto the machine is not fulfilling the requirements of a duty of care for “those with whom we share our technosocial environment”. Similarly, the erasure of jobs for artists and technicians through Creative AI uses (such as in an unconsenting impersonation or use of an artist’s artistic essence, discussed in section 3.2, and particularly in the Lost Tapes example) should be avoided where possible; the primary goal for ethically acceptable Creative AI should be to augment or enhance existing art rather than replacing artists or those who support them.

5.3. Civility and Flexibility

Civility in a technomoral sense is not the simple call to politeness according to the use of it in common parlance, it is, instead, “a sincere disposition to live well with one’s fellow citizens of a globally networked information society: to collectively and wisely deliberate about matters of local, national, and global policy and political action; to communicate, entertain, and defend our distinct conceptions of the good life; and to work cooperatively toward those goods of technosocial life that we seek and expect to share with others” (Vallor, 2016, p. 141). The important part here is to see the possibilities for Creative AI to fulfill this and positively impact society in many different ways. Art has often comprised significant parts of movements for change including political action and raising

awareness of inequalities and injustices in the world. Art has long been an inspiration for improving society, envisioning utopias, improving wellbeing and portraying potential impact that current and possible trajectories could have in the future. Creative AI thus has potential to affect significant positive change through continuing this tradition. This doesn't mean that there won't be problematic uses along the way, but if Creative AI is used responsibly and with the "sincere disposition [...] to work cooperatively toward those goods of technosocial life" it will have a net positive impact on society.

Civility here also includes many of the mitigations and avoidances we have suggested throughout this paper, but we include civility as a specific virtue here because of the incredible possibilities that Creative AI can have as a positive impact on society. Current examples that would fit into this include the Deep Empathy project, and projects to assist with giving voice to MND patients, promote policy and political action, and allow for more people to live a good technosocial life. Respecting artists' identity, essence, and intellectual property are examples of working cooperatively. Even the Creative AI projects that bring long-dead artists "back to life", if done sensitively, can encourage wellbeing through appreciation of the arts.

The final technomoral virtue we want to apply in this chapter is flexibility. It links in with civility well, because it is defined by Vallor as "a reliable and skillful disposition to modulate action, belief, and feeling as called for by novel, unpredictable, frustrating, or unstable technosocial conditions" (2016, p. 145), and thus moderates the zeal for which Creative AI practitioners might want to practice the virtue of civility. Not only is it a moderating force in this way, but it reminds practitioners that the implementations of technologies can be unpredictable and unstable, thus ensuring that they monitor outputs and ensure that these are appropriate. Interestingly, it is here we also see the discussion of forbearance of norms that might be outside of our technosocial experiences, and whether we need to be flexible in terms of our respect for difference in cultural norms.

Vallor addresses this difference by introducing the concept of "capacity for global technomoral agency" to decide which norms should "warrant mutual forbearance", for example, a definition of "feminine virtue" that does not allow for female education or equal participation in technosocial life would not be compatible with global flourishing (pp. 147-148). Instead, active deliberation of the ways that Creative AI can contribute to human flourishing is needed to determine the contributions and methods of contribution it can bring. This translates in practice to remembering that the people who encounter, interact with, are affected by, and use the Creative AI tools and outputs and make up the inputs may have differing social or cultural expectations of these and that, for the most part, these warrant mutual forbearance. Examples from above include the nature of copyright and intellectual property, informed consent, user expectations of disclosure of the nature of the use of Creative AI tools, creation and use of datasets, augmentation of artistic ability vs. replacement of artists, etc. Creative AI practitioners should be easily able to cope with handling these uncertainties - after all, the nature of art is often unpredictable and can sometimes cause discomfort. The key to determining whether this discomfort is acceptable or not is to look at the bigger picture of global technomoral agency and, ultimately, the ability to contribute to human flourishing.

While we have covered many of Vallor's technomoral virtues, we recognise that the others have value within Creative AI practice as well and recommend to the reader to use this application of her approach as a way into further reading and understanding of how virtue ethics can help to frame the future of Creative AI.

