Tool, Collaborator, or Participant: AI and Artistic Agency

Anthony Cross, Texas State University

anthony.cross@txstate.edu

1. AI Takes the Prize

In late August 2022, Jason Allen won first place in the Colorado State Fair's amateur digital arts competition for his artwork, *Théâtre D'opéra Spatial*. (Fig 1.) Ordinarily, such an announcement would garner little attention: there were only 21 entrants in the category, and Allen won a modest cash prize of $300.

However, it soon emerged that Allen's piece had been created using an artificial intelligence (AI) image generation tool, Midjourney. As Allen describes it, he was blown away by the quality of the images he was able to generate using Midjourney; the tool seemed up to the task of generating a visualization of any textual prompt that he could imagine. Speaking to the *New York Times*, Allen described his motivation: "The fair was coming up and I thought: How wonderful would it be to demonstrate to people how great this art is" (Roose 2022)? He selected one of his favorite creations, had it printed on canvas by a local shop, and submitted it to the fair's competition. To his surprise, he won. In his enthusiasm, he posted about his victory on Midjourney's Discord chat; the news soon spread to social media and then went viral.

The response to Allen's award was overwhelmingly negative. Some dismissed his work entirely, claiming that it—along with most other AI-generated images—simply wasn't art at all. This critique hardly seems fair. As others have convincingly argued, there is a long history of artists using technology to assist in the creation of their art, from photography to cinema to procedural art and readymades (Hertzmann 2018; Trujillo 2022). Consider Duchamp's *Fountain*: a readymade urinal, made in a factory, signed by the artist and submitted for exhibition more than 100 years ago. If *this* is to count as art—and, in fact as a touchstone of 20th century art—then there should be no problem in thinking Allen's AI-generated art should clear the bar as well.

A more common complaint was that Allen's submission was somehow deceptive. The judges of the competition were in fact unaware that Allen had used Midjourney to generate the image. Perhaps Allen was guilty of what Dennis Dutton (1979) would refer to in the context of discussing artistic forgeries as a *misrepresentation of achievement*: Allen presented the work as though it were the result of his own craft, the painstaking application of (digital) paint to canvas, whereas it was instead the result of his feeding textual prompts to an AI image generator. Such deception would be objectionable in the same way as forgeries are objectionable. However, it's debatable whether Allen actually intended to mislead the judges in this way. In submitting the work, Allen signed the piece "Jason M Allen via Midjourney." He never claimed that he had created the piece entirely on his own. He broke no competition rules in submitting an AI-generated image. And, finally, the judges agreed that, even if they had known that the piece had been generated using AI, they still would have awarded it first prize.

In his discussion of forgeries, Dutton is very clear on one point: creating a forgery *is* a kind of achievement—just not the kind of achievement that we might have initially thought it was. The same is no doubt true of AI-generated art; Allen's work *is* an artistic performance, and it is to some degree an achievement. But, as Dutton (1979: 309) puts it, "the fundamental question then is, what has the artist done, what has he achieved?" This is a question which, for AI-generated art, has hardly even been asked.

In this paper, I aim to sketch the beginnings of an answer. What exactly is it that AI artists are doing? To put a finer point on the question: How do the affordances of generative AI tools like Midjourney create new opportunities for artistic agency? What does this agency look like, and what role does generative AI play in contributing to it? What might constitute an artistic achievement in creating AI art?

Perhaps unsurprisingly I will argue that there is no single answer to each of these questions, largely because artists are using generative AI to make art in a range of different ways. However, I think that there are several characteristic patterns of agency commonly involved in AI art, each of which involves a different mode of interaction with generative AI, and which correspondingly involves different opportunities for artistic agency and achievement on the part of the (human) artist. After a brief technical primer below, I map out two such patterns by introducing two different

paradigms of AI art which I refer to as the *production* paradigm and the *exploration* paradigm. As I argue in the final sections of the paper, appreciating this point about the diversity of AI art practice has important implications for a broad array of questions concerning the interpretation and appreciation of AI art—not least of which is the conclusion that there may be no *one* artistic category of AI art, but rather several.

2. A Technical Primer

Computer-generated art is hardly headline news; early practitioners of computer art like Frieder Nake and George Nees were using algorithms to generate visual art as early as the 1960s. However, what *is* new in the last decade is the radical expansion of the use of machine learning to generate compelling visual images. In a few short years, we've gone from the trippy outputs of Google's Deep Dream platform, released in 2015, to sophisticated (and award-winning) images like those generated by Allen.

Present-day AI image generators such as Midjourney and DALL-E make use of a variety of sophisticated models for generating images, including generative adversarial networks (GANs) and diffusion models (Cetinic and She 2022). What is perhaps most notable about the current crop of tools is that they generate images from text inputs. This is the result of the application of "transformer-based architectures" which, very briefly, work as follows: a text-encoder first encodes the text input prompt by mapping its contents to a multi-dimensional representational space—"latent space." The tools have been pre-trained to link the encoded text prompt to similarly-abstracted image encodings—to map the encoding of, e.g. "a green apple" to abstract mappings of images featuring green apples. Finally, this image encoding is decoded into an image using either diffusion models or GANs to generate novel images that "fit" the encoded data[1] (Ramesh et al. 2021).

