

**An Artist's Guide to...**

**AI ART**

# Table of Contents



## Introduction

## AI in Creative Careers

- Industrial, Product Design
- Fashion
- Graphic Design
- Fine Art
- Art Analysis

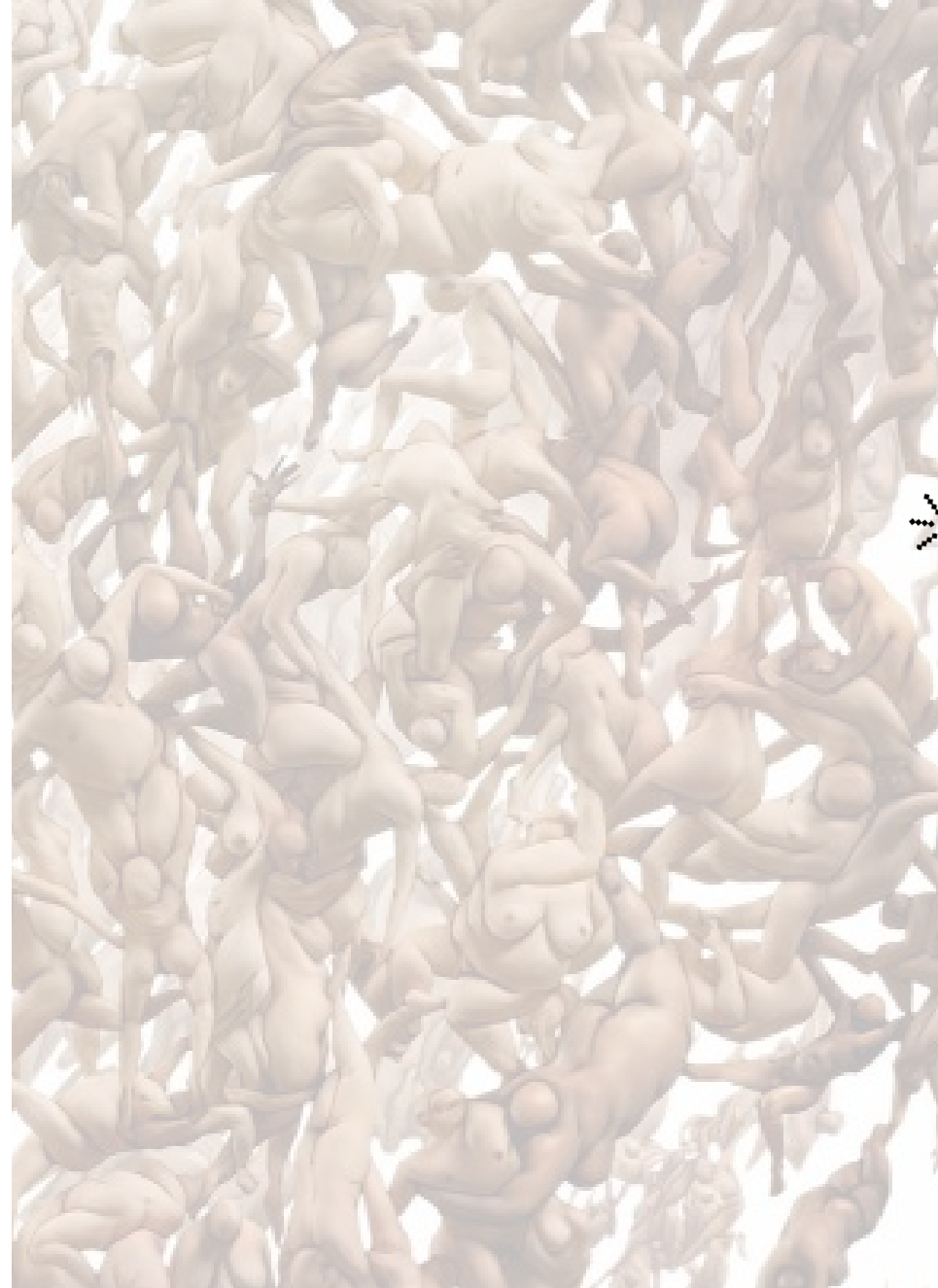
## Where Am I Coming From?