6. Conclusion

The ACM Code of Ethics (which applies to the computing profession generally, therefore encompassing Creative AI tools and applications) promotes a goal of ensuring that the public good is the primary consideration when evaluating ethical decisions (Gotterbarn *et al.*, 2018). It is in this tradition that we have approached this chapter - focusing firstly on specific ethical issues, and then on presenting a future-looking virtue ethics framework to understand how Creative AI can positively contribute to human flourishing within the technosocial environment. It is important to not view this chapter simply as a list of ethical issues and how to solve them - but as a starting point for discussion about what kind of society Creative AI techniques will be creating, and, more importantly, what kind of society Creative AI practitioners *want* to create through their artistic practice and use of AI tools. The virtues discussed in section 5 allow for a theoretical framework to consider future inputs for and applications of Creative AI that we might not have foreseen; keeping in mind the primary consideration of the public good, or human flourishing, within a complex technosocial world.

References

- Ajder, H (2019) *The ethics of deepfakes aren't always black and white*, *TNW | Podium*. Available at: <https://thenextweb.com/news/the-ethics-of-deepfakes-arent-always-black-and-white> (Accessed: 23 April 2021).
- Ben-Tal, ., Tobias Harris, M and Sturm, BL (2020) 'How music AI is useful: Engagements with composers, performers, and audiences', *Leonardo*, pp. 1–13. doi: 10.1162/leon_a_01959.
- Bissoto, A, Valle, E and Avila, S (2020) 'Debiasing Skin Lesion Datasets and Models? Not So Fast', in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Seattle, WA, USA: IEEE, pp. 3192–3201. doi: 10.1109/CVPRW50498.2020.00378.
- Bolukbasi, T *et al.* (2016) 'Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings', *arXiv:1607.06520 [cs, stat]*. Available at: <http://arxiv.org/abs/1607.06520> (Accessed: 10 March 2021).
- Bonifacic, I (2019) *Quips is an AI to help ALS patients speak with their own voice*, *Engadget*. Available at: <https://www.engadget.com/2019-12-18-rolls-royce-quips-als-mnd-speech-ai.html> (Accessed: 16 July 2021).
- Branwen, G (2020) 'GPT-3 Creative Fiction'. Available at: <https://www.gwern.net/GPT-3> (Accessed: 16 April 2021).
- Brodsky (2021) *AI software creates 'new music' from Nirvana and Amy Winehouse*, *The Independent*. Available at: <https://www.independent.co.uk/arts-entertainment/music/news/ai-songs-27-club-music-b1827003.html> (Accessed: 21 July 2021).
- Canazza, S, De Poli, G and Rodà, A (2015) 'CaRo 2.0: An Interactive System for Expressive Music Rendering', *Advances in Human-Computer Interaction*, 2015, pp. 1–13. doi: 10.1155/2015/850474.
- Carroll, R (2016) *Led Zeppelin cleared of stealing riff for Stairway to Heaven*, *the Guardian*. Available at: <http://www.theguardian.com/music/2016/jun/23/led-zeppelin-cleared-stairway-to-heaven-lawsuit-spirit> (Accessed: 16 April 2021).

Cole, S (2020) *Musicians Algorithmically Generate Every Possible Melody, Release Them to Public Domain*, *VICE*. Available at: <https://www.vice.com/en/article/wxepzw/musicians-algorithmically-generate-every-possible-melody-release-them-to-public-domain>.

Collins, N (2020) 'Composition in the Age of AI', *Ideas Sonicas*, 12(23).

Collins, T and Laney, R (2017) 'Computer-Generated Stylistic Compositions with Long-Term Repetitive and Phrasal Structure', *Journal of Creative Music Systems*, 1(2). doi: 10.5920/JCMS.2017.02.

Diakopoulos, N and Johnson, D (2020) 'Anticipating and addressing the ethical implications of deepfakes in the context of elections', *New Media & Society*, p. 1461444820925811. doi: 10.1177/1461444820925811.

Dinan, E *et al* (2020) 'Queens are Powerful too: Mitigating Gender Bias in Dialogue Generation', *arXiv:1911.03842 [cs]*. Available at: <http://arxiv.org/abs/1911.03842> (Accessed: 16 July 2021).

Dockterman, E (2016) *Let's Talk About Grand Moff Tarkin in 'Rogue One'*, *Time*. Available at: <https://time.com/4604996/rogue-one-grand-moff-tarkin-princess-leia-cgi/> (Accessed: 16 July 2021).

Dutch Digital Design (2018) *The Next Rembrandt: bringing the Old Master back to life*, *Medium*. Available at: <https://medium.com/@DutchDigital/the-next-rembrandt-bringing-the-old-master-back-to-life-35dfb1653597> (Accessed: 16 July 2021).

