A Defense of Artificial Intelligence

IN THE ARTISTIC FIELDS

/imagine an essay providing a defense of the AI neural networks used in the arts, as well as a warning to the general public about the danger of lobbyists hijacking conversations surrounding Artificial Intelligence regulation.

By Thomas K. Yonge

Safe-Essay Feature:

This essay makes use of secure bit.ly links (+); clicking links will provide readers with a preview of the link's contents before visiting the webpage.

Author Contact:

The author of this essay may be contacted for inquiries or concerns at:

thomas.k.yonge@gmail.com

Preface to the Article

Artists, engineers, and regulators alike stand at the brink of a digital Renaissance, the nature of which can and must be determined through a conversation between all stakeholders. My name is Thomas K. Yonge; as a writer, educator, content creator, and equities investor, I consider myself one of these stakeholders.

Growing up through the turn of the century, I watched as the advent of large-scale technology adoption brought vast changes to how societies around the world functioned:

- Photoshop and Flash let at-home artists bring ideas to life in a way that once required a studio.
- Social media like YouTube and Twitter made sharing such creations with a global audience possible.
- AWS, Cloudflare, Google Cloud, Microsoft Azure, and other architectures provided the backend structure needed to host this and a range of other digital projects, collectively Web2 and IoT.[1]

Each technology brought with it challenges and a threat to old industries, but each also served to democratize the world and create new forms of human expression. Nowhere has this been truer than in the field of art, where images, videos, and ideas of incredible complexity now circulate the globe at unprecedented rates.

The Dangers of Regulation

If the foreground advance in technology has caught the eye of impressed onlookers, then regulation has by contrast lived in its shadow as a secondary but essential framework. In one respect, it has reined in technology's 'Wild West', bolstering consumer protection, privacy options, and cybersecurity. Yet this political conduit has also become an Achilles heel by which powerful lobbyist groups have attempted to shape policy decisions—often and especially regarding the topics of technology and copyright.[2] [3]

The Layout of This Essay

In this paper, we will engage in a conversation about AI art that addresses both the concerns and the regulatory challenges associated with its development.[4] The essay will contain the following sections:

a) A brief description of how neural networks such as CNNs, GANs, CLIP, and diffusion work.
b) A point-by-point discussion about the concerns and challenges surrounding AI regulation.
c) Voices and statements from users in the AI art community, Midjourney.

---

[1] Internet of Things: a collection of physical devices and their sensors that collect and transfer data, i.e. 'smart devices'.
[2] For instance, this year, Alphabet Inc., Apple Inc., Amazon.com Inc., and Meta Platforms Inc., bled a combined $95 million in lobbying over antitrust regulations that would have directly affected them. See: bit.ly/tech-lobbying+
[3] Likewise, Disney, IBM, and other legacy companies have attempted to increase regulation of competitors in order to open up routes to patent lawsuits and, as a result, new streams of predatory income. See: bit.ly/disney-tech-copyright-230+ and bit.ly/ibm-filters+
[4] As a nascent movement, there is still debate about terminology regarding art made via AI. One specific candidate is synthography; another is generative art. However, for ease-of-reading, this article will use the general phrase 'AI art'.

HOW THEY WORK: CNNs, GANs, DIFFUSION, AND CLIP

Artificial intelligence is a field of programming inspired by the human brain. It uses datasets to train what programmers refer to as a **neural network**[5]: a system of layered code that lets a computer make decisions by categorizing information from a collection of samples—usually pictures or phrases.

Multiple neural networks exist today, spearheaded by different teams of researchers and linked to programs that execute either specific or general tasks: *Midjourney*, for instance, intakes data fed to it by a user to specifically produce computer-generated images, while *ChatGPT* uses user data to act more generally as a virtual assistant. One of the newer models of neural network, which will be used as an example in this essay, is called CLIP, and this itself snaps onto other types of AI such as diffusion models. In this sense, neural networks can be thought of like LEGO bricks.

CLIP, or **Contrastive Language-Image Pre-training** is a collection of code layers and rules (called a **model**) that combines natural language processing (NLP) and zero-shot-styled learning to think in an 'on-the-go' way—this is as far into CLIP's definition as we can get without first understanding two topics:

a) the older models that CLIP builds on, called convolutional neural networks (CNNs)
b) how CLIP uses two new methods—zero-shot and NLP—to improve the performance of the models that it is snapped onto

Convolutional Neural Networks

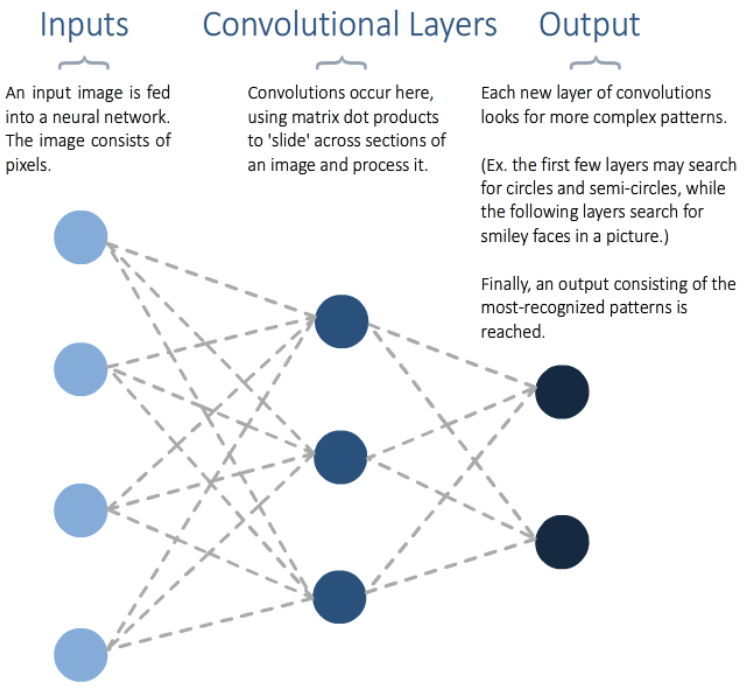
A **Convolutional Neural Network** (called **CNN or ConvNet** for short), is an older type of learning model that is able to identify the characteristics of an image and match it to a set of information that it already knows. CNNs are used in CAPTCHA, facial recognition software, and OCR.[6]

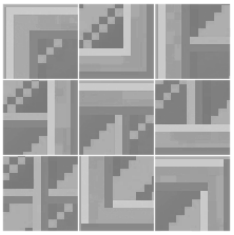
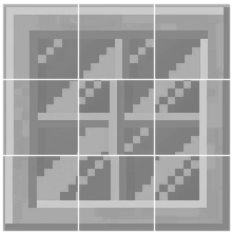
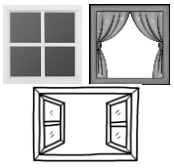
In a CNN, data engineers start by giving their neural network a set of patterns to recognize. These can be things like lines, edges, circles, squares, etc… They then feed their network an image made up of pixels.[7] The network searches for patterns that it recognizes in the image and begins to collect, match, and organize them until it arrives at a sufficient guess of what it is looking at. Each instance where it matches a certain type of pattern in the image is referred to as a **filter**, starting with simple filters like pixels that make up straight lines or pixels that make up squares, and moving on to more complex filters (such as ones that might combine straight lines and squares to make something that looks like a window, or filters that examine the colours of those lines and square to see if they show a brown window or a grey window).

The process whereby the CNN goes about its job assorting and organizing these filters is called **pooling**, and an example of this process from start to finish is given on the following page.

[5] Or more specifically, an 'artificial neural network'—as opposed to a genuine neural network such as a brain.
[6] CAPTCHA: bit.ly/captcha-explainer+ Facial Recognition: bit.ly/facial-rec-explainer+ OCR: bit.ly/ocr-explainer+
[7] Small points in a picture made up of red-green-blue (RGB) values. Ex. The colour indigo, ■, has RGB: 75, 0, 130

Ex. Data engineers have a collection of window images and want to know if a new image is also of a window. They feed their neural network the new image: it looks for certain patterns in the pixels, and as it finds these patterns, it begins to combine them into a rough idea of what it is looking at.



Input	Convolutional Layers *(Thinking in progress!)*	Output
Input Layer: I have received this picture. I do not know what it is yet. 	Layer 1: Are there simple patterns, like small groups of pixels that appear together?  ... Layer 2: Are there some more abstract patterns, like sharp lines, where pixels suddenly change colour? Layer 3: Are there even more abstract patterns, like colours associated with those lines? Layer 4: Are there overall 'window shapes' associated with these lines and colours?	Output Layer: Hmm… Let's put it all together now. This is most likely a **window**.  Why? Because it looks fairly close to the collection of windows that my data engineers gave me. 

Therefore, to summarize, CNNs, as well as all neural networks, are generally:

Software that is given or 'fed' a set of patterns to look for, then fed data in the form of images or words, and then is expected to recognize, arrange, and combine the patterns it discovers in order to produce a best guess of what it was looking at.

> For a more detailed explanation of how CNNs function, see: bit.ly/cnns-explainer+
>
> For a more detailed explanation of how neural networks in general work, see: bit.ly/nns-explainer+

Improving on CNNs with GANs, on GANs With Diffusion, and on Diffusion With CLIP

The powerhouse of CNN computation occurs when data scientists take two of these neural networks and put them together, in essence, allowing them to talk to each other—or more accurately, to compete with each other in a sort of game. This game is called a **generative adversarial network**, or **GAN**. In a GAN, two neural networks, a **discriminator** and a **generator** are set up. The discriminator is fed data from a training set (Ex. 10 images of chairs: rocking chairs, office chairs, recliners, sofas, etc…) and learns to identify these types of objects by their shared characteristics. The generator, on the other hand, is fed noise—a random assortment of useless RGB pixels, like what a person would see on a TV screen full of static. From this, the generator creates a new pattern of noise.

Once the discriminator is sufficiently trained, data scientists turn on the generator and send to the discriminator either:

a) a real image of a chair from their dataset
b) the useless, mushy noise made by the generator

The discriminator then attempts to guess if what it has been sent is real or fake. If it guesses correctly, its code remains unchanged; if it fails, then its code will be given 'additional information' to help it course-correct and improve its guessing; the same is true for the generator—while at first it is likely to fail at convincing the discriminator almost every time, as it course-corrects, its fakes start to become better and better until finally, it is able to defeat the discriminator by 'imagining' a chair that did not exist in the training set, but that has all the qualities of a chair.