## History

- Computer Art
- Deep Dream
- GANs
- DALL-E
- VQGAN CLIP
- CLIP Diffusion

## Ownership

## Accessible Tools



Click on these three dots and you will be brought back to this main menu



Most of the images are hyperlinked with more resources

Throughout the article you will find orbs of color. These will link you to an area with more information

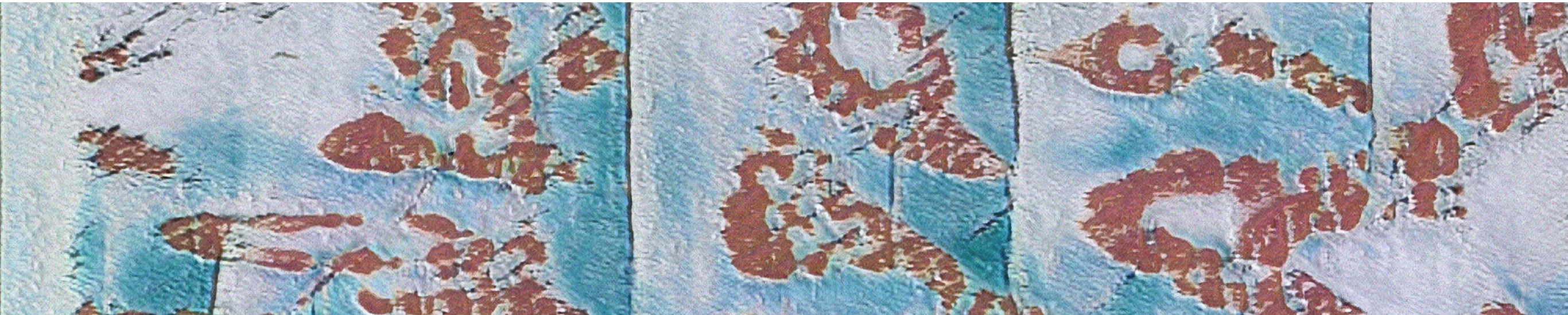


# Introduction



Artificial intelligence and the art world have many different intersections. One of these critical overlaps lies in the community's emphasis on sharing and collective growth. The pillars of open-sourced code and projects in the programming community mirror the collaboration that is present in the art world. This parallel emphasis on sharing information is what has pushed me to write this paper.

I was first introduced to AI with a background in humanities and fine art; it was incredibly hard to find organized information that was digestible while also being technical. We need accessible resources to understand these tools, how they can be used, and how we should evaluate them. It is especially important for artists and creatives to have access to these tools because AI will be affecting all creative careers. It already is.





# AI in Creative Careers



From the runway to the design room to the auction house, AI has become a tool widely used to push the boundaries of creation. A perfect example of this is the introduction of generative design. Generative design employs a system, usually a machine learning system, that has some degree of autonomy in the design process. It can be involved in idea generation, product creation, or both. The designer creates the end goals, sets the parameters, and measures the success of the output. The designer becomes the end decision-maker. This cuts down busy work while enabling more innovative outcomes.

