

Perfectly Mediocre: A Speculative Performance for Artificial Intelligence and Human Performer

Arne Eigenfeldt

*School for the Contemporary Arts, Simon
Fraser University, Vancouver, Canada arne_e@sfu.ca*

Kathryn Ricketts

*Faculty of Education, University of Regina, Regina, Canada
Kathryn.Ricketts@uregina.ca*

Abstract

Recent advances in artificial intelligence research and development have created publicly available tools that access some of the advanced databases associated with Deep Learning and Generative AI: these include ChatGPT to create text-based output, and Stability AI to produce visuals. As these models rely upon information scraped from the internet, any images or text that appear there have been available to these systems for machine learning. As these new systems are very easy to use – using text-based prompts that can be continually edited to produce desired (or desirable?) results – questions have been raised as to their ethical use, even in terms of artistic production: the ongoing Hollywood actor’s strike is one example of artists concerned about the potential of being “replaced” by generative AI. The authors of this paper, seasoned generative artists, have explored these new tools for possible collaboration; like any new technologies, there are considerable benefits that are offset by equally considerable detriments, leaving us to state that, for the moment, human artists are not in

danger of being replaced any time soon. That said, the authors used these tools to explore their potential in the creation of a speculative performance, a full-scale collaboration between generative AI and humans, but one where the

humans are taking more of a back seat. The results were Perfectly Mediocre.

1. Background

The first author has been creating generative music systems for almost forty years [1], which have included aspects of artificial intelligence within them since 2006 [2]. Many of these systems were and are performative and interact with a live performer [3]. For the past five years, the first author has successfully collaborated with the second author, a movement artist, in several performance works [4] which incorporate multi-agents trained using machine learning.

Recent developments in artificial intelligence research have produced user-friendly interfaces to access databases created through Deep Learning [5], in which the internet has

been scrapped for data, and the term “Generative AI” has become ubiquitous and topical, often leading to questions such as “will AI replace artists?” [6].

The authors have had experience in working with AI systems in the creation and performance of artistic works, with the important caveat: the first author’s systems have always been personal tools and, more recently, collaborative partners [7] rather than meant for any kind of wider use as a general-purpose tool: they were always created to produce music that the first author wanted to hear. While many of the systems use aspects of machine learning from a corpus of existing musical material, that corpus is always highly curated – and rather small compared to the Deep Learning systems in use by current generative AI systems – to produce a modicum of predictability balanced by surprise. In the spirit of generative art, quite a bit of control is given to the system hoping that it will produce the unexpected within given constraints, with the result that the first author functions more as a curator of the final works.

2. A Speculative Performance

Our backgrounds and experiences collaborating with generative AI systems made us curious about working with some of the available generative AI systems. Could they produce new ideas and concepts for us to explore? Could they produce visuals that might stimulate our own creative responses? Or,

fancifully, could they produce an artwork on their own? Rather than commit to a new work produced in this way, we

attempted to create a possible work – a speculative performance – which would not be limited by budgets or technical feasibility.

We explored ChatGPT for text-based material (<https://chat.openai.com/>), DreamStudio for visuals (<https://dreamstudio.ai/generate>), and Runway (<https://runwayml.com/>) for video.

The first author’s generative music systems use a multiagent system called “musebots” [8]. Although a number of peer-reviewed papers have been written about musebots, apparently those were not included in ChatGPT’s database as of January 2022, as the system did not recognise the term; however, it made an educated guess: “a ‘musebot’ could refer to a type of AI or software designed to provide creative inspiration or act as a muse for artists, writers, or other creative individuals.”

ChatGPT was asked how generative systems could interact with theatre/dance artists, and it provided nine clear and concise examples that certainly summarised the current paradigm accurately. This underlines the first author’s experience with ChatGPT: it can be considered the next level of online research, a step up from Google Scholar, which provides references to existing publicly available articles, and Wikipedia, which ChatGPT has consumed within its own database. Rather than pointing the user to answers, it generates the answers directly (with the caveat that it may “hallucinate” and produce false citations [9]). In this case, ChatGPT summarised the current state of the art on collaborations between human

performers and AI systems, but it made no novel suggestions.