Elder, A (2020) 'Conversation from Beyond the Grave? A Neo-Confucian Ethics of Chatbots of the Dead', *Journal of Applied Philosophy*, 37(1), pp. 73–88. doi: 10.1111/japp.12369.

Epstein, V (2021) *Text2Image Siren+.ipynb*. Available at: https://colab.research.google.com/drive/1L14q4To5rMK8q2E6whOibQBnPnVbRJ_7#scrollTo=JUvpydy8BWGuM (Accessed: 16 April 2021).

Finucane, B. P. (2019) 'Creating with blockchain technology : the “provably rare” possibilities of crypto art'. doi: 10.14288/1.0370991.

Flossmann, S, Grachten, M and Widmer, G (2009) 'Expressive Performance Rendering: Introducing Performance Context'. doi: 10.5281/ZENODO.849619.

Foscarin, F *et al* (2020) 'ASAP: a dataset of aligned scores and performances for piano transcription.', *21st International Society for Music Information Retrieval Conference*, pp. 534–541.

Franceschet, M *et al* (2019) 'Crypto art: A decentralized view', *arXiv:1906.03263 [cs]*. Available at: <http://arxiv.org/abs/1906.03263> (Accessed: 23 March 2021).

Frid, E, Gomes, C and Jin, Z (2020) 'Music Creation by Example', in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20: CHI Conference on Human Factors in Computing Systems*, Honolulu HI USA: ACM, pp. 1–13. doi: 10.1145/3313831.3376514.

Gorrostieta, C *et al* (2019) 'Gender De-Biasing in Speech Emotion Recognition', in *Interspeech 2019. Interspeech 2019*, ISCA, pp. 2823–2827. doi: 10.21437/Interspeech.2019-1708.

Gotterbarn, DW *et al* (2018) 'ACM Code of Ethics and Professional Conduct'. Association for Computing Machinery. Available at: <https://www.acm.org/code-of-ethics> (Accessed: 19 October 2018).

GPT-3 (2020) *A robot wrote this entire article. Are you scared yet, human? | GPT-3, the Guardian*. Available at: <http://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3> (Accessed: 16 April 2021).

Grachten, M and Krebs, F (2014) 'An assessment of learned score features for modeling expressive dynamics in music', *IEEE Transactions on Multimedia*, 16(5), pp. 1211–1218.

Grachten, M and Widmer, G (2011) 'Explaining musical expression as a mixture of basis functions', in *Proceedings of the 8th Sound and Music Computing Conference (SMC 2011)*.

Greenwald, AG (2017) 'An AI stereotype catcher', *Science*, 356(6334), pp. 133–134. doi: 10.1126/science.aan0649.

Grow, K and Grow, K (2021) 'In Computero: Hear How AI Software Wrote a “New” Nirvana Song', *Rolling Stone*, 2 April. Available at: <https://www.rollingstone.com/music/music-features/nirvana-kurt-cobain-ai-song-1146444/> (Accessed: 21 July 2021).

Hadjeres, G (2021) *The Piano Inpainting Application: An A.I.-enhanced piano performance generation Ableton plugin*. Github. Available at: <https://ghadjeres.github.io/piano-inpainting-application/>.

Hanson, D *et al* (2021) 'SophiaPop!: Experiments in Human-AI Collaboration on Popular Music', p. 7.

Harbinja, E, Edwards, L and McVey, M (2021) *Chatbots that resurrect the dead: legal experts weigh in on 'disturbing' technology*, *The Conversation*. Available at: <http://theconversation.com/chatbots-that-resurrect-the-dead-legal-experts-weigh-in-on-disturbing-technology-155436> (Accessed: 16 July 2021).

Hsu, J (2012) *Why 'Uncanny Valley' Human Look-Alikes Put Us on Edge*, *Scientific American*. Available at: <https://www.scientificamerican.com/article/why-uncanny-valley-human-look-alikes-put-us-on-edge/> (Accessed: 16 July 2021).

Huang, C-ZA *et al.* (2018) 'Music Transformer', *arXiv:1809.04281 [cs, eess, stat]*. Available at: <http://arxiv.org/abs/1809.04281> (Accessed: 23 March 2021).

J Walter Thompson Amsterdam (2017) *The Next Rembrandt, The Next Rembrandt*. Available at: <https://www.nextrembrandt.com> (Accessed: 16 July 2021).