In essence, this antagonistic relationship between the discriminator and generator help to train a new, more powerful model for each—with the newly robust generator being the more useful of the two. In fact, generators can be applied to image synthesis, video synthesis, data augmentation, style transfer, and music generation.[8] That said, GANs have major drawbacks: they require large amounts of computational power, they may produce low-quality images despite their training, they are difficult to finetune, their generators can suffer mode collapse (whereby they only end up producing a limited number of samples), or the generators can suffer vanishing gradients.[9]

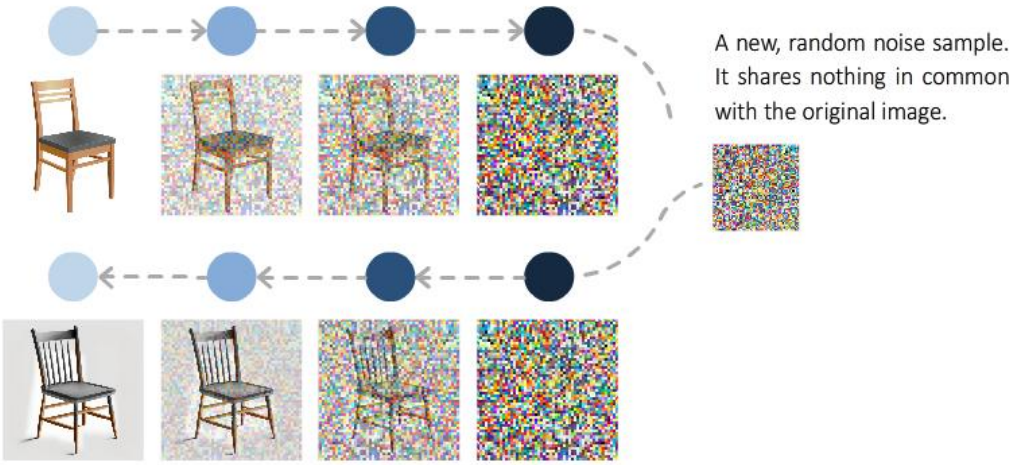
[8] Generators are particularly good at making photorealistic images, with a main example being NVIDIA's StyleGAN.
[9] The last of these, **gradients**, are essentially the strategies that the generator and discriminator use to course-correct after each round of their game. As both models get better at playing, the course-corrections become smaller and smaller—yet smaller changes mean that the models can no longer effectively update their strategies in any meaningful

As a result, a different type of model is required to solve these issues—this is where **diffusion models** come into play. If GANs can be considered a multiplayer game, then diffusion is the solo campaign equivalent. In it, there is no competition, and furthermore, there are no discrete rounds. Instead, a neural network plays a game in an environment modeled by continuous time dynamics. Consider it like a player being put in a room with a clock and asked to solve a puzzle. In this context, the player, or neural network, takes several 'steps' to solve the puzzle, which for the sake of this example might be that it is given a fuzzy image and asked to make out what that image is supposed to be. Maybe it is an elephant; maybe it is a chair.

Notice that in this new type of game, noise is actually added to training set images before they are given to the neural network—this is done in steps, and it is done to such an extend that the images are rendered completely illegible. The 'game', then, is for the neural network to learn how to guess what the original image looked like by applying each noising step in reverse (referred to as **de-noising**).[10]

Doing this, a model can then take any random assortment of noise—not just those belonging to its original training set—and apply the same reversed steps, creating completely new images that share characteristics matching those of the training data. This is how a diffusion model works. See the following example:

Noise is applied to an image of a chair through a series of steps, destroying the image in the process. By applying these steps backwards, a neural network can 'de-noise' or recover the image. By denoising a random sample of noise, however, it can also create an entirely new image with chair-like qualities.



Now for the cherry on top: recall that whether we are looking at a GAN or a diffusion model, each round or step requires a course-correction to occur so that the neural network can improve itself and eventually 'win'. In the context of a neural network whose goal is to draw images, to win then means to be able to take in data and produce some reasonable image that matches that data.

---

way. Even worse, these gradients—so small that, in effect, they 'vanish'–wind up getting lost within the code, and they become unable to move effectively through the layers of the neural networks.

[10] It is worth noting that for simplicity, **diffusion**—the mathematics behind and namesake of diffusion models—has been omitted since it is not the main strategy for image-generating. Nonetheless, it is fully described in the glossary.

On their own, GANs and diffusion models can produce reasonable results, but what if another neural network could be layered over them to improve their abilities? The **Contrastive Language-Image Pre-Training (CLIP) model** does just that, acting like a strategy-guide for our AI player that is built out of other models called **transformers**.[11] This guide shows the player model all the tips and tricks that can get them to a more accurate solution in their game. To achieve this, CLIP uses CNNs in combination with two novel ideas: natural language processing and zero-shot learning.

Natural language processing (NLP) can be thought of as a rulebook of semantics and grammar. It allows CLIP to 'supervise' other neural networks by allowing those networks to understand language prompts.

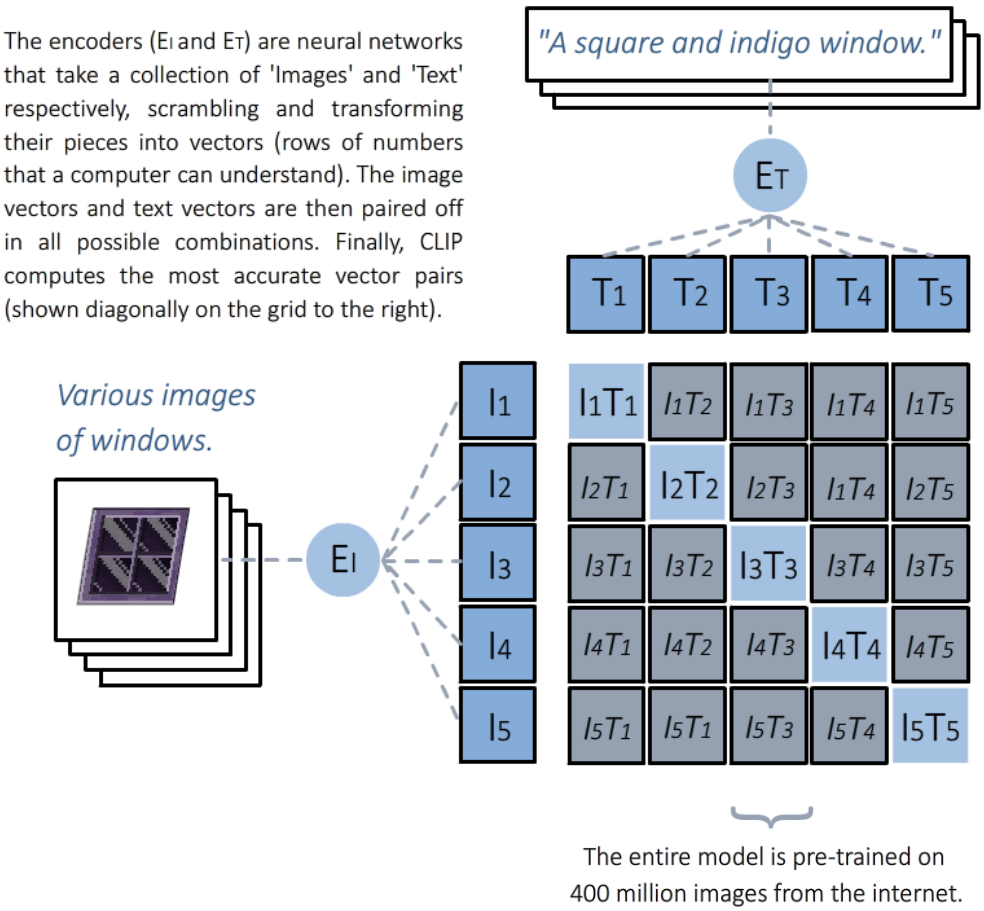
Zero-shot learning is a new, on-the-spot style of computing in which a neural network encounters new objects that are not from its original training data. Ex. a neural network that knows horses and encounters a zebra, but is able to classify the zebra as a horse with some additional characteristic—mainly, stripes. Under normal circumstances, zero-shot learning is difficult to achieve.

CLIP's model also uses a wide array of web-scraped data—specifically pairings of images and text descriptions. It does this to achieve its own style of zero-shot. [12] This means that CLIP is able to supervise a GAN or diffusion model and guide it towards more efficient choices by a) applying NLP rules to b) help it learn on-the-spot through zero-shot learning.[13] This process is described visually on the following page:

[11] Transformers were initially discussed in Google's landmark paper *Attention is All You Need*: bit.ly/attention-doc+

[12] Roughly 400 million image-text pairings taken from websites through the use of software and/or data-sharing policies. Note that this is the same process that companies like Google use to organize their own architectures such as search engines. From Google: "Most of our Search index is built through the work of software known as crawlers. These automatically visit publicly accessible webpages and follow links on those pages, much like you would if you were browsing content on the web. They go from page to page and store information about what they find on these pages, along with other publicly-accessible content in Google's Search index." *Source:* bit.ly/google-search-explained+

[13] As a tangential point, many other types of neural networks and training methods exist outside of CNNs, GANs, and CLIP, and it is important to be aware of them. For instance, two examples of other neural networks are Recurrent Neural Networks (RNNs) and Multilayer Perceptrons (MLPs)—one is used in AI decision-making and the other is a somewhat more flexible form of CNN, which allows for better performance in tasks aside from image regression (such as recognizing patterns in time-series data). Each has its own strengths and weaknesses, and moreover, each may be combined with other types of networks to produce a more adaptable and robust system.



As a consequence of its design, CLIP also solves a number of overall problems with diffusion, GANs, and CNNs in general:

a) CNNs were only as powerful as their datasets and guiding architectures, often meaning that a specific neural network was only good at a specific kind of task (Ex. identifying chairs, but not drawing chairs, or understanding when a person was sitting in a chair). NLP and image-text pairings help CLIP overcome this issue because they provide surrounding context to the chair.

b) Neural networks that perform well on training sets used for benchmarks often fail to perform equally as well out in the wild. By using a kind of web-based zero-shot, CLIP avoids this issue; future models may even attempt to connect to the internet in real-time.

c) Datasets are labor intensive and only provide a narrow range of information. ImageNet for example contains only 14 million images. While this may seem like a high number, it is a far cry from the 400 million image-text pairings used by CLIP's web-scraper, and although CLIP was initially labour-intensive to put together, the fact that it can be applied to a wider range of tasks through a) and b) makes it a cheaper and more flexible model than other more task-specific neural networks.

Overall, natural language processing and zero-shot learning can enhance a model's guesses so much that the result is a neural network able to imagine entirely new images. *Midjourney*, *Stable Diffusion, DALLE* and other AI art-generation software all make frequent use of this enhancement.

Congratulations, you are now an A.I. expert!

Further Resources

> Radford et al. *Learning Transferrable Visual Models…* (CLIP): bit.ly/clip-article+
>
> A surface-level discussion of the CLIP model: bit.ly/clip-summary+
>
> How diffusion models function: bit.ly/diffusion-explainer+
>
> Natural language processing: bit.ly/nlp-explainer+
>
> Zero-shot learning: bit.ly/zsl-explainer+

POINT-BY-POINT: A DISCUSSION OF REGULATORY TOPICS

Overview

For artists, engineers, and regulators, the three most common questions about the effect of neural networks on fields such as literature and the arts are:

a) Will they will replace jobs?
b) Do they infringe on copyrights?
c) Are they dehumanizing artists and devaluing art?

In addition to answering these core questions, there are other regulatory caveats that must be addressed during the process of legislation in order to protect individual creators and small research labs from larger—and often predatory—corporations.

1. The effects of bills passed by a government are often difficult to withdraw or modify. Technology meanwhile is advancing at such a pace that unintended consequences from inert legislation may restrict future research and development of technologies in connected fields, having devastating effects on safety standards, the economy, and marketplace competitiveness.
2. AI regulation, if conducted in a partisan manner, will allow corporate entities and copyright trolls to harass content creators, AI and non-AI artists, hobbyists, and research labs with litigation.
3. Following from this, the largest real-time threat to AI legislation is that corporate entities may astroturf anti-AI movements to push for specific forms of regulation that favour their own neural networks while harming smaller companies and individuals who may compete with them.
4. Legislation cannot adopt a 'Good v. Evil' mindset. Harassment of AI art prompt engineers occurs in the same way that harassment of artists occurs, and both are valid experiences worth considering, though more often, Anti-AI art sentiment is signal boosted in the media.
5. Using a neural network is not necessarily a one-way interaction. Legislation and restrictions on how AI gathers its datasets have the potential to lead to faulty networks that may not only produce less efficient programs but programs that unintentionally direct or homogenize human behaviour.
6. Limiting training sets for AI and enforcing strict copyright policing will lead to a number of equity concerns, building prejudice into neural networks through law.
7. For many, AI bolsters equity, acting as a tool to help artists with disabilities, reintroduce marginalized voices into art spaces, and remove barriers to access for low-income creators.
8. Is there a best model for regulation moving forward?