---

[1] The differences between diffusion models and GANs are largely unimportant for the purposes of my argument. However, it's worth noting that most currently popular AI image generators rely on diffusion models, in part because of ability of these models to produce higher quality and more diverse output images. GANs such as StyleGAN are still popular as a means of generating a narrower class of images, e.g. realistic human faces. Thanks to an anonymous referee for encouraging me to discuss this difference.

The effectiveness of these tools is the result of training them on massive datasets of image-text pairs; this training allows the tools to create the latent space mappings between encoded text and encoded images discussed above. Previously such datasets were often custom prepared for AI training—a costly and time-consuming process. The ImageNet database, for example, contains over 15 million images organized around thousands of hierarchically categorized words and concepts—all reviewed and quality-controlled by human input. Assembling this database has taken more than a decade of work completed by undergrads and paid workers on Amazon's Mechanical Turk platform (Fei-Fei and Krishna 2022). Unlike previous models, the current crop of AI image generation tools rely on training data gathered directly from the internet without human input. For example, DALL-E relies on OpenAI's CLIP model (Contrastive Language-Image Pre-training). CLIP scrapes text-image pairs from image captions and alt-text online; this has resulted in a mapping tool trained on hundreds of millions of image-text pairs, dramatically increasing its performance (Radford et al. 2021).

To use these platforms to create new images, the user begins by typing in a string of text. After some delay, the user is then presented with several variations of the AI's output. The user can then adjust the initial string, generating new outputs, or instead select one of the initial outputs to use to generate further variations. By way of this iterative process, the user can then attempt to zero in on a final image. What is perhaps most striking about the process is its *ease*: it is possible to create hundreds of images quite quickly. This often facilitates a sort of playful exploration and generation; users are encouraged to iterate multiple times to see what the algorithm generates, exploring the effects of slight alterations to the prompt. By his own account, Jason Allen spent hours developing and refining the prompt he used to generate *Théâtre D'opéra Spatial*—a prompt which he considers to be central to his work, and which he has to this point declined to share with the public.

3. The Production Paradigm

What exactly is it that AI artists are doing when they engage with an AI image generator in the way I've just described? What kinds of agency is involved, and what are the prospects for creativity and artistic achievement? As appreciators, where should we direct our attention and engagement?

Let's begin with what I'll call the *production* paradigm of AI art, which I take to be the dominant understanding of AI art both in common discussions of the topic, as well as in existing philosophical and critical discourse. According to this paradigm, human artists make use of generative AI as either a *tool* or a *collaborator* in the generation of an artistic product. As I'll discuss below, there isn't consensus as to which role AI plays in the process—that of tool or collaborator—but either option is compatible with the production paradigm. What matters, ultimately, is that AI contributes to a creative process guided by a human artist which is ultimately oriented towards the *production* of some artistic output. It is the output of this process—the production itself—that is the site of artistic appreciation, and which may constitute a creative artistic achievement on the part of the artist.

The production paradigm is often motivated by way of analogy. First, consider cases in which artists make use of a technological *tool* to realize particular artistic products. Perhaps the most natural starting point in the context of thinking about AI art would be photography. A photographer uses her camera equipment to realize a particular kind of image. In principle, the photographer might realize the image through other means—by, say, creating a photorealistic drawing of the scene in question. But in this case, she creates an image indirectly, allowing what Henry Fox Talbot called the "pencil of nature" to capture the scene for her. Alternatively, consider a musician making use of a modular synthesizer to create a piece of music; by adjusting knobs and patching cables, she makes use of the synthesizer as a tool to create particular kinds of sounds that will feature in her composition. In each of these cases, the artist's creative agency guides a process oriented towards the production of an artistic output—a photograph, a musical performance—that incorporates the use of a tool.

As an alternative analogy, consider cases in which artists rely on *collaborators* to realize artistic products. Take the case of the studio assistant: painters from the Renaissance onwards have often employed assistants to help realize their artworks. Usually, artists delegate more mundane tasks like painting backgrounds or repetitive details. However, in some cases artists have delegated nearly the entirety of the creation of artworks to their assistants. For example, Jeff Koons has admitted to relying entirely on assistants for their technical skills in creating and fabricating art; his role, he has claimed, is that of the "idea person." One might also consider the case of film production. A film's director is generally held to be responsible for the overall production of the film. However, film directors

commonly delegate a number of key activities to collaborators who are themselves responsible for various aspects of the film such as lighting, photography, costuming, set design, and so on. What the examples of studio art and film production demonstrate is that artists often delegate some portion of the process of creating a final product to their artistic collaborators; however, as I've noted, there may be different levels of responsibility for the final work afforded to these collaborators. Some artists might attempt to exert complete creative control over the output of their collaborators, while others might allow their collaborators a significant degree of latitude in making artistic choices while reserving only the right to make overall decisions about the completion of the work (Livingston and Archer 2010).