Kain, E (2020) 'Star Wars' Deepfake 'Fixes' That Crazy 'Mandalorian' Season 2 Finale Cameo, *Forbes*. Available at: <https://www.forbes.com/sites/erikkain/2020/12/23/star-wars-fan-fixes-that-crazy-mandalorian-cameo-with-a-stellar-deepfake/> (Accessed: 16 July 2021).

Keats, J (2020) 'Film Photography Can Never Be Replaced', *Wired*, 21 April. Available at: <https://www.wired.com/story/film-photography-can-never-be-replaced/> (Accessed: 16 April 2021).

Kiessling, A (2021) 'Alex Kiessling: gallery', *Alex Kiessling*. Available at: <https://www.alexkiessling.com/gallery/> (Accessed: 15 April 2021).

King, T (2019) *Human-Like Playtesting with Deep Learning*, *Medium*. Available at: <https://medium.com/techking/human-like-playtesting-with-deep-learning-92adafffe921> (Accessed: 15 April 2021).

Knotts, S and Collins, N (2020) 'A Survey on the uptake of Music AI Software', in *Proceedings of the International Conference on New Interfaces of Musical Expression. New Interfaces of Musical Expression*, Birmingham, United Kingdom: NIME, pp. 499–504.

Larsson, S. and Heintz, F. (2020) 'Transparency in artificial intelligence', *Internet Policy Review*, 9(2). doi: 10.14763/2020.2.1469.

Liebman, E and Stone, P (2020) 'Artificial Musical Intelligence: A Survey', *arXiv:2006.10553 [cs, eess]*. Available at: <http://arxiv.org/abs/2006.10553> (Accessed: 7 April 2021).

Lost Tapes of the 27 Club (2021). Available at: <https://losttapesofthe27club.com/> (Accessed: 21 July 2021).

Louie, R *et al* (2020) 'Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models', in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. CHI '20: CHI Conference on Human Factors in Computing Systems*, Honolulu HI USA: ACM, pp. 1–13. doi: 10.1145/3313831.3376739.

MacCarthy, M and Propp, K (2021) *Machines Learn That Brussels Writes the Rules: The EU's New AI Regulation*, *Lawfare*. Available at: <https://www.lawfareblog.com/machines-learn-brussels-writes-rules-eus-new-ai-regulation> (Accessed: 21 July 2021).

Martinez, FJ (2019) 'AI and Creativity—Is It Ethical; Will It Kill Creativity?', *Ethics in Graphic Design*, 3 June. Available at: <http://www.ethicsingraphicdesign.org/ai-and-creativity-is-it-ethical-will-it-kill-creativity/> (Accessed: 10 March 2021).

Meskys, E *et al* (2019) *Regulating Deep Fakes: Legal and Ethical Considerations*. SSRN Scholarly Paper ID 3497144. Rochester, NY: Social Science Research Network. Available at: <https://papers.ssrn.com/abstract=3497144> (Accessed: 23 April 2021).

MIT Media Lab (2017) *Deep Empathy*. Available at: <http://deepempathy.mit.edu> (Accessed: 16 July 2021).

Ncube, CB and Oriakhogba, DO (2018) 'Monkey Selfie and Authorship in Copyright Law: The Nigerian and South African Perspectives', *Potchefstroom Electronic Law Journal*, 21, pp. 1–35. doi: 10.17159/1727-3781/2018/v21i0a4979.

Neilan, D (2018) *These maniacs will not stop until Nic Cage is in every movie*, *The A.V. Club*. Available at: <https://www.avclub.com/these-maniacs-will-not-stop-until-nic-cage-is-in-every-1829302082> (Accessed: 16 July 2021).

Roberts, A *et al* (2019) 'Magenta Studio: Augmenting Creativity with Deep Learning in Ableton Live'. doi: 10.5281/ZENODO.4285266.

Rosner, H (2021) *The Ethics of a Deepfake Anthony Bourdain Voice*, *The New Yorker*. Available at: <https://www.newyorker.com/culture/annals-of-gastronomy/the-ethics-of-a-deepfake-anthony->

bourdain-voice (Accessed: 20 July 2021).

Scribner, C (2000) *Peter Paul Rubens* | *Biography, Style, & Facts*, *Encyclopedia Britannica*. Available at: <https://www.britannica.com/biography/Peter-Paul-Rubens> (Accessed: 23 March 2021).