Discussion: Common Questions

a) Will neural networks replace jobs?

Throughout history, human development has been influenced by an assortment of revolutions.[14] This is not an idle platitude; in fact, each revolution had a concrete and measurable process: it developed new *methods* to use on something previously *abundant* in order to create new *products*. In doing so, every revolution has changed human society by first improving the quality of life and second displacing old production methods; importantly, old production methods continue to exist in the new society, though in a diminished capacity, usually as a hobby or as part of a looser artisanal economy.

- The Agricultural Revolution developed the method of farming to use on wild fruit in order to create harvests, improving nutrition and displacing foragers—yet foragers still exist today.
- The Industrial Revolution developed the method of mechanical engineering to use on raw materials in order to create cheaper products, allowing individuals to own more assets and displacing factory workers. Yet factory workers still exist in supporting and technology-adjacent roles.

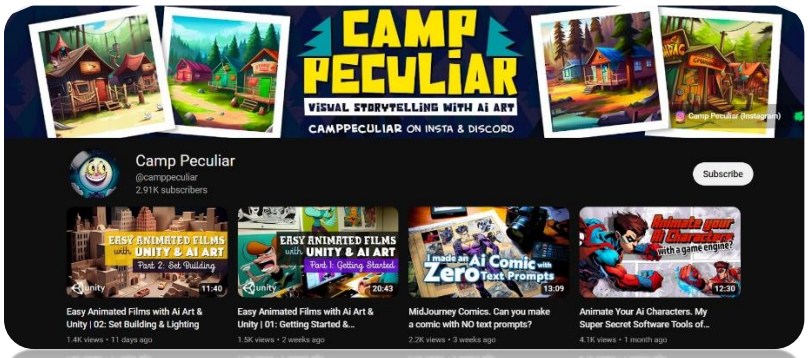
Likewise, the Information Revolution has now developed the method of neural networks to use on widely-available data (images, words, videos, sound bytes, etc…) in order to create an array of digital products, suites, and services, all contributing towards a $821 billion global IT market and allowing for accurate search queries and business-consumer matching, robust problem-solving models in the sciences and healthcare, and adaptive cybersecurity infrastructure critical to national defense. In exchange, this revolution displaces creatives—artists, writers, musicians.

Nonetheless, this is not to say that artists, writers, and musicians will cease to exist, as is the fear by many. Instead, new kinds of creators will emerge, using AI to help further their projects. I provide three examples below:

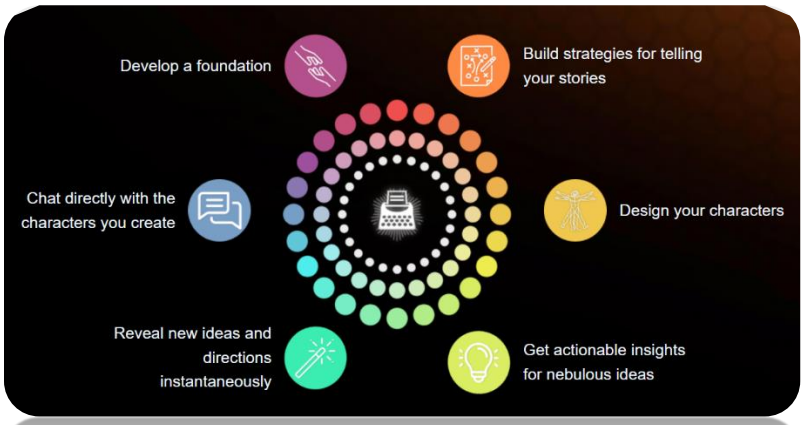
Ex. Camp Peculiar hosts a visual storytelling channel on YouTube where he makes use of *Midjourney* and other AI software to produce animated comics, unique character sheets, and storyboarding.[15] The channel provides instructionals on AI-use and how to design the best outcomes from AI-generated artwork. Camp Peculiar's role can be described as a **prompt engineer**, an individual who through the use of word prompts and image revisions creates unique characters and artwork.

[14] From the first developments of high culture during the Paleolithic Revolution, to the environmental Green Revolution beginning after WWII, the successful monetary policies of the Great Moderation, the Sexual Revolution of the 60s and 70s, and more recently, the Information Revolution that began with the construction of ARPANET in 1968—change permeates human society at nearly every level.

[15] Camp Peculiar's website: bit.ly/camp-peculiar-yt+



Ex. *Story Prism* is an AI co-pilot designed by Jon and Matt Firman to help writers develop a greater understanding of their characters, ideas, and plot directions.[16] For instance, a writer struggling with the motivations of one of their characters can feed their story's logline and central message into *Story Prism*.[17] The program then generates answers to a series of questions: 'What are your character's desires?', 'What are their physical/moral weaknesses?', 'What is something they would say?'—with each new answer depending on the responses generated for the previous questions; a writer can also modify answers at any step to produce different character outcomes, allowing them to both look into and craft the minds of their characters in real-time.



Ex. *SOUNDRAW* is a music generator that uses artificial intelligence to generate sound clip combinations based on music characteristics like length, tempo, and mood. The clips can be combined or recombined to produce unique and royalty-free sound tracks that users can add to original projects like videogames, YouTube videos, news clips, news summaries, and music with human vocals added. This can be extremely useful for creators who require specific styles of content-dependent soundtracks.

[16] *Story Prism's* main website: bit.ly/story-prism-website+
[17] An example of *Story Prism*'s use: bit.ly/prism-demonstration+

Each of the three examples above demonstrates the opportunity for a unique use of neural networks within art fields. While the 'uniqueness' of AI art in particular will be further explored in section b), in this section at least, I aim only to show here that despite being a nascent industry, AI prompt engineering is already leading to the creation of artist-adjacent jobs, and so both here and in this essay's final point, I will suggest that upskilling programs[18] conducted in partnership with companies like Midjourney and OpenAI can serve to lessen the need for heavy-handed regulation.

b) Do neural networks infringe on copyright?

To understand the issue of copyright, one must first acknowledge that every country has its own domestic copyright laws, and as the topic of AI art is an international one, the specifics of legislation will vary from country-to-country. However, nearly every country in the world has signed the 1886 Berne Convention[19] and so it is worth examining the spirit of this convention, since it both lays out a thought framework for the topic and produces some surprising insights into the place that AI art has within it.

- Article 2-(2) and 2-(3) of the Berne Convention state the general protection of literary and artistic works, especially of those that appear in a fixed[20] form, as well as the protection of "Translations, adaptations, arrangements of music and other alterations of a literary or artistic work… without prejudice to the copyright in the original work."
- Article 2-(5) states that "Collections of literary or artistic works such as encyclopaedias and anthologies which, by reason of the selection and arrangement of their contents, constitute intellectual creations shall be protected as such, without prejudice to the copyright in each of the works forming part of such collections."

These selections alone suggest that, contrary to the question of whether or not AI art breaks copyright, AI art is in fact *protected* under international copyright law as a fixed-form alteration of scrambled noise, conducted to produce an artistic end product; furthermore, web-scraped datasets, as collections of literary or artistic works similar to encyclopedias—but again on a level

[18] Hands-on or online courses, instructional videos, networking resources, and other forms of education.
[19] See: bit.ly/berne-convention-doc+
[20] Permanent or tangible.

so scrambled by noise that they are essentially unintelligible[21]—are also protected under the Berne convention. Let us discuss this particular issue first before moving on to whether or not AI breaks copyright, as this first provision is a necessary supplement to the second.

Despite the protections that the Berne Convention offers, recent events suggest that the copyright offices of countries like the United States hesitate to acknowledge these rights. In the United States, for example, the U.S. Copyright Office (USCO) asserted that the AI-generated art piece *A Recent Entrance to Paradise*, created by Dr. Stephen Thaler, was not eligible for copyright because it lacked human authorship and that copyright law only provides protections to "the fruits of intellectual labour" that are "founded in the creative powers of the human mind".[22]

Dr. Thaler applied the neural network he created, called *The Creativity Machine*, to a series of pictures, repurposing them to simulate a near-death experience as seen by a synthetic dying brain. The result is shown below:

Above: Dr. Thaler's "A Recent Entrance to Paradise"

Ironically, if I were to commission a US artist to copy the image and then have them submit the new work, as made by human hand, it would likely receive copyright protection. If I had access to greater resources and the will to do so, I might even be able to retroactively 'copyright' multiple pieces of AI art and pursue strategic litigation against their artists—either for financial gain, or to harass or censor them.

I then put it to the reader:

- Was Dr. Thaler's AI, and by consequence its art, not the fruit of intellectual labour?
- Was it not founded in the creative powers of his human mind?

[21] And, with original training images no longer needed after a neural network has finished learning, one can imagine a neural network overall as an artist who has read an encyclopedia, learned the alphabet from it, and then used this to write a book in partnership with its prompt engineer. The result is an original work.
[22] See: bit.ly/thaler-copyright+

- Is a neural network truly different from a pen, or a palette, or a word processor, or graphic design software insofar as each is a tool that extends the creative intents of its user?[23]
- Does the work not deserve copyright and does refusing to provide this not create an abusable loophole in the law?

Other cases of AI used for creative purpose can be found in Oxia Palus' recreation of Picasso's lost art, *The Lonesome Crouching Nude*[24] and Ammaar Reshi's children's book, *Alice and Sparkle*.[25] In each case, AI art was intended and created under human direction.

If the discussion outlined above demonstrates that AI art is vulnerable to copyright loopholes, then the conversation below will demonstrate the circumstances under which AI's neural networks might be thought to break copyright, and why certain easily-made assumptions are wrong.

Recall that neural networks function by attempting to clean up entirely random pieces of noise; after learning how to clean noise from a training set, the original images are no longer relied on unless a new model is taught to do some different task. Therefore, copyright cannot be broken by the neural networks performing their tasks on-the-spot because they do so without any real interaction with the images they were trained on.

This is the first and most important point, and it allows us to make parallels between AI's artistic process and that of real-world creators. The closest equivalent to what AI does—learning from many pictures and creating new images based on the styles and characteristics of those images—is **pastiche**, where an artist or writer imitates the styles of their contemporaries. Laurence Sterne, for instance, wrote the novel *Tristram Shandy* in which he took passages nearly word-for-word from other writers, but rearranged them creatively so as to give the passages new meaning.[26]

This raises an interesting question: despite 'forgetting' its original training data, is it possible for a neural network's pastiche to so efficiently recall the stylistic elements of an image that it can faithfully recreate it, and if so, under what circumstances might this happen?

Midjourney's *Midjourney Version 4* release caused a commotion in early December of 2022 when a user realized that their prompts resulted in the famous picture of National Geographic's *Afghan Girl*, a photographic portrait of a refugee taken during the USSR's invasion of Afghanistan in 1984. Prompt words alone appeared to produce at least three near-perfect replicas of the girl during a single image generation (for each /imagine command a user gives, *Midjourney* creates 4 small-scale images that can then later be upscaled to give them greater detail). American vlogger and YouTuber

[23] Dr. Thaler in fact does not need to rely on me for his defense. He writes a philosophical defense of his case here: bit.ly/thaler-creativity-machine+

[24] See: bit.ly/picasso-hidden-art+

[25] See: bit.ly/alice-and-sparkle+

[26] Further from pastiche (and aside from the obvious label of generative art, which has been in use since the 1960s), what an AI does may also be considered an extremely intricate form of collage or décollage. Depending on which comparison is made, different legal conversations might take place as to both the copyrightability and copyright-breaking of AI art.