What role does generative AI play in production-oriented AI art? Some have argued that generative AI is best understood as just another tool that may be enlisted by artists for the creation of images. Kathryn Wojtkiewitz makes this argument explicit:

> Generative AI programs are the latest in a long line of technology that facilitate new automatic image-making techniques. Like these previous technologies, generative AI programs change the way a creator participates in the creation of the image: where previously artists had to draw and paint from what they saw, the camera obscura allowed for tracing direct (albeit flipped) images of an existing scene; photographic cameras removed physical movements entirely, allowing for the artist to press a button and capture the image; photo screen-printing put the creator's actions at two removes, using mesh and paint to reproduce a photographic image; now a creator composes a string of words to produce the image (Wojtkiewicz 2024).

The tool analogy isn't perfect, however. Artists using generative AI seem to have less control over the process of image creation than the artist using a camera-obscura or a photographic camera. This is because text-to-image transformation algorithm are *opaque* to artists in ways that other image-generation tools are not. Consider a photographer who applies a filter to the lens of their camera; in principle, the photographer could know exactly what sorts of changes would result in the completed photograph as a result of changing the way they operate their image-generating tool. Now, compare this with the case of AI image generation. Although an artist might know in general terms what would result from, say, requesting a "photorealistic" image, they couldn't in principle know *exactly* what sort

6

of image would be the product of the prompt due to the way that the algorithm generates a novel transformation at each prompting.

This might push us towards thinking of generative AI as offering a more robust form of collaboration that goes beyond the contributions of a mere tool. On this understanding, AI functions with some degree of autonomy, offering a distinctive contribution to the creative process and that perhaps, as Claire Anscomb has argued, deserves some degree of production credit for the resulting artwork (Anscomb 2022). Of course, this proposal would face the challenges associated with attributing something like autonomy or creativity to generative AI—a proposal that many philosophers have rejected (Brainard 2023; Kelly 2019).

My point here is not to argue that one of these analogies is more convincing than the other; instead, what is common to both of them as manifestations of the production paradigm is that each pushes us towards a focus on the *output* of the creative process as the site of artistic appreciation. Consider Anscomb's overall appraisal of the contribution of AI to artistic production:

> Although AI agents are not themselves artistically creative, they can work iteratively without human intervention to non-accidentally generate the formal features of an image, and thus contribute to the realisation of some of the salient properties of a work….What has emerged is a picture where the interactions between AI and human agents are key to the realistion of the artistic value of the kinds of work under discussion (2022: 35).

The site of appreciation for Anscomb is the set of properties possessed by the output of the creative process; the resulting images *are* the artwork, and the artist's achievement lies in their ability to guide the creative process to its conclusion in the generation of these images.

It's fairly clear that this is the understanding that Jason Allen himself was operating under while creating *Théâtre D'Opéra Spatial*; what initially motivated his submission of the work for competition was the quality of the images he was able to create using Midjourney. Beyond this, his creative activity seemed focused on refining and improving the output itself; beyond the initial generation of the image using Midjourney, Allen also edited it using software tools such as photoshop and oversaw its printing. To evaluate Allen's work from the perspective of the production paradigm requires first that we scrutinize the resulting image: we attend to its qualities, looking at what it depicts and taking

7

in its formal properties. We might then ask what role Allen played in creating this image, and the extent to which the AI image generator he used—in this case Midjourney—deserves at least some credit for the output. I also think that widespread acceptance of the production paradigm accounts for a great deal of the outrage towards Allen's victory; because of Allen's reliance on an AI tool in the production of the image, the amount of artistic credit he deserves for it might appear to be significantly less than it would be had he not relied on generative AI at all.

4. The Exploration Paradigm

I'll now turn to an alternative paradigm of the artist's role in AI art—what I'll call the *exploration* paradigm. According to this paradigm, artists primarily interact with generative AI as a means of exploring and interrogating the way that the generative AI platform "sees" and "represents". What I think is most significant about this paradigm is that it shifts the site of artistic appreciation towards the artist's interaction with the AI and the way in which the artist structures this interaction. The products of this interaction—say, the images that result from the process—are not the primary object of appreciation, but rather function as a kind of documentation or evidence of the artist's exploration. I'll first elaborate on the paradigm, then illustrate it by way of both analogies with other art forms and examples of AI art that seems to focus on this sort of exploration.

Many AI artists understand their work as a kind of dialogue with an algorithm. What they are engaged in is a kind of interactive exploration in which the human artist, by way of adjusting their prompts and instructions, maps out the contours of the algorithm's representational space. Here is a characterization of this activity from two AI researchers, Eva Cetinic and James She:

> Most of the current AI Art works can be understood as results of sampling the "latent space." Perhaps the most novel aspect of AI Art is this possibility to venture into that abstract multi-dimensional space of encoded image representations….How one orchestrates the design of this space and what one finds in it, eventually becomes the major task and distinctive "signature" of the artist. In this context, it is important to understand the role of the human in this collaborative process with the machine (Cetinic and She 2022: 12).

What is centrally important about this characterization is that it is less centrally focused on the *output* of the AI image generator. Instead, what matters is the interaction between the artist and the algorithm: by adjusting inputs, iterating, and sampling, an AI artist is engaged in a process of mapping—and perhaps interrogating—the way that the algorithm sees and understands. AI art involves thematizing this process of mapping in a way that makes it a vehicle for meaning.