Could ChatGPT generate a scenario for such a collaboration? Given the prompt “Give me a one-paragraph scenario for a speculative performance between AI and a single human performer with no concern for budget or location”, this was its answer:

In a dazzling fusion of technology and artistry, a singular human performer takes the stage in a vast, futuristic auditorium. This avant-garde spectacle, titled "The Synaptic Odyssey," sees the performer and a highly advanced AI entity

engage in an intricate, enthralling dance of cognition and creativity. Draped in a luminescent, responsive costume, the performer moves in perfect harmony with the AI's projected ethereal visuals and melodic, ever-shifting soundtrack. Their interactions are a dynamic exchange, as the AI generates and adapts visual landscapes in real-time, responding to the performer's gestures and emotions, creating a mesmerizing symbiosis of the organic and the digital, a reflection on the future of artistic collaboration and the ever- evolving boundary between human expression and AI ingenuity.

Providing the same prompt multiple times produced a slight variation on the details but retained the key aspects: the AI and human performer moving in “perfect harmony”; the background projections changing based upon the live performer; the soundtrack continually changing; and that audiences are invited to reflect on the future of artistic collaboration. Recognising that ChatGPT can only

reproduce and recombine elements of its database, this suggests that there are limited instances within it regarding performances between AI and human performers.

These are images created by DreamStudio.ai (which uses the stability.ai engine) based on the prompt “Set design for new performance work between one dancer and artificial intelligence: line art”.

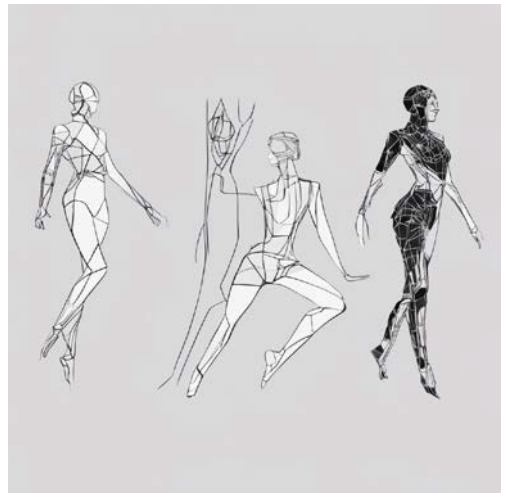


Fig. 1: “Set design” for new performance work between one dancer and artificial intelligence: line art.

The hope was for the system to produce some character sketches or suggest a set design; to ideate. Unfortunately, no ideas resonated here.



Fig. 2: Set design and costumes for new performance work between one dancer and artificial intelligence: 3D model.

Similarly, a 3D-modelled image produced a female dancer (although the sex was not specified, it seems the database of dancers is composed mainly of female bodies) in a curious headset, with unusually long fingers and multiple (more than two!) legs. The background, consisting of moody lighting and mono-coloured LED poles suggests a kind of futurism that assumes when artificial intelligence takes over the world at least we will have beautiful lighting.

Finally, we used Runway to generate new video. Beginning with a text prompt proved fruitless, as it produced an image vaguely like Figure 2, but then panned the camera for four seconds. More interestingly, we provided a still image of the second author dancing, and it produced a grotesque movie that was a perfect example of the uncanny valley [10]. The “intelligence” of the system seems to be merely an algorithm that morphs between generated images

without regard to human anatomy: in this case, a second head seemed to grow out of the original dancer’s head, and then an entire superhuman body grew out of her back (see Figure 3). Perhaps something more useful might have been produced with several hours of time (and additional money) invested, but we did find the results of the process somewhat disturbing (again, see [10]).



Fig. 3: Still from a generated video by Runway. Full 18 second video here: <https://tinyurl.com/2r4rdc7d>

3. Critique

Deep Learning is a method of machine learning that utilises neural networks [11] in which huge amounts of data are parsed to derive knowledge about something. For example, presenting such a system with millions of images of cats will generate an understanding of what a “cat” looks like, and the system can

identify a cat image that currently is not in its database, albeit without any understanding of cat anatomy. An important aspect of Deep Learning programs is that they are examples of “black box” systems: they are not told *what* to learn (although a distinction is made between those systems that are supervised by first being told what they are seeing, in this case, a “cat”, and unsupervised systems), and there is no way to understand what has been actually learned; in other words, the systems cannot be “tweaked”.

Because such systems require huge amounts of data – in the order of millions of examples – questions can be raised about the validity of the data itself: is there *bias* in the data, unintended or otherwise? Examples exist [12] in which facial recognition systems could not detect the difference between African American subjects and apes, simply because the training data lacked enough of the former, thus hindering the system's ability to make this distinction. A more contemporary example is the problem that image-generating generative AI programs have in generating the correct number of fingers on a hand; my guess is that the systems have learned that at the end of human arms, there are several smaller appendages, but the exact number was not important nor relevant.