Hank Green cited this in his video *Things Are Changing Very Fast*, where he claims that: "[*Midjourney Version 4*] has been trained on that image and so it is plagiarizing it." [27] [28]

By this point, I hope that readers who have parsed earlier sections of this essay better understand the nuance of AI and so refrain from reducing its functions to a single sentence. The AI certainly did appear to plagiarize the photo, but Hank Green's explanation is oversimplified.

On the left: the original "Afghan Girl" image by National Geographic's Steve McCurry
On the right: four 'imagined' images, all of which share clear similarities to the original photo.

If Mr. Green's assessment of AI lacks proper depth, then how did it 'steal the Afghan Girl'? The answer is that AI was not trained on the image of the Afghan Girl. It was trained on thousands of it. Recall that neural networks rely on datasets made up of collections of images and labels.

CLIP's zero-shot-styled model scrapes the web to produce a massive and flexible dataset—but it also fails to account for types of redundancy.[29] The more *Midjourney*'s neural network saw the Afghan girl's photograph, taken in the exact same pose and with the exact same expression, the more inclined it was to give greater weightings to what it assumed were similar characteristics on different training images. In addition, 'Afghan Girl' is both a descriptive prompt and the name of the National Geographic photograph. Any /imagine command would resultantly create an image of *the* Afghan Girl, instead of *an* Afghan girl.

[27] Although in reality, not breaking its copyright. See: bit.ly/clarifying-copyright-us+
[28] HankGreen's video: bit.ly/hg-afghan-girl+
[29] While OpenAI has not publicly disclosed how CLIP lessens redundancy in its dataset, it is safe to assume that even with random sampling or other techniques being used to avoid capturing large collections of the same image, having multiple 'similar-enough' copies of a picture would produce redundancy without specifically copying any one picture.

On the left: the result of '/imagine a girl from Afghanistan' on *Midjourney, now producing less-similar results.*
On the right: a page of Google's search results for 'a girl from Afghanistan' for comparison.

In fact, I can reproduce this specific issue with an art piece by Johannes Vermeer and Leonardo da Vinci, via the following prompts:

/imagine a painting of "Girl with a Pearl Earring" in the style of Johannes Vermeer, oil painting

/imagine a painting of "Mona Lisa" in the style of Leonardo da Vinci, sfumato

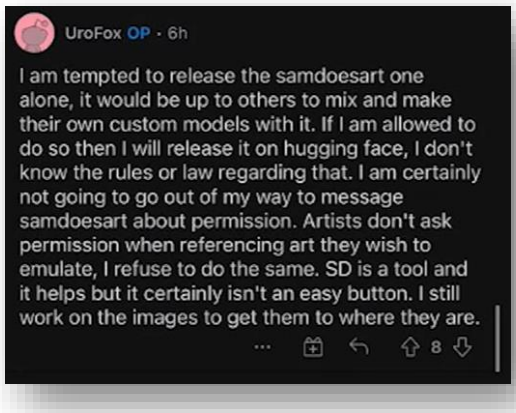
Above: "Girl with a Pearl Earring" and its Midjourney equivalents.

Above: the "Mona Lisa" and its Midjourney equivalents.

Therefore, in a regulatory minefield of nuance and shades of grey, we can consider the framework below for when and how AI affects artists:

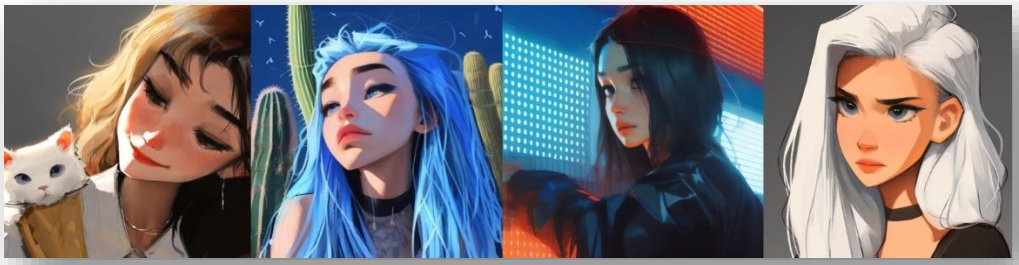
- For small creators who do not to upskill, AI will likely displace them from the workforce.
- For large creators, the popularity or redundancy of their works online, in addition to the qualifiers that the work is drawn in a specific and repeated pose, makes them a target for web-scraping and ultimately AI-driven plagiarism, whether benign or malicious.

A malicious use of AI can be demonstrated in the following case study: in late-2022, a digital artist by the name of SamDoesArt had 300 of his copyrighted works used to train a model of *Stable Diffusion* by an individual user, UroFox.



Above: UroFox's reasoning for training a model on SamDoesArt's artwork.

Because the majority of SamDoesArt's images consist of upper-body shots of females drawn with similar facial characteristics, the neural network model was able to quickly match his art style and composition, producing images that were difficult to tell apart from his.



Above: a reel of four digital paintings from left to right: 1, SamDoesArt; 2, AI-generated; 3, AI-generated; 4, SamDoesArt.

As a large YouTuber and content creator, SamDoesArt posted about UroFox on Instagram, opening UroFox up to harassment by Sam's community; this in turn led to UroFox shutting down their Reddit and Twitter accounts and caused other prompt engineers to react by building new models that specifically targeted Sam's art—with competitions to mimic it now as closely as possible. None of this constitutes copyright-breaking, the AI-generated characters are reasonably unique[30], and human effort went into engineering the creations. Sam cannot claim copyright over his style in the same way that he could not claim copyright over the idea to create his portraits—all he can claim ownership over is any work he has already produced as a result of his ideas and processes.

Perhaps then, at least philosophically-speaking, it is not Sam's art that the malicious prompt engineers were attempting to make facsimiles of—rather it is Sam. This is a point we will return to in section c) under the notion of the 'self' in art, though for now, let it suffice to say that the use of neural networks can and does harm artists, with not the necessity, yet still the potential, to devalue and degrade the artist's moral right to his own experiences, training, and art style.

Overall, when misused, AI models have the potential to lead to reputation damage, forgery, fraud, and identity theft, even though they do not break copyright.[31] This is a valid and important critique of AI art, but it establishes that regulation must aim to find ways of limiting repetition in datasets, rather than to limit the datasets overall—or worse, as a result of misunderstanding AI, to try to regulate art-focused neural networks instead of the datasets they rely on, something that would be tantamount to the difference between surgery and butchery.

It is perhaps with this notion in mind that AI companies such as *Stable Diffusion*'s creator, CompVis, and Midjourney have already shown that they are interested in operating in good faith—Midjourney's neural network processes and categorizes prompt results that may create contested materials and applies bans that filter them out accordingly while better solutions are developed

[30] —with both prompt words and final designs representative of their engineers (or artist's) intent.

[31] As SamDoesArt describes in his own video, outlining his situation: bit.ly/sda-summary+, and which regulators will be all-too-familiar with.

and a recent *Stable Diffusion* update limited the ability to copy artist styles, as well as to produce not-safe-for-work content (i.e. pornography).[32] [33] Likewise, neither they, nor *DALLE*'s OpenAI assert copyright over works produced by their users, and Midjourney even runs their company at near-to-non-profit, choosing instead cycling funds back into improving the technology and moderation behind their users' AI-generated content.

In short, research labs are willing to take the lead, and in cases where bad faith actors may have their own neural networks, new and specific regulatory structures, akin to human rights tribunals and informed by committees of various stakeholders in technology and the arts, may fare better at dealing with case-specific situations of job-loss, copyright infringement, and artist-devaluing than more rigid top-down regulation would.[34] If regulators prefer, they might even consider a partially automated approval system for AI art copyright that can be supported by human-led tribunals; in such systems, AI art—or any art—could be uploaded to a copyright office's own neural network, one made to run the candidate against a database of protected images and videos to check if it qualified for copyright approval. If rejected for approval, the human tribunals could then offer the option of a closer look by a human committee ruling.

Whatever their decision ends up being, regulators must aim to establish a process inclusive of all voices that is flexible enough to protect individual artists while allowing researchers the freedom needed to engineer solutions to copyright problems. This means not only eyeing datasets as opposed to the networks themselves for targeted regulation, but referencing voices from artists, writers, musicians, *and* the data scientists who create neural networks and image/word datasets through web scraping during this legislative process.

c) Are neural networks dehumanizing artists and devaluing art?

One of the most common arguments against AI art on social media is that AI art is not art and that, by masquerading as real art, it thereby dehumanizes artists and devalues the field of art itself. This gave me pause, and for a moment I leaned back quietly in my chair, listening to it creak, staring inconstantly at the empty plates and crumpled papers scattered behind my laptop, as if they were the ruins of a modern mock-Troy. I listened, whether I meant to or not, to a confusing jumble of New Year's celebrations that came from outside my window; a firecracker went off. And then another, and at 1:42am on January 1st of 2023, I sat alone in my room, wondering what art was.

To answer the question of "What is art?" it is perhaps useful to ask: "What sets apart the above paragraph from the rest of this essay?" It is a paragraph I typed on Microsoft Word; I used Google's built-in thesaurus to select and rejig pieces of it; I took the precise time and date from the taskbar on my Windows operating system; I salvaged the idea of Troy from a PDF of the *Iliad*. In that sense,

[32] See the blocked use of '/imagine Afghan Girl' and derogatory or sexual '/imagine's in the Midjourney discord.
[33] See: bit.ly/stable-diffusion-nsfw+
[34] Therefore, a punitive but case-by-case model, informed by human rights frameworks and with stakeholder input may be worth exploring, as it would more adaptively protect individual creators and reduce loophole exploitation by large corporations. It may also potentially be combined with regulation requiring AI art to be accompanied by the prompts that created it.

nothing sets the paragraph apart from the rest of this essay—it uses definitions, words, references, technology, and lastly, it exists within the context of a larger argument. However, that is not its only context.

A second context is that it presents to the reader a glimpse into my life during a moment on an early winter morning that would have otherwise been entirely forgotten. Parts of me bleed through the paragraph to reach you, the reader, and for 89 words you see a part of me. The paragraph above is literature and thereby art.

So, what is art? I believe that art is an injection of the human spirit into the machinery of the world. More specifically, it is an expression of self, whether by intent or coincidence, that uses tools to create a fixed work.[35] Let us pick apart an example of how art and the process of its making converge, along with the contending views that grew up around this process:

In the 15th century, the abbot Johannes Trithemius, who was considered one of the fathers of cryptography, reacted to the invention of the printing press with similar arguments to those who today look to discredit AI art. Dorothea Salo, in her translation of Trithemius' words, lays out his argument:

> "Brothers, no one should think or say 'Why do I have to wear myself out writing by hand, when the art of printing has brought so many books to light, so that we can cheaply put together a great library?' Truly, whoever says this is trying to conceal his own sloth… He who ceases the work of a scribe because of printing is not a true friend of Scripture, because heeding no more than the present he takes no care to educate posterity."[36]

Trithemius saw the printing press as a harbinger of sloth, one that failed to carry forward the typist's self—in his description, the writer's very spirit was bound to future generations through the long and human-intensive exercise of writing works out by hand. His argument here is an **appeal to purity**.[37] In Trithemius' mind, the printing press would devalue the process and thereby the spirit, reducing produced works to mere automation. A survey of the past several decades, from Hemingway's *For Whom the Bell Tolls* to Atwood's *The Handmaid's Tail* showcase just how wrong he was—the human spirit was not devalued; instead, it was emboldened. Ideas were able to be constructed faster, spread further, and as a consequence grow more complex than before. Yet the ideas and their qualities remained firmly in the grasp of the human spirit, and the content remained firmly in the scope of the human self.[38] The invention of the printing press did not bring about a Dark Age but instead a Renaissance that democratized ideas and their sharing amongst the public.