Consider, by way of analogy, one of the most well-known examples of performance art: Marina Abramović's *Rhythm 0*. The piece, initially performed in Naples in 1974, consisted of a set of objects placed on a table—ranging from flowers, wine, and bread to a gun, blades, and construction tools—along with a set of instructions provided to the audience. Roughly, the audience was instructed to use the objects on the table on Abramović herself in any way desired; the artist claimed full responsibility for any outcome of this process over the six hours of the performance. The outcome of the process was rather horrific: according to Abramović's own account, what began as mild interaction soon turned to attempted humiliation and assault as the audience removed her clothes, marked her body, cut her, and ultimately threatened her life (O'Hagan 2010). While it isn't my aim here to offer a nuanced interpretation of the piece, I think one doesn't have to reach far to see what Abramović's general aims were: by structuring her interaction with her audience through provided objects and instructions, she aimed to elicit particular sorts of participation from them. It was the participation itself that would illuminate issues of collective responsibility, susceptibility to suggestion and instruction, and the perceived division between art and life.

My suggestion is that many AI artists approach their interaction with generative AI in the same manner. By way of their selection of prompts, they elicit a sort of "participation" on the part of the AI in generating images. This participation allows them to interrogate the algorithm's latent space—the abstract model allowing for the mapping of text strings to images. Why might such an interrogation of the latent space of the algorithm be significant? Alternatively, why should we care about the way that the algorithm "sees"? Recall how these image generators are trained: they are the result of deep learning algorithms applied to a massive corpus of image-text pairs sourced from the internet. When we learn about the algorithm's ways of seeing and representing, we are in some sense learning about *our*

own ways of seeing and representing. We see ourselves, through a glass darkly—as expressed in the visual culture of the internet and encoded by the algorithm's deep learning.

Let me illustrate with two examples. Sofia Crespo is an artist using AI to interrogate our visual representation of the natural world. *Critically Extant* (2022) is a series of works in which she uses an AI image generator to create representations of endangered or extinct species. Even though these image generators are trained on millions of publicly available images on the internet, the generated representations look very little like the species themselves. Consider the example of the Taurus Gudgeon—a small freshwater fish native to Turkey which the International Union for the Conservation of Nature lists as "critically endangered." Crespo's image of the fish doesn't even look like a fish—let alone *this specific species* of fish. What accounts for this is the near-total absence of images and representations of these species in our ordinary online spaces. In mapping the blind spots of the algorithm, Crespo is able to illustrate how we collectively see—or more to the point, fail to see—these endangered species.

Another example is the work of Mario Klingemann, a well-known AI artist whose work has often involved programming and training his own AI algorithms. Consider his 2018 work *Memories of Passersby I*: a sculptural installation featuring portraits automatically generated by way of an algorithm that Klingemann himself coded. Klingemann trained the algorithm on thousands of painted portraits from the 17th to 19th centuries; although each is a novel creation, the generated portraits bear an uncanny resemblance to a great many classic works of portraiture. Klingemann's work doesn't make use of a text-to-image transformation algorithm, but nevertheless it functions as an exploration of the way the AI model has learned to "see" and represent from its training data. Through it, we may come to see and appreciate the conventions of painted portraiture, as distilled and recombined through the algorithm.

The analogy with performance art is also useful in thinking about how, under the exploration paradigm, we should understand the status of the work and the main focus of artistic appreciation. Consider again Abramović's *Rhythm 0*. Tate Britain includes in its collection the table with each of the 72 objects Abramović selected, as well as a series of slides featuring text and images of the 1974 performance. What's important for our purposes is simply to note that these serve as *documentation* of the

work itself; they help us to appreciate and understand what transpired. Returning to AI art: the exploration paradigm treats AI-generated images similarly to photographs of a piece of performance art. Rather than being the main focus of artistic appreciation, they serve instead as documentation of the artist's process of exploring, interrogating, and mapping the algorithm.

What exactly is the AI artist's role in such a process, and what do they contribute? Sherri Irvin has convincingly argued that, for much contemporary art, the artwork consists (at least in part) of the provision of specific rules for engagement with the work. These often include rules for *participation* in particular:

> Rules for participation can be an integral part of the artistic project pursued within an artwork, affecting the kinds of experience it affords and the kinds of meaning one can reasonably find in it. The prospect for participation, even if not taken up by a particular audience member, deeply shapes the expressive possibilities and themes of the work (Irvin 2022: 78).

Although Abramović herself couldn't fully predict her audience's participation in *Rhythm 0*, appreciating her work consists largely in appreciating how she established a space for participation by way of the instructions she provided to her audience member. According to the exploration paradigm, the AI artist contributes something similar by way of their prompts to the generative AI model: they create a space for the AI to participate in ways that may be challenging, illuminating, or engaging.