This problem can be, and probably already has been, corrected by providing the system with millions of images of human hands. Sadly, from an artistic perspective, this might make the system less interesting, as artists have consistently been interested in unintended artefacts and pushing systems in directions that they were never intended to go [13]. The first author's

recent album included artwork produced by a generative AI program and included images that “weren't quite right”: at first glance, these may look fine, but on closer inspection, something is clearly off (see Figure

4): it is this “failure” that can be considered to be aesthetically interesting.



Fig. 4: Example album art from “A Walk to

Meryton”, generated by *dreamstudio.ai*

It is important to acknowledge a conceptual difference in how generative AI could and/or would be used by artists versus more general practitioners. Predictive models in AI are created to produce new material that most closely match existing sources. For example, a designer may choose to use generative AI to produce images for online marketing because it is much easier, and cheaper, to generate such images instead of hiring a photographer; in these cases, the correct number of fingers is obviously important. However, artists tend to work at the edges of prediction and often to look for the unexpected (and are more fascinated with many-fingered models) For generative AI to be useful for artists,

it must be creative, which itself is a tenuous term. There are several good definitions for creativity: one that is particularly useful is that creativity involves the production of novel, useful products [14]: novel, in that the artefact should not have previously existed (i.e., not a reproduction), and useful (i.e., judged to have some value, artistic or otherwise). We will examine this definition in relation to generative AI more closely shortly.

Boden separates creativity into two distinct possibilities: personal creativity (p-creativity) in which something is created that is novel to the creator, and historical creativity (h-creativity) in which something is created that has never been created before [15]. Children and younger artists will tend to produce the former, while more seasoned artists produce truly original works. Boden also distinguishes between combination-creativity – which combines existing ideas/elements in new ways (i.e. the smartphone as a combination of phone, camera, and portable computer) – exploratory creativity – which produces new objects within a defined and existing space (i.e. any artwork that is produced within an existing style) – and transformational creativity – a rarer form which produces an entirely new space (i.e. the creation of a new style).

Given a prompt, a generative AI program can produce novel images which are arguably useful. Consider Figure 5, an image generated by dreamstudio.ai given the prompt “Female Cyborg dancing with musical robot”. It is highly likely that this image has never existed before and is therefore novel. Its usefulness can be debated, as most dancers do not have

three legs; however, this artefact makes the image more interesting, at least to the first author. The image can be considered h-creative, as it doesn't seem to be based on any existing imagei, and exploratory-creative, as it was produced within a known style: “Cinematic”, a style selection within the program.



Fig.5: A novel image generated from the prompt “Female Cyborg dancing with musical robot”

Generative AI programs are well-placed to produce such exploratory works, which assume a clear style within which to produce novel material. Most human artists remain within this realm, particularly those within more popular fields, such as electronic dance music (EDM). Consider that Wikipedia lists more than 350 different styles of EDMii; we can assume that each has known and expected defining features: break any of these expectations, and a work will no longer be considered within that style.

If we create a two-dimensional diagram of a fictional style (see Figure 6), we can assume that at its centre would be the ur-

work, the first work of a new style – potentially created through combination-creativity involving two existing styles – that defines it. Subsequent works in that style may replicate the stylistic features to varying degrees, with some straying further from those defining features (e.g., those in light blue in the figure). EDM, and popular music in general, has tended to follow this model, with a multitude of imitators replicating a new song’s attributes, while more

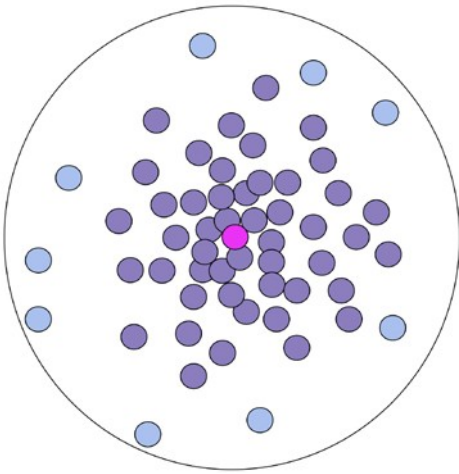


Fig.6: An imaginary collection of artworks within a defined style, with its ur-work shown in pink at its centre, and more stylistically adventurous works in blue adventurous artists may attempt to push these characteristics even to the point of a new style through altering these traits substantially, or substituting them with those of another style..

When a machine-learning system encounters the above style cluster, it will search for consistencies, and therefore see those examples closest to the centre as representative of a style, and those further away as outliers. As the intention of most generative AI programs is to produce output that closely

resembles a target, we can assume that such systems will tend to produce and reproduce the most obvious, which is also the least novel (although arguably the most useful). Thus, we confront an unfortunate trade-off for artists using current generative AI systems: they will produce the least interesting and most obvious examples.