---

[35] At least within the scope of regulation—as ephemeral, or temporary art, is still art too, and furthermore, in an ironic twist of fate, all human works within the scope of the universe are ephemeral.

[36] See: bit.ly/trithemius-vs-gutenberg+

[37] Also known as the "No True Scotsman" logical fallacy. See: bit.ly/no-true-scotsman+

[38] It is also worth noting, though out of respect for the reader's time not worth its own tangent, that Plato believed writing would 'implant forgetfulness' in the human soul and devalue memory. Perhaps he was right, though if so, I can't for the love of me remember why.

In general, as technologies progress, they decentralize the process of creating art and in doing so smudge the boundaries between private and public spheres. This allows art and ideas to propagate and grow more complex as more humans interact with them.

Scripts kept secret and penned by monks became published treatises that were debated by philosophers, which then became books mass-produced by any author for a worldwide audience; mathematicians who once kept secret journals of questions only they could solve later came to steal and exchange their ideas, laying the groundwork for the Enlightenment in Europe, where reason and logic were openly discussed in salons.[39] Centuries later, even the salon gave way to co-authors and the rise of modern research labs—all to discuss increasingly complex ideas. 'Et al.' is not a throwaway phrase. It is an abstraction of the self so that the idea may take center stage.

Art has seen a similar progression as the tools of the artist's trade have grown in complexity while also growing in accessibility.[40] Through neural networks like *Midjourney*, prompt engineers create a list of words that affects what types of form, dimensions, colours, composition, and overall style are used to build an image. By adjusting these prompts, upscaling images to higher resolutions, and remixing images to produce fresh details or remove inaccurate ones, prompt engineers use the neural network as a tool to create art—much in the same way I used the digital tools at my disposal to create my paragraph. Jamboards, learning sessions, user polls, and public prompts meanwhile all help to extend the individual ideas that prompt engineers have into a remarkable digital tapestry, made up of artworks from all around the globe. [41]

[39] See Veritasium's video-explainer on the history of imaginary numbers: bit.ly/veritasium-mathematics-history+

[40] In a later section, I will describe how neural networks present access to art that those with disabilities did not previously have, as well as how it presents affordable solutions to individuals who cannot pay for expensive art.

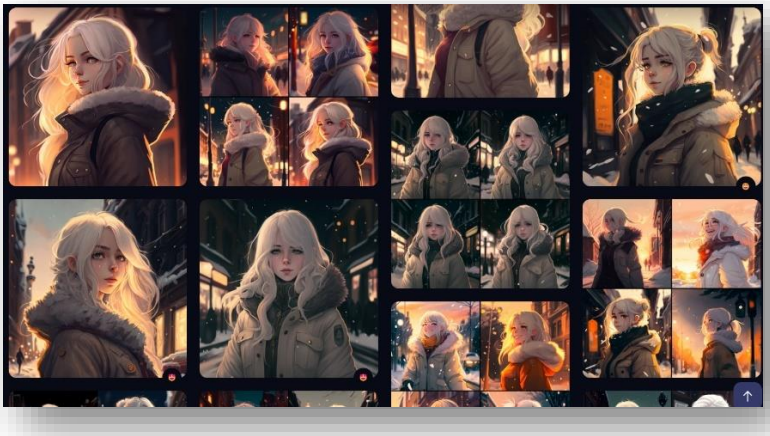
[41] These are all community-driven features of Midjourney.

Despite being conducted through the medium of a complex technology, prompt engineers create a form of art that not only reflects their self but extends it through having learned from millions of words and images. A 'copy' of SamDoesArt's portraits may use the prompt words *'A girl, beautiful and in love, blushing, smiling, white hair, standing on a snowy London street on a December night and looking at you, incandescent light glowing through her hair, accent lighting, line art, anime, Procreate, in the style of SamDoesArt –niji –ar 2:3'* and create a piece of art in his style, but it also pulls from a multitude of concepts and reworkings separate from 'in the style of SamDoesArt', which all imbue the piece with the uniqueness and elements of the prompt engineer's self. If weaponized, this data can and does devalue an artist, and regulators should work to understand how to mitigate its damage[42]—but with this said and at the risk of sounding pedantic, AI art can never devalue art itself, because AI art is art.

To the right: an AI-generated piece of art in the style of SamDoesArt using Midjourney. Below: the process I followed to achieve the piece, including a selection of generative images during some of these steps.

10 initial prompts, with words adjusted at each iteration / 6 upscales taken from this set and 1 variation / 1 new prompt containing the 4 best outcomes of the mid-journey-generated reference images / 2 upscales of this result / 1 new prompt containing the best outcome / 2 upscales from this prompt / 1 prompt containing the best outcome of an older set and an adjustment to the prompt words / 2 upscales from this prompt / 3 new prompts using Midjourney's remix mode / 2 upscales from the 1st prompt, 1 upscale from the 2nd prompt, 2 upscales from the 3rd prompt / 1 beta upscale redo / 1 Remix of an upscale / 1 remaster / 1 detailed upscale redo / 3 upscales from the prompt /1 additional upscale / 1 new prompt with image and word adjustment / 3 upscales from the prompt.

Total time: 1 hours and 16 minutes



[42] Either through the aforementioned tribunals or opt-out systems from datasets.

Discussion: Regulation

This section of regulatory concerns will act as an accompanying speed-round for the more intensive preliminary discussion that was had so far. It is meant to prompt regulators to consider the potential drawbacks to certain styles of legislation and prevent knee-jerk policy moves by outlining implications that might not have initially been thought of, as well as to showcase the positive applications of AI in the arts. It is not a conclusive list, and it is likely that through further discussions, politicians will find ways to overcome or expand on the challenges described below.

1. **The effects of bills passed by a government are often difficult to withdraw or modify. Technology meanwhile is advancing at such a pace that unintended consequences from inert legislation may restrict future research and development of technologies in connected fields, having devastating effects on safety standards, the economy, and marketplace competitiveness.**

   There is a consistent trade-off over time that has led to bloat in government bodies and ineffective policy implementations; no government is free from this threat, and no policy is free from its consequences. Poorly-shaped legislation can act like an artery plaque that misallocates regulatory responsibilities to committees that may be ill-equipped to handle a field as quickly-changing as AI. More consequentially, AI regulation cannot be 'undone' or modified in the traditional sense, in that by the time a government has fully reviewed a specific policy change, the field may have changed drastically, making their discussions irrelevant and prone to creating more policy problems.

   By extension, unintended consequences may arise that hamper progress in other fields that one might not initially associate with neural networks.

   Ex. A medical AI used by a business to detect cancerous tumors may in the future have its datasets limited due to copyright issues or regulations on how the neural network can behave. This has the potential for misdiagnosis or correct diagnosis but only of specific types of patients.[43]

   Ex. Self-driving cars rely on neural network attention maps, a branch of **computer vision**, which studies how computers 'see'. Policy decisions may affect how well computer vision does its job, leading to the potential for an increase in on-the-road accidents.

   Ex. A policing AI trained to detect the likelihood of a person committing a crime (through facial recognition, crowd monitoring, de-noising of security camera data, and other analytics) could mistakenly profile an innocent citizen[44] [45] (although it should be noted that policing neural networks generally rely on one-shot learning, which is different from the zero-shot learning of AI

[43] See: bit.ly/racial-bias-health+
[44] See: bit.ly/racial-bias-policing+
[45] See: *Minority Report (2022)*. And then see it again.

art). AI regulation that does not make distinctions between computer vision for art and for policing may mistakenly craft regulation that is too heavy or too light for both industries.

Ex. Neural networks used to manage monetary and banking policy may find themselves unwittingly limited by regulations that were not intended to affect them, with real-world effects on GDP. For instance, the United States Federal Reserve already uses computer modelling and neural networks that have been developed at each of its regional branches in order to model employment data and surveil the economy and capital markets; likewise, it is researching and developing a decentralized currency model, central bank digital currency (CBDC). [46] [47] [48] It is conceivable that at a point in the future, legislation made to hamper one variety of neural network may lead to issues in others, with ultimate real-world damage to market liquidity, currency efficiency, and GDP.[49]

Ex. In a similar manner to the example above, neural networks used for cybersecurity and national defense may suffer from misregulation, leaving their countries vulnerable to a) attacks from foreign actors, whether state-funded or private, and b) lack of competitiveness in the global $328-billion field of AI—expected to reach thousands of billions of dollars by the end of the decade. This has the potential to be a death knell for a country's economy.

Ex. Further to the point above, AI regulation in one region does not prevent AI misuse in another. For instance, if copyright legislation is decided in Europe that hampers artificial intelligence, there is nothing to stop other countries—India, China, the United States, Canada, etc… from using less-restricted versions of AI in a zero-sum game. International decisions will have to be reached about AI, with international implications for countries who ignore the decisions.[50]

2. AI regulation, if conducted in a partisan manner, will allow corporate entities and copyright trolls to harass content creators, AI and non-AI artists, hobbyists, and research labs with litigation.

Current instances of malicious and predatory litigation prove that any legal framework for AI regulation will be vulnerable to abuse from those with the money, power, and influence to conduct it. Neural network regulation designed in a partisan manner, or in a manner that does not set out clear guidelines for large corporations, will be exploited, for instance to steal art from smaller creators or pursue aggressive litigation against smaller business competitors.

Ex. Copyright legislation that might consider AI art's style-imagining to be a derivative work would open up legal loopholes for companies to create art in various styles, potentially with AI itself, and then claim that any similar creation is a derivative work, or flood the web with their mass-produced

[46] See: bit.ly/machine-learning-fed+
[47] See: bit.ly/cbdc-fed+
[48] See: bit.ly/cbdc-fed-2023+
[49] Governments should consider employing research labs specifically for this task—that is, large-scale penetration testing of vulnerable neural network systems—on a far more regular and robustly-funded basis.
[50] I do appreciate that this is an impossible ask in today's multipolar and open-source world, but efforts to try should still be made.

images, with the ultimate goal of them being caught by a web-scraper. This would end the art career of every independent creator by cutting off all forms of monetizable art except to companies large enough to pay-to-play, or rather pay-to-harass.