One question that may arise under the exploration model is whether or how the AI artist's activity differs from, for example, AI content moderators who are often in the business of exploring AI models as well. What distinguishes the AI artist's work as art?[2] I don't want to fully commit myself to any particular account of what makes an activity an artistic activity, but for the sake of argument I'll adopt one approach that I think is promising: Richard Wollheim's account of the importance of thematizing activities in his discussion of painting as an art. On Wollheim's account, an artist thematizes their activity when they come to regard it as contributing in some way to the overall meaning of the artwork. According to Wollheim, it is in this process of thematization—of "attempt[ing] to organize an

---

[2] I'm grateful to an anonymous referee for pressing me on this question, as well as the two further questions discussed in the following two paragraphs.

inherently inert material so that it will become serviceable for the carriage of meaning"—that a painter comes to practice painting as an art, rather than mere painting (1987: 20-25). Wollheim's account is hardly uncontroversial, but I think his notion thematization is useful here: artmaking requires a kind of critical engagement with one's agency, especially insofar as it contributes to the meaning of the resultant work as a whole. I think this might help to clarify the difference between the AI artist and the AI content moderator: while the former aims to explore AI models for the sake of e.g. ensuring that these models are safe to use, the AI artist instead treats this exploration as a vehicle for artistic meaning. As I've illustrated by way of the examples above, many AI artists are interested in using AI as a kind of mirror in which we might explore ourselves—at least insofar as we are reflected in the models' training data. This is one among many ways in which an AI artist might thematize their exploration of AI algorithms.

Two further questions for the exploration paradigm concern the limits of the analogy with participatory performance art: First, one might be concerned that AI cannot actually "participate" in an artwork, insofar as the AI is incapable of conscious choice or more robust agency. In response, I will concede that that the analogy with performance art isn't a perfect one. However, this seems to be no more of a problem for the exploration paradigm than the concerns I raised in the previous section concerning the analogies motivating the production paradigm: AI seems to be incapable of genuine "collaboration" just as much as it is incapable of genuine "participation." One option at this point would be to think of AI image generators not as a participant but rather as a very specific kind of tool—one that we use, perhaps, to explore our reflection in latent space. This approach might capture some of the motivation for the exploration paradigm without having to attribute participation to an AI model. However, I think that this approach would still face the concern that the opacity of these algorithms makes the analogy with tools imperfect as well. None of these analogies is perfect—a point I discuss in more detail in the conclusion below—but even so I believe that the exploration paradigm provides us with useful tools for understanding and appreciating AI art.

A second, related question concerns the audience for AI art according to the exploration paradigm: in traditional performance art, the art is performed for an audience that, in addition to participating in the work, is also capable of experiencing the work. Unlike such an audience, an AI model

seems incapable of experiencing a work. This question again points to the limits of the analogy with performance art, but in response I will simply observe that performance art is often meant to be appreciated by an audience that extends well beyond those present at the initial performance. To return once more to the example of Abramović, audiences today appreciate her work by way of encountering documentation of it—photographs, recordings, and the physical objects that played a role in the performance. Similarly, I'd suggested that audiences can appreciate AI art by way of accessing the outputs of the artist's exploration of the algorithm—that is, the images and other artifacts generated. These function as a kind of documentation enabling appreciation, even if they are not themselves the primary focus of artistic appreciation.

5. Implications

So far, I've introduced two different paradigms of agency in AI art: the production paradigm and the exploration paradigm. In the remainder of the paper, I will illustrate the significance of these two paradigms. Which paradigm we adopt will lead us to very different answers to a number of important questions concerning the understanding, appreciation, and evaluation of AI art. I'll argue further that it may be more productive for both artists and appreciators in many contexts to adopt the exploration paradigm rather than the production paradigm. To make my case, I'll focus on three such questions: the question of style in AI art; concerns about algorithmic bias in AI art; and questions about credit and copyright in AI art.

Before I begin, I want to put forward two important disclaimers: First, the production paradigm and the exploration paradigm are not entirely exclusive. In practice, I suspect that many actual instances of AI art may involve concern both for the product of the process of interaction with the AI *as well as* the process of exploring the way the algorithm sees. Return to the example of Sofia Crespo's work: it is no accident that she presents us with images that are visually appealing and which draw our attention to their formal qualities. While these images certainly serve as documentation of her interaction with generative AI, they can also be objects of appreciation in their own right—a point which the artist herself has emphasized in her own presentation of the work (Crespo 2022). In this paper I take some inspiration from Dominic Lopes's approach in *Four Arts of Photography*, in which he

characterizes four "arts" of photography—four ways that photographers might practice photography as an art. As Lopes puts it:

> The four arts of photography are very rough guides to how to appreciate different kinds of photography, or how to appreciate photography in different ways… By no means does the method imply that photographs should be made as exemplars of only one art at a time. Nothing much recommends making a game of arguing over which pigeonhole a particular photograph goes into. Philosophy is not the carpentry of pigeonholes. Idealization sharpens the powers of human understanding without dictating behavior (Lopes 2016: 34-35).

Like Lopes, my goal is not to enforce a rigid categorization or pigeonholing of AI art, but rather to facilitate appreciation by drawing our focus to additional dimensions of artistic concern and agency.

Even so, this is not mere categorization for its own sake. I suspect that the default approach to the appreciation and creation of AI art will be something akin to what I've called the production paradigm. This may result in a limited appreciation of the significance of AI art, insofar as we neglect important aspects of this art's artistic value. Similarly, I suspect that AI artists themselves might find that adopting the exploration paradigm offers richer prospects for artmaking. In offering the exploration paradigm, my aim is to open up a sense of the possibility and promise of AI artmaking that goes beyond an exclusive focus on the objects that are the output of the creative process.