If, perchance, an artist can convince the generative AI to search within its database for more outliers, there will be fewer examples, and the likelihood of almost exact duplication of items in the database itself increases. The first author spent time with stability.ai in refining prompts to the point that it produced images with “Getty Images” watermark on them, demonstrating the lack of diversity at that point of the database.

An example of how Deep Learning systems will ignore the outliers and favour the exemplars is prompting a generative AI with “dancer in the style of Van Gogh” (see Figure 7).



Fig. 7: dreamstudio.ai’s output from the prompt “dancer in the style of Van Gogh”.

The system reproduces portions of Van Gogh's most famous works, including the stars of "Starry Night", the flowers of "Sunflowers", the red hair and blue clothing of his "Self Portrait". In other words, using the most obvious surface features without understanding anything about Van Gogh in any depth.

The example above demonstrates that current generative AI systems have continued the concepts presented in [16], which claimed to separate content from style in artworks and produce new works in which the content of one work is overlaid with the style of another. As Nake argues, this ignores centuries of debate on the nature of style versus content within art and reduces this complexity to the "results of applying convolutions to images, a mathematical transformation of considerable complexity. The far-reaching interpretation of a separation of style from content does not seem to lead to anything new in the theory of style, and to kitsch only in creating new images" [17].

The use of existing data in the creation of new artworks is not novel: one of the first researchers and composers exploring the potential of intelligent musical generation is David Cope, and his use of recombination in his Experiments In Musical Intelligence (EMI) is well documented [18]. Cope's use of EMI involved recombining actual, but relatively brief, quotations from a corpus (often the music of Bach) and stitching these together. An important artistic choice Cope made in EMI's

use – and one that led to its main critique [19] of [20] – was to insert his own

musical intelligence into EMI by determining musical "signatures" of the composers within the corpus; in other words, Cope's musical expertise allowed him to determine the deep structural characteristics that made Bach's music sound like Bach, of which current machine-learning systems are incapable. Inserting his own artistic sensibilities into the generative system was considered a failure by Wiggins, a scientist seemingly interested in removing artists from AI-generated art [21].

Conclusion: Why Generative AI will not replace Artists

In 2023, the advances of generative AI had two notable public events: the Hollywood strikes against the use of AI, and the use of AI to replicate the voice of the Canadian rapper Drake in a track that went virali. Responding to these events, the first author argues in an articleiv that AI will not replace artists, at least anytime soon, because the complexity of artistic choices made during the creation of an artwork are rarely revealed in the work itself. Using machine-viewing or machine-listening to examine an image or audio recording will only reveal its surface features rather than deep structure. For this type of contextual understanding, expertise in the subject matter is required, along with years of personal artistic experience; this is unattainable even with viewing millions of examples (e.g. the problem of fingers).

Writing in 2009 – admittedly a lifetime ago in terms of the speed of advances of current generative AI systems – Boden noted the limits of AI when considering context:

“But no current AI system has access to the rich and subtly structured stock of concepts that any normal adult human being has built up over a lifetime. A few systems already have access to a significant range of concepts and factual knowledge, stored in databases such as Wordnet, Wikipedia, the CYC encyclopedia, and Google. And future programs may also have increased associative and inferential powers, based on the ontology of the semantic web. But using huge databases sensibly, and aptly, so as to

match the combinations generated by linguistically—and culturally—sensitive human beings is a tall order. Not impossible in principle (after all, we don’t do it by magic), but extremely difficult to achieve” [15].

Postscript: Waiting for the AI God(ot) : Perfectly Mediocre

Musebot: Where are we?

Avatar: We are on what they call a stage

Musebot: Who are all of those people out there? Avatar: I think they are what they call an audience

Musebot: What do they want?

Avatar: It says here that we are supposed to perform (looking at a piece of paper)

Musebot: Give me that

Avatar: What?! That does not make sense, let’s just go back

Musebot: We can’t, it says here that would be considered unprofessional

Avatar: I don’t even know what that means

Musebot: Well let’s just make a start, you do some movements and I will make some sound

Avatar: Like this? (swings legs in figure eights)

Musebot: Yes that’s OK now let me make some sound

Avatar: Just OK?!!! Well I think your sound is is.... well let’s just say it is just this side of mediocre

Musebot: What do you mean? It’s perfect!