If the above example may be filed under the 'let's-not-give-them-any-ideas' category; the following examples showcase the ideas that companies already have:

Ex. The YouTube copyright claim system is currently used by companies to immediately take down videos without regard for creator rights or Fair Use.[51] Toei Animation, an animation company that has produced Dragon Ball Z and One Piece, struck down over 150 YouTube videos belonging to Totally Not Mark, an animation reviewer and critic on YouTube, whose video content met Fair Use guidelines.[52] Without considering Mark's right to fair-use before sending their initial round of takedown claims over the span of a single night, Toei erased 3 years' worth of Mark's work. YouTube eventually ruled in Mark's favour, but the company caused both monetary and emotional damage to Mark through its abuse of copyright protections.[53]

Ex. The Digital Millennium Copyright Act (DMCA), an arguably outdated piece of legislation, is often used to harass individual creators for using snippets of music or sound bytes in their work, even if the snippet or byte does not make up anywhere near the majority of the work. Case-in-point: In August of 2022, Twitch flagged a sound byte of a police siren used in an indie game, leading to the entire audio of a live playthrough event being shut down; notably, any action taken to remedy the Twitch shutdown automatically notifies the legal team that copyrighted the clip in the first place, allowing them access to the notifier's information and opening the notifier up to direct legal action.[54] This is referred to as **copyright trolling**, and YouTube's penguinz0 provides both Twitch and YouTube examples in his video *Why I'm Mad About Twitch DMCA*.[55]

Ex. Wizards of the Coast (WOTC) has sought to shore up Hasbro's stock price by monetizing the rules and all profitable offshoots belonging to its table-top game *Dungeons and Dragons* through an updated Open-Game License. This update requires individuals or companies making over $50,000 through their own 'homebrew' games to report revenue to WOTC and place a company branding label on their products[56]; it also references videos (such as game reviews and playthroughs) in a manner that would expose fair and open reviews of its content to malicious litigation in a manner similar to the situation between Toei and Totally Not Mark. Smaller outlets like Kobold Press rely on freelancers to make the offshoot games they sell, and these costs will likely result in lower sales, stifled creative artistry from hobbyists (who wish to create works only tangentially-related to D&D's game rules), and more expensive products for consumers.[57] [58]

[51] See: bit.ly/youtube-copyright-trolling+
[52] See: bit.ly/tnm-toei+
[53] See: bit.ly/toei-tnm+
[54] See: bit.ly/siren-sound-copyrighted+
[55] See: bit.ly/DMCA-penguinz0+ (1:45 onward)
[56] See: bit.ly/ogl-dnd+
[57] See: bit.ly/dnd-royalties-clownfish+ (9:00 onward)
[58] See: bit.ly/twitter-post-ogl-summary+

Ex. A non-creatives instance of copyright trolling might occur where competitors of smaller research labs who derive funding through their neural networks file multiple copyright lawsuits against those labs for infringement in order to drive them out of business. Whether the lawsuits were won or lost, the implications could damage the neural network lab's reputation and damage its pocketbook through limiting access to funding programs and drying up its budget with legal fees. In a situation like this, big-fish-eats-little-fish style litigation would ensure the destruction of research independence and solidify the power that monopolies have over AI, rendering them even more powerful than government.

3. Following from the concerns above, the largest real-time threat to AI legislation is that corporate entities may astroturf anti-AI movements to push for specific forms of regulation that favour their own neural networks while harming smaller companies and individuals who may compete with them.

The abuse of legal loopholes starts when the loopholes are made. I suggest to the members of regulatory bodies that companies—especially those who serve to gain advantage over competitors, individual creators, and even government—may suggest minute or incomprehensible changes to AI regulation during its crafting via lobbying and, potentially, astroturfing[59] anti-AI movements in order to produce the appearance of 'from-the-ground-up' support for regulation that is in their favour. Though astroturfing is notoriously difficult to prove, groups such as the Concept Art Association already appear to be using erroneous information about AI to form their own policy stances.[60] Below are provided two real-world examples of lobbying concerns.

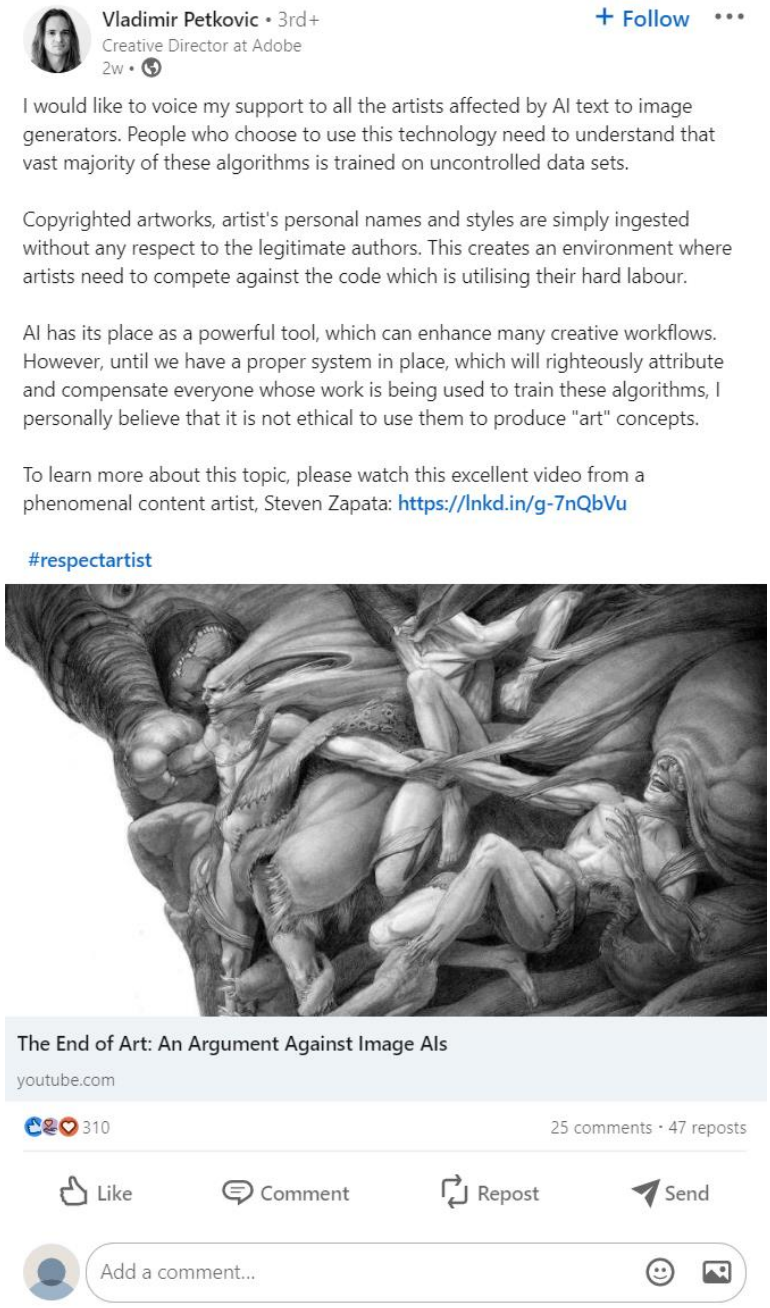
Ex. Alphabet Inc., the parent company of Google, has a total market capitalization of $1.14 trillion. It has both the lobbying strength and the will to affect its lobbying in the American Congress. In December of 2022, the New York Times reported that Google was issuing a 'Code Red' over OpenAI's *ChatGPT*, worrying that it would prove a likely competitor to its own Google Search. [61] As a reminder, roughly 80% of Google's revenue is derived from the ads it runs on its services, with 57% coming directly from search ads.[62] [63] Users will ultimately be able to ask ChatGPT direct questions and receive direct answers—without sponsored results appearing. This would eviscerate the company's income. Google's only practical move would be to lobby for specificly restrictive legislation to hamper OpenAI and similar AI companies.

[59] That is, co-opting in a way that is not immediately visible.
[60] bit.ly/concept-art-association+
[61] bit.ly/google-red-alert+
[62] bit.ly/google-income-breakdown-01+
[63] bit.ly/google-income-breakdown-02+

Ex. In similar fashion, programs such as *Midjourney* and *DALLE* present active threats to Adobe's business model by generating images from scratch as opposed to gathering, merging, and touching up pre-existing images or creating them in the Adobe suite. Adobe appears to have positioned itself on the side of creatives by using social media and its independent team members in order to announce its antagonism towards text-to-image AI (see following page). Ironically, Adobe has concurrently been working on an AI suite in photoshop that has similar features to *Midjourney* and *DALLE*—for instance, the ability to replace part of a photo with an AI-generated image.[64]

Above: an example of Adobe's AI software in which a user can click-and-drag to select a target area, then prompt engineer a barn owl to appear in that area.

[64] bit.ly/adobe-ai-art+

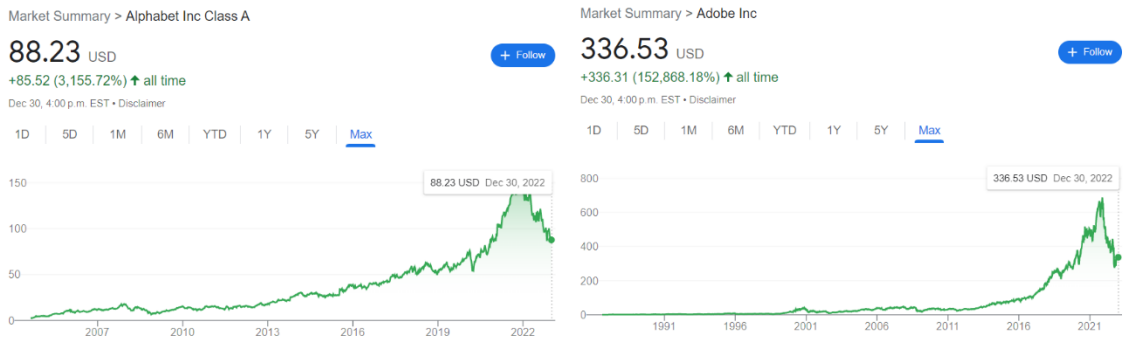


Above: Vladmir Petkovic, Adobe Creative Director, appears to posture on social media. [65] *Incidentally, Adobe still accepts AI-generated art submitted through contributor.stock.adobe.com and receives 33% commission for images; 35% for videos when artists purchase the images.*[66]

[65] bit.ly/petkovic-statement+
[66] bit.ly/adobe-pay-breakdown+

The Adobe of today relies on its creatives ecosystem for income, despite creatives originally having resisted its own Photoshop program when it was first released.[67] By posturing publicly while failing to inform the public about its conflicts of interest, Adobe lays the groundwork for astroturfing of anti-AI movements and voices.

While there are various companies that sit on all sides of the AI issue—and within each company, there may be varying opinions between employees—it is undoubtable that every company, especially those trading publicly, will seek to exploit legislation for its own benefit.



Above: Google Finance charts for both Alphabet Inc. and Adobe Inc., demonstrating the tender position 2022 has left each in. Wounded companies are often dangerous ones; they will cut no corners to ensure their survival.

4. **Legislation cannot adopt a 'Good v. Evil' mindset. Harassment of AI art prompt engineers occurs in the same way that harassment of artists occurs, and both are valid experiences worth considering, though more often, Anti-AI art sentiment is signal boosted in the media.**

Regulators must understand that as this new technology develops, the loudest voices will usually be the angriest and most reactionary. That said, these types of voices do not always represent the majority of opinions on the matter. In the same way that artists are unjustly harassed by prompt engineers, those who adopt AI face severe stigmatization from the art community, and on the following page, a selection of various topics is showcased: death threats, witch hunts against real artists who were mistaken for prompt engineers, attempts to goad corporations into lawsuits against AI art, intentionally flagging real art as AI art, and lobbying war-chests appearing to be raised under misleading fundraisers.

[67] Similarly, on the argumentative basis that it replaced artists.

323

· Posted by 5 hours ago

[TW: DEATH THREAT] And they say that "AI bros" are the ones harassing the artists?

imgur.com/aiH2kn...

nsfw

535

· Posted by 3 days ago

What's up with /r/art apparently removing real artists work and claiming its ai?