The second disclaimer to offer is that, to this point, I've only been focusing on "pure" applications of AI image generation tools by artists. In these cases, the creative process almost exclusively consists in the artist providing prompts to a generative AI platform and using it to create images. It's important to note that many contemporary AI artists do much more than just this. Consider Sougwen Chung's work, which combines AI image generation with traditional painting, Chinese calligraphy, and robotics; or Anna Ridler's work, which involves AI processing in an interrogation of large data sets of various kinds, many of which have been assembled and collected by the artist herself. Neither of these examples fit easily into the discussion I've offered so far, simply because they incorporate additional modes of artistic practice that extend beyond providing prompts to a generative AI model. AI tools can of course be incorporated into artistic practice in a wide variety of ways, and I do not mean to rule these out as cases of genuine artmaking. Instead, I have limited my consideration only to

the case of using AI image generators to show that, even when all an artist does is make use of such tools, there is still nevertheless a possibility of genuine artistic agency and meaningful artistic production.

5.1 Style

Is there such a thing as an AI art style? And if so, what might it consist in? My suggestion is that adopting the exploration paradigm points us in the direction of richer prospects for the development and appreciation of artistic style in AI art.

In discussing artistic style, I first want to distinguish between two senses of the term: we can distinguish between *general* style and *individual* style (Wollheim 1987: 26). In speaking of a general style, we might look to distinctive characteristics or notable features common to a body of work—think here of the characteristic features of Art Nouveau, Brutalism, Impressionism and the like. On the other hand, individual style is usually held to be a characteristic mode of agency present in the work of an individual artist—a particular choice of technique, a specific set of artistic resources, and so on, which perhaps express the artist's ideals for their work (Riggle 2015). Developing a style depends on the ability to reflect on one's agency over time, to conceptualize it, and to consistently deploy it in artmaking with predictable results.

Adopting the production paradigm might push us to look for artistic style—both general and individual—expressed in the *outputs* of AI art. But are there any distinctive characteristics or features of the images and other artifacts generated by way of generative AI platforms? Alice Helliwell has argued that a distinctive feature of such artifacts is their *weirdness*—the fact that they often involve significant norm-violations for depiction. For those familiar with AI-generated images this should be no surprise: it's easy to find examples of AI-generated hands with too many fingers or garbled strings of text. However, as Helliwell herself notes, such weirdness may only be transitory:

> As AI systems increasingly succeed in producing works that resemble the human works that they are trained to replicate, they will cease to produce weird outputs. As aestheticians, this might be of concern to us; soon, the distinctive aesthetic of weird AI art may be lost (Helliwell 2022: 243).

Such improvement is likely not too far in the future, given the rapid pace at which the output quality of generative AI is improving. I suspect that, over time, the category of "AI art" will come to be irrelevant from the perspective of artistic style, at least insofar as such style is understood to consist in commonalities to be found among the images produced using generative AI.

The exploration paradigm points us in a different direction in looking for style: the suggestion introduced above is that such style would lie in "how one orchestrates the design of [the latent space of encoded image representations] and what one finds in it" (Cetinic and She 2022). Such a style would exist at the level of the prompts one chooses to explore, the choice of image generation tool, and the like. It would not necessarily be discernible at the level of the visual outputs of one's work—that is, the generated images. There are two important implications here: first, individual style for AI artists, should it exist, has less in common with the artistic style of traditional visual artists than we might have thought. Individual style for AI artists might instead be closer to the individual style of conceptual artists whose work includes a visual component, like Sol Lewitt or John Baldessari. Second, in looking for a general style of AI art we might do better to focus on non-visual commonalities among work—we should focus instead on artists' choices of AI models, prompting strategies, training data, and so on. Looking at AI art from the perspective of the exploration paradigm helps to make these points clearer, and I believe that doing so will lead to richer ways of conceptualizing and appreciating AI art.

5.2 Algorithmic Bias

It is by now widely known that AI tools trained on large datasets take on the biases that are inherent in those datasets. Without correction, many AI tools reliably reproduce biases similar to well-known implicit bias patterns in humans, including biases with respect to race, sex, and age (Johnson 2021). It should be no surprise that similar biases have emerged in AI image generation tools: generated images tend to feature persons with lighter skin tones who conform to traditional gender norms (Srinivasan and Uchino 2021). Algorithmic bias has attracted a great deal of attention insofar as such bias reflected in the outputs of an AI image generator seems morally objectionable (Fazelpour and Danks 2021).

What is less appreciated is that, in the case of AI image generation, such bias may also be *aesthetically* objectionable. What I mean by this is that such bias might lead to significant concerns about the aesthetic value of the outputs of AI image generation tools. Why might this be the case? Given that the current crop of tools has been trained on image-text mappings scraped from the internet, they are *very good* at producing conventional modes of representing particular objects. This is true quite generally, but also at the level of general visual styles of the sort discussed above: you can get an ordinary picture of a hot-dog, and you can also generate a hot dog picture in the style of Van Gogh, film noir, cyberpunk, or clip-art. But you cannot use these tools to generate a new mode of representation or a new visual style entirely. The concern is that these algorithms are therefore "locked in" to established modes of representation. The worry, then, is that AI generated art cannot be original.