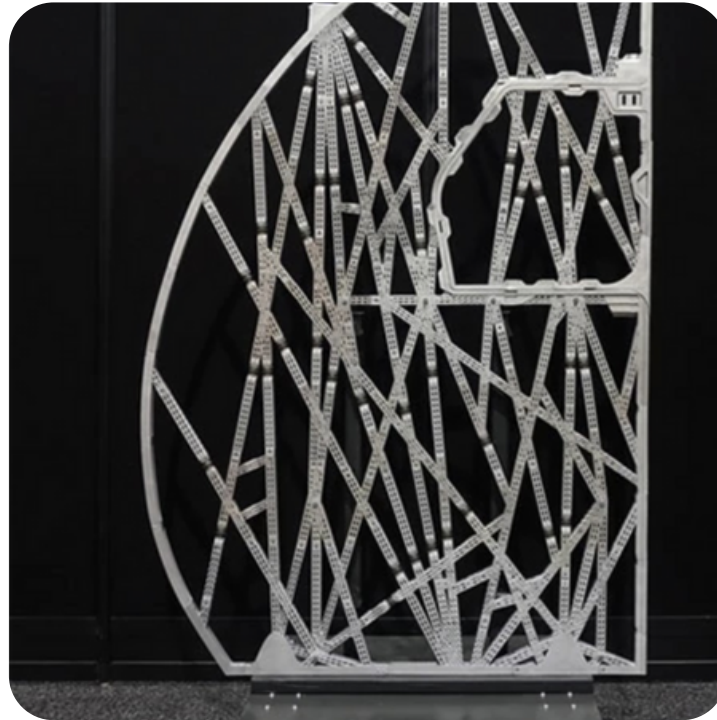
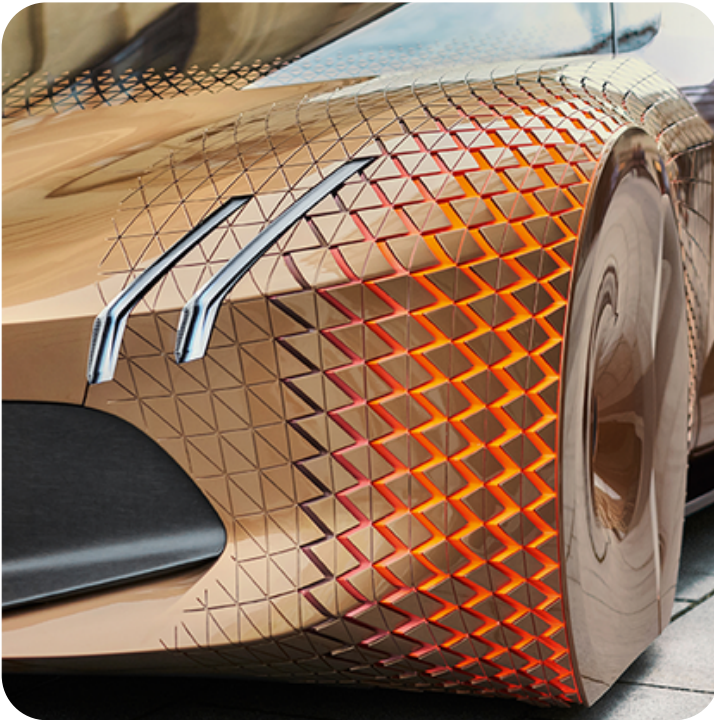
"AI Design will allow creatives to concentrate on their real jobs"  
- Apple CEO Tim Cook



# Industrial and Product Design



A perfect example of the use of generative design is found in industrial, product, and architectural design. Using AI in the design room allows us to approach a problem from an entirely different perspective. It allows us to go beyond the limitations of human creativity. Companies like Adobe and Autodesk have popular software that are commonly used by designers and programmers alike.



# Fashion



AI has already been integrated into the world of fashion, one of the most personal forms of design. As one of the biggest markets in the world, fashion has many areas that benefit from AI. From design, manufacturing, logistic supply chain, marketing to consumer experience, recommendations, and styling AI is guaranteed a big spot in the fashion industry.

Just one of the many examples of AI used in fashion is the ACNE Studios collaboration with AI artist Robbie Barrat to create their fall 2020 line. The artist trained the machine learning model on the archive of ACNE Studios' previous designs and used a model called GANs to generate brand new designs based on the archive. Those designs were then fabricated and shown on the runway.





# Graphic Design



And finally, graphic design is innately connected to technology. One of the many examples of machine learning used in graphic design is when the Nutella brand used a neural network to design product packaging. After the neural network was trained on designs, colors, and patterns, it then created 7 million different packaging designs. This highlights the sheer scale that AI can work on. Generative design like this opens up new possibilities for unique graphics. If this was done by humans it would be super labor-intensive and it would cost a lot of money.