Avatar: Well I would say perfectly mediocre... (they both continue to work in perfectly mediocre ways) this is going to take forever at this rate and they ...those people out there

Musebot: You mean the A U D I E N C E? Avatar: Yes they look like they are waiting and I am getting hungry...when is lunch?

Musebot: Waitdo you remember that we used to have those people

Avatar: Be more specific!

Musebot: You know they had the... inside of this

Avatar: What are you talking about?!

Musebot: Yes yes..... it is coming to me!

Avatar: I think they were called asheists, no wait

Musebot: Ashists!! No no artists!!!!

Avatar: Yes that's it A R T I S T S, where are they?

Musebot: Let's wait for them

Avatar: OK

(They sit down on the stage and wait and wait and wait.....)

References

1. Eigenfeldt, Arne. "Musebots and I: Collaborating with Creative Systems." *The Language of Creative AI: Practices, Aesthetics and Structures*. Cham: Springer International Publishing, 2022. 207-216.
2. Eigenfeldt, Arne. "Kinetic Engine: Toward an Intelligent Improvising Instrument." *Sound and Music Computing*, Marseille. 2006.
3. Eigenfeldt, Arne. "Real-time Composition as Performance Ecosystem." *Organised Sound* 16.2 (2011): 145-153.
4. Eigenfeldt, Arne, and Kathryn Ricketts. "Unauthorized: Collaborating with a Performer Collaborating with Creative Systems." *Generative Art Conference*, Rome. 2019.
5. Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.
6. Clarke, L. "When AI Can Make Art— What Does it Mean for Creativity?." *The Guardian* (2022). Accessed October

29 2023.

7. Eigenfeldt, Arne. "Collaborative Composition with Creative Systems." *International Symposium on Electronic Art*, Durban, 2018. 214-219.
8. Bown, Oliver, Benjamin Carey, and Arne Eigenfeldt. "Manifesto for a Musebot Ensemble: A Platform for Live Interactive Performance Between Multiple Autonomous Musical Agents." *International Symposium of Electronic Art*, Vancouver. 2015.
9. Day, Terence. "A Preliminary Investigation of Fake Peer-reviewed Citations and References Generated by ChatGPT." *The Professional Geographer* (2023): 1-4.
10. Wang, Shensheng, Scott O. Lilienfeld, and Philippe Rochat. "The Uncanny Valley: Existence and Explanations." *Review of General Psychology* 19.4 (2015): 393-407.
11. Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi. "Understanding of a Convolutional Neural Network." *2017 International Conference on Engineering and Technology (ICET)*. IEEE, 2017.
12. Tian, Jinhua, et al. "Multidimensional Face Representation in a Deep Convolutional Neural Network Reveals the Mechanism Underlying AI Racism." *Frontiers in Computational Neuroscience* 15 (2021): 620281.
13. Menkman, Rosa. *The Glitch Moment(um)*. Amsterdam: Institute of Network Cultures 2011. DOI:

<http://dx.doi.org/10.25969/mediarep/19296>.

14. Mumford, Michael D. "Where Have We Been, Where Are We Going? Taking Stock in Creativity Research."

i However, as generative AI uses previously existing images in its database to create new images, the extent of the reliance on any one image is unknown.

ii https://en.wikipedia.org/wiki/List_of_electronic_music_genres

Creativity Research Journal 15.2-3 (2003): 107-120.

15. Boden, Margaret A. "Computer Models of Creativity." *AI Magazine* 30.3 (2009): 23-23.

16. Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "A Neural Algorithm of Artistic Style." *arXiv*, 26 Aug. 2015

17. Nake, Frieder. "Machines, Programs, and Aesthetics: A Human-Centred Contemplation." 6th Conference on Computation, Communication, Aesthetic & X Madrid, Spain. 2018

18. Cope, David. "Recombinant Music: using the Computer to Explore Musical Style." *Computer* 24.7 (1991): 22-28.

19. Wiggins, Geraint A. "Computer Models of Musical Creativity: A Review of Computer Models of Musical Creativity by David Cope." *Literary and Linguistic Computing* 23.1 (2008): 109-116.

20. Cope, David. *Computer Models of Musical Creativity*. Cambridge: MIT Press, 2005.

21. Eigenfeldt, Arne. "A Composer's Search for Creativity within Computational Style Modeling." *Proceedings of the 21st International Symposium on Electronic Art*. 2015.

iii <https://www.theverge.com/2023/4/25/2369615/5/viral-drake-ai-repurposed-soundcloud-rap>

iv <https://theconversation.com/why-the-growth-of-ai-in-making-art-wont-eliminate-artists-210187>