How we respond to this concern will depend on which paradigm of AI art we adopt. If we default to the production paradigm, then we would need to respond to this concern by demonstrating that, in fact, generative AI can avoid such concerns about being locked into established representational styles. Such a strategy might be promising; it is after all possible to train an AI image generator to *maximize* visual novelty while maintaining some recognizability as art (Elgammal et al. 2017). Such a process would no doubt provide fodder for an artist looking to develop a new mode of visual representation. Even so, it's not my purpose here to pursue this line of argument; I simply want to note that the need to make it in the first place depends on adopting something like what I've called the production paradigm of AI art.

On the other hand, suppose that we adopt the exploration paradigm. If we do so, then it looks like this objection hardly applies. As I've argued, the artist's agency in AI art is largely a function of interacting with, exploring, and interrogating the algorithm. This means that there is a great deal of potential for creativity and novelty *even if* the visual outputs of this process largely cohere with existing modes of visual representation. In fact, an artist might make use of such algorithmic biases as an object of her exploration and investigation; from the perspective of the exploration paradigm, these look more like artistic *resources* than they do artistic limitations.

5.3 Copyright and Credit

Lastly, I will focus on the implications for artistic credit and copyright that follow from adopting these paradigms of AI art. Let me begin with a follow-up to the original story about Jason Allen: in September, 2023, after a year of requests and appeals, the US Copyright Office's Copyright Review Board denied Allen's request for copyright of *Théâtre D'Opéra Spatial*. In the statement that the review board issued, they cited newly-issued guidance for AI-generated artworks: roughly, if the "traditional elements of authorship" are for the most part generated by a human, the work may be copyrighted by them. On the other hand, if these elements are largely AI-generated, then the work may not be copyrighted by a human. In the case of Allen's work, the board found that "the Midjourney Image, which remains in substantial form in the final Work, is not the product of human authorship." Although the board recognized that Allen inputted a prompt to generate the image, they found that "Mr. Allen's actions as described do not make him the author of the Midjourney Image because his sole contribution to the Midjourney Image was inputting the text prompt that produced it" (United States Copyright Office 2023b).

I won't try to resolve whether the US Copyright Office's policy is the correct one; instead, I simply want to point out that its focus on the contribution of generative AI to the final product is squarely in line with the production paradigm. On this view, the work just *is* the output—and therefore it makes sense to think about issues of copyright in terms of the degree to which generative AI is responsible for the creation of this output.[3]

Adopting the exploration paradigm leads us into slightly more challenging—but also more interesting—territory: what the work *is*, according to this paradigm, isn't so much the output as it is the interaction between the artist and the generative AI platform, as structured by the artist's prompts. What would need copyright protection in this instance would be the prompt itself—intriguingly,

---

[3] The production paradigm also lends weight to the oft-aired complaint that AI models are trained on the work of human artists. To the extent that the output of these models often copies or duplicates distinctive features of the work of these artists, there may be cause for complaint on the part of the human artists that their work has been unfairly copied.

something that the US copyright office has already considered, allowing in its guidance on AI authorship that some prompts may be "sufficiently creative to be protected by copyright" (United States Copyright Office 2023a). My suggestion is that adopting the exploration paradigm orients us towards these concerns, allowing for an expanded conversation about what sorts of copyright protections AI artists deserve.

6. Conclusion

We often seek to understand new and emerging artistic practices by comparing them with old ones. Photography was compared to painting and film was contrasted with theater, just as the early novel was measured by the yardstick of verse fiction. One concern about such comparisons is that we might miss what's distinctive about the new practices, simply because our focus on comparison alone might make us insensitive to new varieties of artistic agency and the prospects they afford.[4] Although this concern is a legitimate one, I think that these comparisons are unavoidable: they are often all that we have at the outset to make sense of emerging forms of artistic agency, and they can also be helpful starting points for consideration of their prospects.

This paper has focused on a slightly different concern, namely: the choice of comparison matters greatly, and in picking the wrong set of comparison cases we may miss out on a great deal. What I've called the production paradigm of AI art is a paradigm motivated by comparing AI art to existing artistic practices which, like photography, depend on the usage of a mechanical tool, or which, like film production, involve reliance on skilled collaborators. Comparison with these existing practices provides us with a particular orientation towards AI art—one focused on the artistic significance of the output which is the product of the creative process. Given the prevalence of these artistic practices in our culture, it's quite natural that they are the default basis for comparison when it comes to AI art.

In sketching out the exploration paradigm, I hope to have offered a different set of artistic practices for our comparison: these practices are drawn from the rather more esoteric world of contemporary art. In such practices, artists explicitly create opportunities for *participation*—and it is the nature

---

[4] For such an argument applied to the consideration of television as an artistic medium, see (Nehamas 1999).

of this participation that is illuminating. In comparing AI artmaking to such practices—and inviting us to think of AI as a kind of participant in dialogue with human AI artists—I hope to have broadened our sense for the possibilities of artistic agency in AI art.[5]

Reference List

Anscomb, C. (2022) 'Creating Art with AI', *Odradek*, 8/1: 13-51.