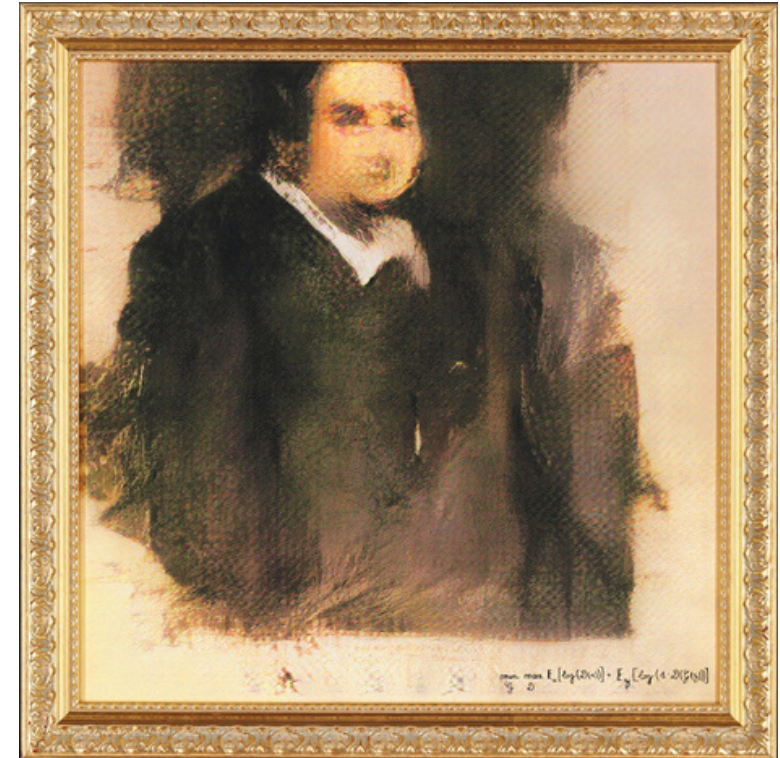
# Fine Arts



Not only is technology affecting the design world, but it has also entered the fine arts world. In 2018, Christie's auction house sold the first piece of AI art for 432,500 dollars, 40 times the estimated amount. The piece, "Portrait of Edmond Belamy", is digitally printed on canvas and placed in a decadent gold frame. The print was made by the French art collective called Obvious. The explosion of this sale stirred many feelings within the art community. The AI artist community has grown to dislike Obvious, seeing them as marketers that poorly represented the AI art community.



French artists that make up Obvious art collective: Hugo Caselles-Dupré, Pierre Fautrel and Gauthier Vernier



"Portrait of Edmond Belamy"

Before Obvious was a well-known collective, they got attention because they used exaggerated marketing techniques. The group implied that the art they were making was autonomously created by AI. That is simply not the case.

The generation of their art was done with a machine learning model called a GANs, a generative adversarial network. The model was trained on a corpus of 15,000 images of portraits made between the 14th and 20th centuries. The model then generates brand new images based on the training set. Those generated images are then run through a discriminator which determines if the image was made by the generator or the dataset. The cycle typically ends when the computer can not decipher between human-generated and computer-generated images.

The GANs model is a tool, it is not an autonomous artist. The user uses their artistic license to decide what the model is trained on, how many epochs to run the program for, and what output image to use.



Using image generation, the artist Mario Klingemann generated this piece entitled "Memories of a Passerby 1" --Click the image to hear Klingemann speak about this piece

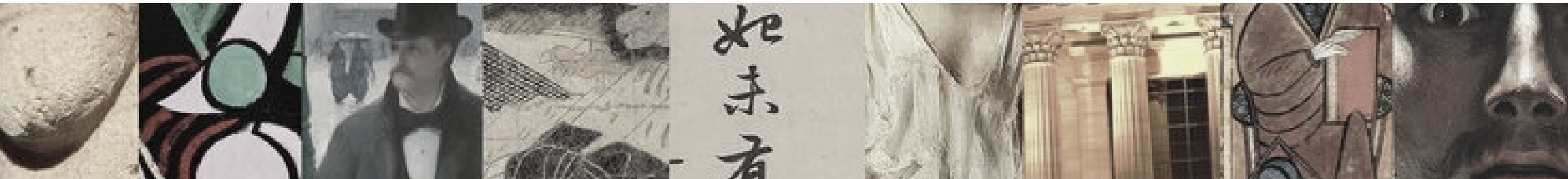
[The Arguments  
Against Obvious](#)



# Art Analysis



In addition to art creation, art analysis and art history are heavily impacted by AI. Most major art institutions have teams dedicated to using machine learning technology to forward curatorial, research, and community activities. Using art collections as a corpus of data allows for different techniques to be implemented. Different methods of data exploration allows us to organize, analyze, find patterns and create recommendations between different pieces of art. Artificial intelligence, augmented reality and virtual reality are becoming heavily used to further our understanding of art in the field of digital art history.





# Stepping Back....



Until my first year at Kenyon College, I had always been skeptical of technology. I felt protective of my favorite traditional art forms such as painting, intaglio printmaking, and darkroom photography. I felt that digital art was working to erase the beauty behind human creativity.