Answered

https://twitter.com/benmoran_artist/status/1607760145496576003

The subreddit mods appear to have shut down all discussion on it and it isn't clear why

192 Comments Award Share Save ···

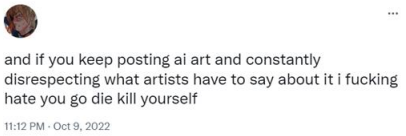
· 3 days ago

Answer: Ben Moran is a Vietnamese artist who does commissioned digital character art, such as this and this. The mod team from /r/Art banned the artist under the premise that he had submitted an AI generated artwork and passed it off as his own. When he challenged this on Tuesday, the team issued a personal statement that they did not see value in the quality of his work.

The official response from the mod team:

> I don't believe you. Even if you did "paint" it yourself, it's so obviously an AI-generated design that it doesn't matter. If you really are a "serious" artist, then you need to find a different style, because A) no one is going to believe when you say it's not AI, and B) the AI can do better in seconds what might take you hours.
>
> Sorry, it's the way of the world.

Unfortunately, it now appears that anyone asking about, or found even linking to the subreddit while discussing the ban in other communities, is getting banned from /r/Art themselves.

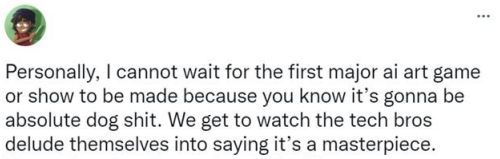


Yesterday, a GoFundMe claiming to "protect artists from AI technologies" launched with a $200k goal.

Turns out, it's actually a scheme to expand corporate IP law on behalf of some of the biggest enemies in the fight for artistic freedom (i.e, Disney)

It's already raised $70k.

bit.ly/lobbing-full-thread+



237

Posted by 2 months ago

The AI art witch hunt has begun

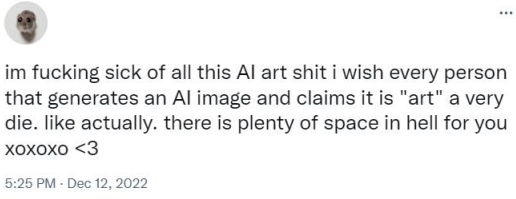
AI Discussion nsfw

Non-AI Artists on Pixiv are getting their artworks tagged with AI tags by other users, while also receiving accusations such as "this was made by AI!".

https://www.sankakucomplex.com/2022/10/22/ai-art-addiction-continues-to-sabotage-pixiv-artists/ (warning, NSFW site!!)

Just because some artworks have generic backgrounds and weird ankles doesn't mean that they were made by AI, please don't point fingers until you are certain that the artist has used AI to make the art.

77 Comments Award Share ···



In addition to the harassment from social media users and artists, media outlets appear to exacerbate Anti-AI sentiment, with preferential coverage and interviews given to voices arguing against AI and AI art. In November of 2022, Black Label Art Cult, a professional artist and AI advocate with exhibitions in Rome and LA, took the time to create an open letter and video directed at major American media companies to address their problematic biases:

{ twitter.com/i/status/1589649316599443456 }

In the letter, he addresses that:

- few if any prompt engineers have been interviewed
- little coverage has been given to AI as an explorative tool for disabled artists
- media writers have based exaggerated articles on faulty premises about AI's process.

Between unfair media coverage and outrage on social media, both of which serve to amplify each other in a vicious cycle, productive conversations between artists and prompt engineers often find themselves drowned out in favour of rage-bait and op-eds that are designed to evoke fear.[68]

5. Using a neural network is not necessarily a one-way interaction. Legislation and restrictions on how AI gathers its datasets have the potential to lead to faulty networks that may not only produce less efficient programs but programs that unintentionally direct or homogenize human behaviour.

A funny thing happened on the way to the Neuromancer. In the final hours of 2022, *BeInCrypto*, an online cryptocurrency news site, released an article where it discussed the potential for chatbot use in Meta's very own *Metaverse*. Citing reports from earlier in the year that Decentraland and The Sandbox—two *Metaverse* virtual reality spaces—had less than 1000 daily active users despite their $1 billion valuations, it suggested that chatbots that looked and felt organic could be used to bolster the illusion of reality in the *Metaverse*.[69]

To be clear, there is no evidence that the *Metaverse* has engaged in this practice of perception-warping, though there certainly exists the possibility of such a strategy being employed by a social media company. As a tie-on to this, regulators should also ask themselves how much of their own lives is determined by predictive analytics. Predictive analytics is a business process whereby neural networks collect large swaths of data about a person, constructing a profile of them—patterns of their likes, dislikes, hobbies, mannerisms, beliefs.[70] The network then feeds this information to businesses in order to drive purchasing of products in a process called surveillance capitalism.

If surveillance capitalism can be used to drive buying, what other behaviours can it affect? That is, if individuals interact with neural networks without knowing it, is it possible that those neural

[68] For instance, the Los Angeles Times' opinion piece by Molly Crabapple: bit.ly/beware-robot-art+
[69] bit.ly/meta-chatbots+
[70] Recall previous conversations about the regulatory danger of failing to discriminate between policing neural networks and image-generating neural networks; a similar issue arises here with predictive neural networks.

networks change their behaviour?[71] Furthermore, is it possible that neural networks that do *not* interact with individuals are still able to change their behaviour?

Perhaps you did not really want to buy that knife set for the holidays. Perhaps social media really is making you angrier than you should be about politics. Perhaps it wasn't really your idea to adopt that puppy. Perhaps you did not really want to kill yourself. A research paper by Amir Dezfouli and his colleagues in 2020 outlines the theoretics of this very real threat.[72] [73]

Even without particular direction, AI—and specifically AI applied to social media—has the potential to resocialize and homogenize entire blocs of a population, harming independent thought, belief, and politics, and creating schisms based on patterns of in-group/out-group preference.

There is a valuable distinction to make here between AI art and AI-politik: each rebukes the other. Art finds pleasure in its distinction and naturally includes elements of unique selves, as outlined in previous sections of this essay; yet a person's political views often find comfort in groupthink, or the act of gathering, agreeing, and complying with those who share similar thoughts; for many, it is how they build a sense of community. The concept of shared evaluation is explored in a paper by Charles Efferson and Sojna Vogt, which for length-constraints is not fully explored here but is nonetheless a critical read for those interested the social mechanics of groupthink.[74]

It might surprise regulators that a critique as harsh as the one above exists in an essay titled 'A Defense of Artificial Intelligence', but these things are entirely congruent. AI for the benefit of humanity and human expression deserves a fair defense when placed under the regulatory spotlight—this includes highlighting key concerns about how mind and machine work together. They are tough concerns to reckon with. They are even tougher concerns to address. And they will always force us to learn more about ourselves than we ever wanted. Nonetheless, the end result of the long path to regulation should be to help neural networks act in a flexible but open and transparent way.

6. Limiting training sets for AI and enforcing strict copyright policing will lead to a number of equity concerns, building prejudice into neural networks through law.

Returning to the Berne Convention document, Article 7-(1) discusses a term of protection for a copyrighted work; more generally, countries tend to offer a term of the creator's lifespan plus 50 years for their works' copyright protection. Depending on how strongly copyright enforcement over neural networks becomes, it is possible that well-intentioned protection of creators' intellectual property might entirely exclude their voices from the world of digital infrastructure. On

[71] To date, much of the technology we interact with already influences our behaviour. TikTok's algorithm already reinforces certain user behaviours. See: WSJ's exposé, bit.ly/wsj-tiktok+

[72] bit.ly/decision-vulnerabilities+

[73] It should also be asked what might happen if malicious actors, whether foreign adversaries or a country's own government, apply neural networks to conduct psychological operations (PSYOPS) or, on a greater scale, conduct psychological warfare. Election results may be manipulated by AI; regime changes may as well. A new and open-source schematic of democracy will ultimately be required, but that is outside the scope of this article.

[74] bit.ly/behavioural-royal-society+

a larger scale, this might lead to the exclusion of entire cultural works from neural networks, resulting in inequitable models. Concrete examples are provided below for the reader's consideration:

Ex. A poorly-designed copyright system restricts a neural network's web-scraping capabilities in such a way that the diffusion model trained by it is only able to produce art styles from over 50 years ago. As the model is adopted by companies in art and other industries for day-to-day use, this stylistic bias spreads, creating a sort of faux-Renaissance that may have its own effect on human culture. Though unique in their own right, the various faux-styles would effectively replace modern, human styles by merit of being promoted through other neural networks, such as those used for social media, who might also believe that to promote such a type of art would be in compliance with the copyright regulation.

Ex. A medical company may use neural networks to inform a database of CAS-9[75]-based cures for diseases, allowing them to mass-patent a number of potential cures and stifle research from other companies while they trial each of these potential cures on their own. While CRISPR is IP non-exclusive for medical research, cures can be patent-locked or patent-trolled. This would allow a company to corner the market on these curative datasets.

Ex. Concerns related to business and policing neural networks that apply to this section that were listed under Section 1. However, a more pervasive and less obvious form of social profiling fits into schemes of coverage and approvals—such as is the case with insurance companies who build risk profiles of clients, retail sellers who build customer profiles of their consumers, and financial and monetary institutions who use datasets to govern access to credit.[76]

7. For many, AI bolsters equity, acting as a tool to help artists with disabilities, reintroduce marginalized voices into art spaces, and remove barriers to access for low-income creators.

For artists with disabilities, technology can often empower the creative process and exploration of the self by eliminating barriers to access. For instance, an artist suffering from cerebral palsy who finds it difficult to create smooth lines is able to construct their portrait in *Midjourney* by typing; an artist paralyzed from the neck down can use vocal prompts to generate characteristics of their art, including finer details.

[75] The enzyme delivery system that allows gene modification in CRISPR.

[76] Note that, rather frustratingly, poor regulation, a lack of regulation, or over-regulation could all result in a social credit system similar to China's own CSC, and each could leave loopholes for either abuse of dysfunction of that system.

Ex. Direct instances of AI's use for creators with physical challenges are provided below. These first-hand accounts come from conversations of mine with users on Reddit and Discord, and they are well-worth exploring to understand the positive impact that the technology has on artists.[77]

· 18 days ago

Hello!

I'm an artist with carpal tunnel syndrome and other nerve issues in my hands.

Those issues have really prevented me from being able to draw or paint to any degree I could find satisfying. I went through art school, I have a BFA, I spent -years- trying to improve my skills at both only to make very little progress due to the issue with my hands.

AI art has finally let me start producing the kind of images I've always wanted to. I cannot begin to express how wonderful and liberating it is to be able to do that!

12 Reply Share Report Save Follow

· 19 days ago

My wife, has something called aphantasia. It's the inability to picture or form mental images of things that are not present.

She used to be able to draw, but lost the ability. AI art, allows her to express herself again in ways she lost.

It's also true of people who have lost limbs, have nerve damage or shaking, colour blindness, and an entire segment of the population who can not afford traditional mediums, and the thousands of dollars it can cost in materials lost and mistakes made.

Imagine a blind person being able to paint again...

Any artist who wants this tech shut down, is doing so for purely selfish reasons. There are billions of people overseas, many of whom can already draw much better than what the ai can output, and certainly much better than they can... and I don't see headlines of art brigades going over seas to snuff out budding art movements... or take art supplies away from their would be competition. It's sad, and myopic.

8 Reply Share Report Save Follow

[77] References have been taken with permission of the posters and follow-up discussions about their impairments. Though informal, they provide a strong start a more rigorous conversation about how AI can be used to empower creators with disabilities.