Brainard, L. (2023) 'The Curious Case of Uncurious Creation', *Inquiry: An Interdisciplinary Journal of Philosophy*: 1-31. https://doi.org/10.1080/0020174X.2023.2261503

Cetinic, E., and She, J. (2022) 'Understanding and Creating Art with AI: Review and Outlook', *ACM Transactions on Multimedia Computing, Communications, and Applications*, 18/2: 1-22. https://doi.org/10.1145/3475799

Crespo, S. (2022) 'AI-generated creatures that stretch the boundaries of imagination', *TED: Ideas Worth Spreading*. Available at:

https://www.ted.com/talks/sofia_crespo_ai_generated_creatures_that_stretch_the_boundaries_of_imagination (Accessed 16 October 2024).

Dutton, D. (1979) 'Artistic Crimes: The Problem of Forgery in the Arts', *The British Journal of Aesthetics*, 19/4: 302-14. https://doi.org/10.1093/bjaesthetics/19.4.302

Elgammal, A. et al. (2017) 'CAN: Creative Adversarial Networks, Generating "Art" By Learning About Styles and Deviating from Style Norms', in: Goal, A., Jordanous, A., Pease, A. (eds.) *Proceedings of the 8th International Conference on Computational Creativity, ICCC 2017*, pp. 96-103. Atlanta, GA: Georgia Institute of Technology.

Fazelpour, S., and Danks, D. (2021) 'Algorithmic Bias: Senses, Sources, Solutions', *Philosophy Compass*, 16/8: e12760. https://doi.org/10.1111/phc3.12760

Fei-Fei, L., and Krishna, R. (2022) 'Searching for Computer Vision North Stars', *Daedalus*, 151/2: 85-99. https://doi.org/10.1162/daed_a_01902

Helliwell, A. C. (2022) 'Art-ificial: The Philosophy of AI Art'. PhD thesis, University of Kent.

Hertzmann, A. (2018) 'Can Computers Create Art?', *Arts*, 7/2: 18. https://doi.org/10.3390/arts7020018

Irvin, S. (2022) *Immaterial: Rules in Contemporary Art.* Oxford: Oxford University Press.

Johnson, G. M. (2021) 'Algorithmic bias: on the implicit biases of social technology', *Synthese*, 198/10: 9941-61. https://doi.org/10.1007/s11229-020-02696-y

Kelly, S. D. (2019) 'What Computers Can't Create', *MIT Technology Review*, 122/2: 68-75.

Livingston, P., and Archer, C. (2010) 'Artistic Collaboration and the Completion of Works of Art', *The British Journal of Aesthetics*, 50/4: 439-55. https://doi.org/10.1093/aesthj/ayq029

Lopes, D. M. (2016) *Four Arts of Photography: An Essay in Philosophy.* Malden, MA: Wiley.

Nehamas, A. (1999) 'Plato and the mass media', in: *Virtues of Authenticity: Essays on Plato and Socrates*, pp. 279-99. Princeton, N.J: Princeton University Press.

O'Hagan, S. (2010) 'Interview: Marina Abramović', *The Observer*: 32.

Radford, A. et al. (2021) 'Learning transferable visual models from natural language supervision', in: *Proceedings of the 38th International Conference on Machine Learning*, pp. 8748-63. PMLR.

Ramesh, A. et al. (2021) 'Zero-shot text-to-image generation', in: *Proceedings of the 38th International Conference on Machine Learning*, pp. 8821-31. PMLR.

Riggle, N. (2015) 'Personal Style and Artistic Style', *The Philosophical Quarterly*, 65/261: 711-31. https://doi.org/10.1093/pq/pqv026

Roose, K. 2022) 'An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy.', *New York*. Available at: https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html (Accessed 16 October 2024).

Srinivasan, R., and Uchino, K. (2021) 'Biases in Generative Art: A Causal Look from the Lens of Art History', in: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and*

Transparency, pp. 41–51. Association for Computing Machinery.

https://doi.org/10.1145/3442188.3445869

Trujillo, J., G.M. (2022) 'AI Art Is Art', *Aesthetics for Birds*. Available at:

https://aestheticsforbirds.com/2022/11/02/ai-art-is-art/ (Accessed 16 October 2024).

United States Copyright Office (2023a) 'Copyright Registration Guidance: Works Containing

Material Generated by Artificial Intelligence', in *Federal Register*, pp. 16190-94. Available at:

https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-
works-containing-material-generated-by-artificial-intelligence (Accessed: 16 October 2024).

United States Copyright Office (2023b) *Re: Second Request for Reconsideration for Refusal to
Register Théâtre D'opéra Spatial (SR # 1-11743923581; Correspondence ID: 1-5T5320R)* Available
at: https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf
(Accessed: 16 October 2024).

Wojtkiewicz, K. (2024) 'How Do You Solve a Problem like DALL-E 2?', *The Journal of Aesthetics
and Art Criticism*, 81/4: 454–67. https://doi.org/10.1093/jaac/kpad046

Wollheim, R. (1987) *Painting as an Art.* Princeton, NJ: Princeton University Press.