Learning about AI this past year has completely changed how I understand technology as a whole. I began to see AI as an extension of our humanity, instead of a sterile series of 0 and 1s. I urge you to combat the humancentric bias that most of us have against AI.

In learning about AI, separately from art, I grew a passion for understanding the human experience through this lens. Artificial intelligence is a tool similar to art, one which we use to further explore what it means to be human.

The AI world is here and it is only growing. We are fortunate to be at the forefront of this AI movement. We have the privilege of seeing and understanding what AI is before it has dissolved and disappeared into our world.

One of the things that helped me understand how mind-blowing and fascinating this field is was learning its history.



Above is some of my art

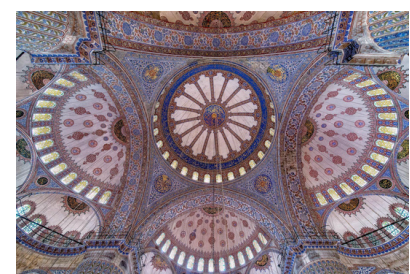
# History



Generative art has been around since the beginning of art itself. The use of generative ideas dates all the way back to the Middle Paleolithic era. Temples from the 1400s display algorithmic art, as the artist gives up control to follow a pattern. With the advent of computers, generative and algorithmic art only enhanced.

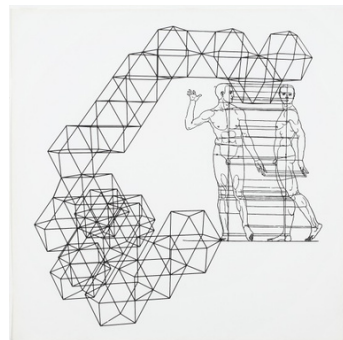
1400

Patterns



1960

COMPUTER ART



2014

GANS



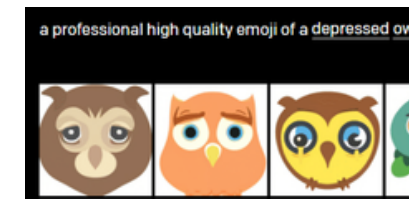
2015

DEEP DREAM



2021

DALLE



PRESENT

VQ-GAN CLIP



"The Yellow Smoke That Rises Its Mouth On The Window-Panels" from VQ-GAN-CLIP [@vqganclip](#) on Twitter





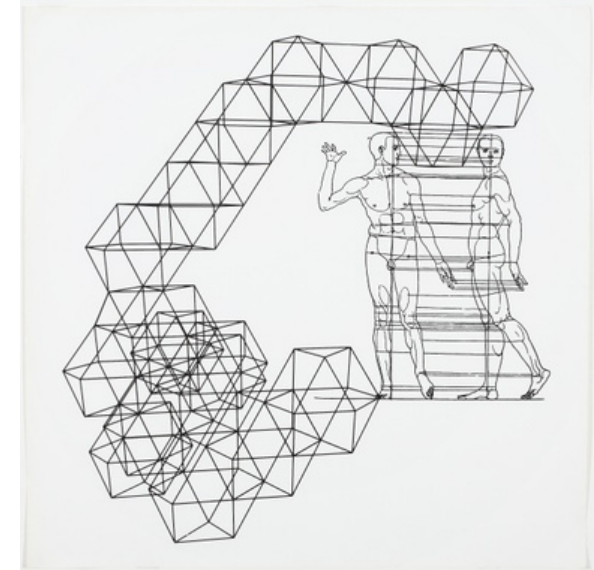
# Computer Art



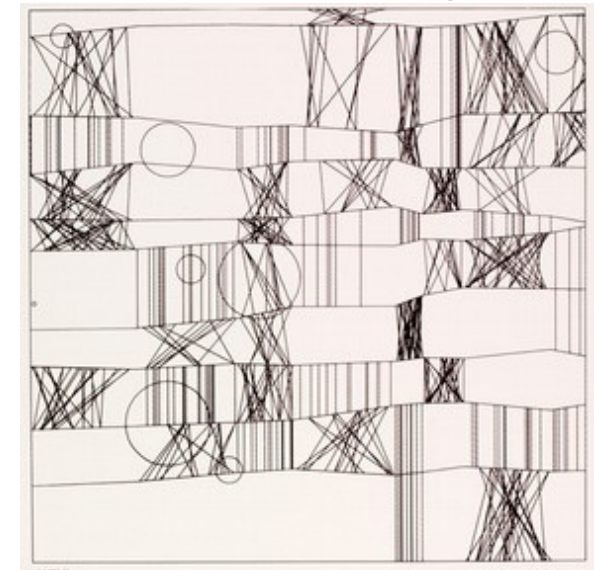
From the beginning, computers have been used as a tool to create art. In the 1960s, the commercial computer was very large and expensive making them difficult to access. Most computers were only available to large labs, universities, and corporations. This is why so many of the first computer artists were researchers.