Smith, AE (2021) *GANs + Sanrio = Ganrio!*, *Medium*. Available at: <https://amyelizabethsmith01.medium.com/gans-sanrio-ganrio-21e263666929> (Accessed: 16 April 2021).

Stoffregen, TA *et al* (2008) ‘Motion Sickness and Postural Sway in Console Video Games’, *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(2), pp. 322–331. doi: 10.1518/001872008X250755.

Suciu, P (2020) *Deepfake Star Wars Videos Portent Ways The Technology Could Be Employed For Good And Bad*, *Forbes*. Available at: <https://www.forbes.com/sites/petersuciu/2020/12/11/deepfake-star-wars-videos-portent-ways-the-technology-could-be-employed-for-good-and-bad/> (Accessed: 16 July 2021).

Tangcay, J (2021) *Anthony Bourdain Voice Re-Creation in Documentary Raises Controversy*, *Variety*. Available at: <https://variety.com/2021/artisans/news/anthony-bourdain-fake-voice-roadrunner-documentary-backlash-1235020878/> (Accessed: 16 July 2021).

The Method Case (2013) ‘Long Distance Art One artist, two robots and three paintings’, *THE METHOD CASE*, 23 November. Available at: <https://www.themethodcase.com/video-long-distance-art-one-artist-two-robots-three-paintings/> (Accessed: 15 April 2021).

Toews, R (2020) *Deepfakes Are Going To Wreak Havoc On Society. We Are Not Prepared.*, *Forbes*. Available at: <https://www.forbes.com/sites/robttoews/2020/05/25/deepfakes-are-going-to-wreak-havoc-on-society-we-are-not-prepared/> (Accessed: 15 April 2021).

United States Department of Labor (2020) *OSHA Technical Manual (OTM) - Section IV: Chapter 4 | Occupational Safety and Health Administration*. Available at: <https://www.osha.gov/otm/section-4-safety-hazards/chapter-4> (Accessed: 16 April 2021).

Vallor, S (2016) *Technology and the virtues: a philosophical guide to a future worth wanting*. New York, NY: Oxford University Press.

Vear, C (2020) ‘Creative AI and Musicking Robots’, in: *1st Workshop on Creativity and Robotics, 12th International Conference on Social Robotics (ICSR 2020)*, Colorado, USA. Available at: <https://dora.dmu.ac.uk/handle/2086/20477> (Accessed: 20 July 2021).

Wareing, J (2021) ‘I made GPT-3 read every @emfcamp talk title and then asked it to write new ones. A thread. <https://t.co/M6xZgmyuCN>’, @jonty, 5 March. Available at: <https://twitter.com/jonty/status/1367932738691596295> (Accessed: 16 April 2021).

Whitwam, R (2020) *New Tool Generates Every Possible Melody for Public Domain Use*, *ExtremeTech*. Available at: <https://www.extremetech.com/extreme/306575-new-tool-generates-every-possible-melody-for-public-domain-use>.

Widmer, G (2002) ‘Machine Discoveries: A Few Simple, Robust Local Expression Principles’, *Journal of New Music Research*, 31(1), pp. 37–50. doi: 10.1076/jnmr.31.1.37.8103.

Widmer, G and Goebel, W (2004) ‘Computational Models of Expressive Music Performance: The State of the Art’, *Journal of New Music Research*, 33(3), pp. 203–216. doi: 10.1080/0929821042000317804.

Wikipedia (2021) *Ship of Theseus*. Available at: https://en.wikipedia.org/wiki/Ship_of_Theseus#Thought_experiment (Accessed: 4 July 2021).

Wolf, MJ, Miller, KW and Grodzinsky, FS (2017) ‘Why We Should Have Seen That Coming: Comments on Microsoft’s Tay “Experiment,” and Wider Implications’, *The ORBIT Journal*, 1(2), pp. 1–12. doi: 10.29297/orbit.v1i2.49.

Yin, Z *et al* (2021) ‘“A good algorithm does not steal – it imitates”: The originality report as a means of measuring when a music generation algorithm copies too much.’, in *Proceedings of EvoMUSART. EvoMUSART 2021*, (in press).

Zhou, Y *et al* (2020) ‘Generative Melody Composition with Human-in-the-Loop Bayesian Optimization’. doi: 10.5281/ZENODO.4285364.