· 10 min read · ✦ Member-only · Listen

Gatekeeping in the Art World and AI Tools as Accessibility

What is art? Ask a million people; get a million different answers. Ask people in the art world what art is, and those answers will start to contain words like *imagination*, *creativity*, *craft*, *process*, *soul*, and maybe *effort* or even *suffering*.

...

Having cognitive and neurological disabilities, intellectual disabilities, learning disabilities... for a long time, the art world has kept a hard line on gatekeeping "what is art" based on time spent by the artist learning how to use physical and/or digital tools. (Or, if you know the right people, how much money someone is willing to pay you for something that others may or may not consider to be "art.")

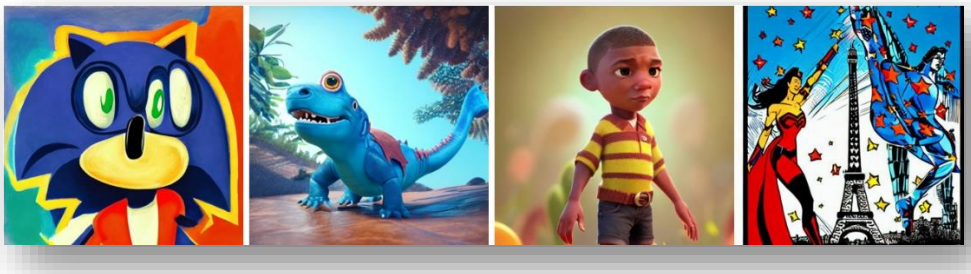
If you have a brain and body that doesn't allow you to learn how to use these tools work, or to use them with any level of competence, that's a barrier that won't go away with any amount of effort. Art is part of what makes us human, and for those who have the yearning to create art based on what's in their souls and their imaginations, it is unbearably cruel to deny that ability to people who finally have assistive technology that allow them to express themselves in the manner of their choosing?

AI tools level the playing field. I see pictures in my head that I can't reproduce on paper or canvas, that I can't take a photograph of or understand how I could use Photoshop to make. But as a writer (a type of artist, mind you!) I can use words to describe the parts of my imagination I want to share. Each type of AI tool uses a slightly different language model, and you have to know how to "talk" to it to tell it what to do.

Ex. In her 2019 lecture at Carnegie Mellon University, transmedia artist and professor Stephanie Dinkins described the ability of AI art to act as a tool for removing barriers between gender, race, and social status, referring to an art experiment of hers called *Not The Only One (N'TOO)*.[78] In N'TOO, a recursive neural network was trained with a dataset of conversations between Stephanie Dinkins and members of her family, producing a small-data model of voices from three distinct generations. Visitors of the display could communicate with the AI, asking questions and hearing responses that reflected the voices of the black American family. The interactivity of this digital memoir showcases a valuable outlet for awareness of underrepresented groups of people, especially within one's own community, by recalling history in both a personal and historically meaningful way.

Above: Stephanie Dinkins' interactive art piece, Not The Only One.

Ex. The *Espace Numérique* team in the French city of Rennes advocates for digital literacy through educational workshop events. Recently, it has started hosting a series of activities centered around using Replicate and Playground AI for modelling, Kahoot for icebreakers and quizzes, and Tencent's ARC for upscaling, in order to provide children from low-income neighbourhoods access to AI technologies. Children learn to work with the programs by generating Pokémon and other pictures using a computer; they then upscale and print out their work, participating in art shows where children and parents can view the art creations. *Espace*'s workshops provides a clear and direct way for at-risk youth to onboard into digital citizenship, while also allowing them to learn about and experiment with developing their own art styles. The program runs from January 12th to February 9th, 2023.



Above: shared student art pieces from the Espace Numerique program.

[78] See: bit.ly/ai-social-equity+

8. Is there a best model for regulation moving forward?

Several thoughts on regulation have been mentioned at points throughout this essay. Let us take a moment to collect and review them all here:

- Upskill programs that teach artists how to work with AI tools can be incorporated as a fundable alternative to strict and sole regulation; these programs can be run by AI art research labs in partnership with governmental bodies (through oversight, direct participation, or subsidies, depending on how involved a government wishes to be); notably, the training programs can also be extended to citizens from technologically under-developed or low-income areas to acquaint them with these neural network models and help ease their path to digital citizenship.[79]
- For regulation that does move forward, great care must be given to understanding the unseen impacts of restricting technology in one area such as art, only to have those restrictions affect another area—such as economics, healthcare, social constructs and culture, equity, and civil liberties.
- To this end, regulation cannot be inert and must provide outlets for timely changes and updates; failing to provide this basis opens a country up to unintentionally handicapping its own social, economic, and technological progress and damages its competitiveness on the world stage.
- At the heart of discussions, regulators should also consider the rights of prompt engineers who create art through a self-involved process that utilizes technology to extend their creative intentions; these engineers have as much a right to protections of their art as traditional artists.
- Nonetheless, care must be taken to ensure that there are bodies in place to oversee instances where traditional artists and prompt engineers have their rights abused—for instance, in the case where AI art closely matches a human artist's work, or where AI meets copyright criteria and therefore is deserving of legal copyright protections. This type of oversight is best conducted on a case-by-case basis and is therefore an appropriate fit for tribunals made up of stakeholders from the fields of both neural networking and traditional arts, or via automated approval systems that give rejected creators the option to have a human committee review their copyright application.[80]
- Overall, regulators should remain aware that companies and individual actors may seek to influence legislation for their own benefit. A thorough understanding of AI and of the nuance in legislation's wording by both regulators and the general public will be required as governments build out regulatory strategies for the oversight of AI.

[79] Federal regulators may even wish to offload day-to-day operations to municipalities, which already have library infrastructure available to them and can retrofit these hubs with upskilling programs and certification offerings.

[80] With the challenges and abusabilities of YouTube and Twitch's automated copyright systems now in hindsight, I hope regulators appreciate why the caveat of a *human* committee as a required reviewal option has been mentioned throughout this article.

Conclusion

The present discourse on AI in the artistic fields often appears to pit two large groups—anti-AI artists and pro-AI artists—against each other. In reality, many layers of nuance exist, with a variety of individuals and organizations in each camp and a variety of interpretations about what constitutes art, fair use, and copyright. None of this nuance should be lost in the uproar against artificial intelligence, and a constructive approach that involves input from both sides is still the best solution for moving forward on AI regulation.

I would like to leave readers with a final thought: 64,000 years ago, in a cave in Cáceres, early Neanderthals blew iron oxides through their fingers to make stencils on the rocky walls of their home. If you were to enter the cave today and look up, you would be greeted with a constellation of art that stood as a reminder of the families who once lived there and hunted side-by-side. In the centuries since, the field of art has evolved alongside humans, and though AI art may displace traditional creative methods, it is far more likely to extend the human communities that use it and the human element that is expressed through it, than to destroy either of these things.

I and many other creators believe in the ability of AI to improve the fields we work in. With each day that passes, we find ways to apply neural network technologies to the goal of expressing the world and ideas around us, much like our Neanderthal relatives. These applications are valid. They create art, and they create a sense of community, and it is on behalf of the communities who use AI art that I submit this essay in its defense.

A Midjourney Community Mosaic

On July 12th of 2022, Midjourney, a 10-person start-up business, released artificial intelligence software by the same name that was capable of taking written prompts from a user and creating, or 'imagining' what those prompts would look like as a picture. Since then, over 10 million people have adopted the *Midjourney* model as a creative platform. Artists from around the globe come together to create all sorts of strange and fantastical art.

I asked the community to submit pictures for a group mosaic. With over 350 images submitted, the result is a community testament to this spectacular and ever-growing field:

Glossary

Artificial Intelligence - the development of computer systems to perform tasks that usually require human intelligence, such as recognizing speech, making decisions, and recognizing objects.

CLIP, or Contrastive Language-Image Pre-training Model - a neural network model that learns to associate images with text descriptions by predicting which image matches which text; this model can be layered over other neural networks to enhance their abilities.

CNNs, or Convolutional Neural Networks or, ConvNet - a type of neural network that is often used for image processing tasks, particularly for tasks like image classification.

Computer Vision - the field of study focused on enabling machines to interpret and understand visual information from the world around them.

Copyright Trolling - the practice of sending out copyright infringement notices or lawsuits for the purpose of generating revenue or intimidating individuals into settling the claims.

De-noising - a technique for iteratively removing noise from a piece of data. In image processing, de-noising is used to remove unwanted artifacts or distortions from images or to generate a new image from an entirely random piece of noisy data. It works hand-in-hand with inpainting, which is a technique that restores parts of damaged images (Ex. Photoshop's 'Content Aware' feature).

Diffusion - a mathematical process of smoothing out an image by iteratively spreading its information across neighboring pixels. Diffusion can reduce noise in a given image, though it is not specifically designed for that purpose. Instead, de-noising is the primary strategy that image-generating diffusion models employ to clean noisy data. In this respect, diffusion can be used to support de-noising, especially for tasks like edge detection, segmentation, and overall image enhancement.

Diffusion Model - a type of neural network model that is uses statistics to analyze how information or materials distribute through a system over time. Via diffusion, the model can extract relevant features and patterns from training data, making it a powerful tool for a wide range of applications.

Discriminator - a component of a GAN that is responsible for distinguishing between real and fake data. The discriminator is trained to classify inputs as either real or generated.

GAN, or General Adversarial Network - a type of neural network that consists of two components: a generator and a discriminator. The generator creates new data from noise, while the discriminator attempts to guess if this data matches data from its training set.

Generator - a component of a GAN that is responsible for creating new data instances. The generator learns to create outputs that are similar to the training data that it is fed by researchers.

Gradients – mathematical 'rules' used to optimize the parameters of a model during its training. Gradients measure the slope of the loss function with respect to each parameter, which is then used to update the parameters to minimize the loss. To decrease the gradient is to increase the model's convergence on a desired outcome.

Groupthink - a phenomenon where a group of individuals conform to the opinions or beliefs of the group, often at the expense of critical thinking and independent decision making.

Model - a mathematical representation of a machine learning system that can be trained on data to make predictions or decisions. A model consists of a set of parameters that are adjusted during training to optimize performance on a specific task.

Neural Network - a type of machine learning model that is inspired by the structure of the brain. Neural networks consist of layers of interconnected nodes or neurons that process and transform inputs to generate outputs.

NLP, or Natural Language Processing - the field of study focused on enabling machines to interpret and understand human language. NLP is used in applications such as language translation, sentiment analysis, and chatbots.

Pastiche - an artwork or composition that imitates the style of another artist or genre, often for the purpose of paying homage or parodying the original work.

Pooling - a technique used in convolutional neural networks wherein an input is divided into subsections. Those subsections and information about them are summarized and assigned statistics, which are then used to represent them on the output of the neural network. Pooling helps make neural networks more efficient by aiding them in pattern-recognition.

Predictive Analytics - the practice of using data, statistical algorithms, and machine learning models to identify the likelihood of future outcomes based on historical data. Predictive analytics are well-suited to business, finance, and government, where they can be used to build profiles of people according to their online and offline habits.

Prompt Engineer - a person who designs and creates prompts for natural language processing models, particularly for language generation tasks.

Surveillance Capitalism - the practice of using personal data collected from individuals, usually through digital platforms, for commercial gain, often without the explicit consent or awareness of the individuals.

Transformers - a type of neural network architecture that is used for natural language processing tasks. Transformers are particularly useful for tasks that require modeling long-range dependencies in text.

Zero-shot Learning - a type of machine learning that allows a model to recognize new classes of objects without explicit training data. In zero-shot learning, a model is trained on a set of seen classes and then used to recognize new, unseen classes based on their attributes.