With no user interface, the computer scientists and artists would write their own programs, the output devices were limited. Most used pen plotters. This was a mechanical device that held a pen and is connected to the computer that controls its movements.

The famous Bell Labs was at the forefront of this innovation. Artists, such as Frieder Nake and Georg Nees, would create a set of rules for the computer to follow and the plotter would output the work.



"Kubo-Oktaeder" Georg Nees

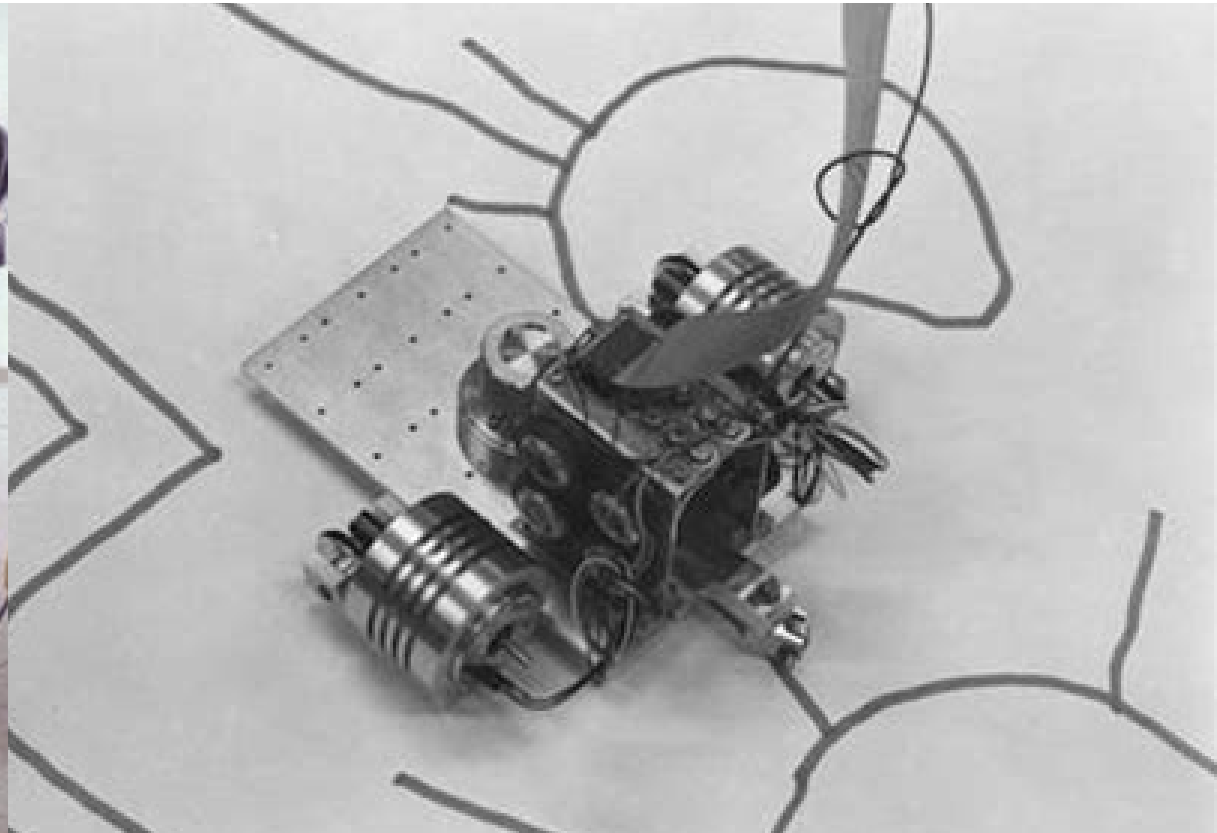
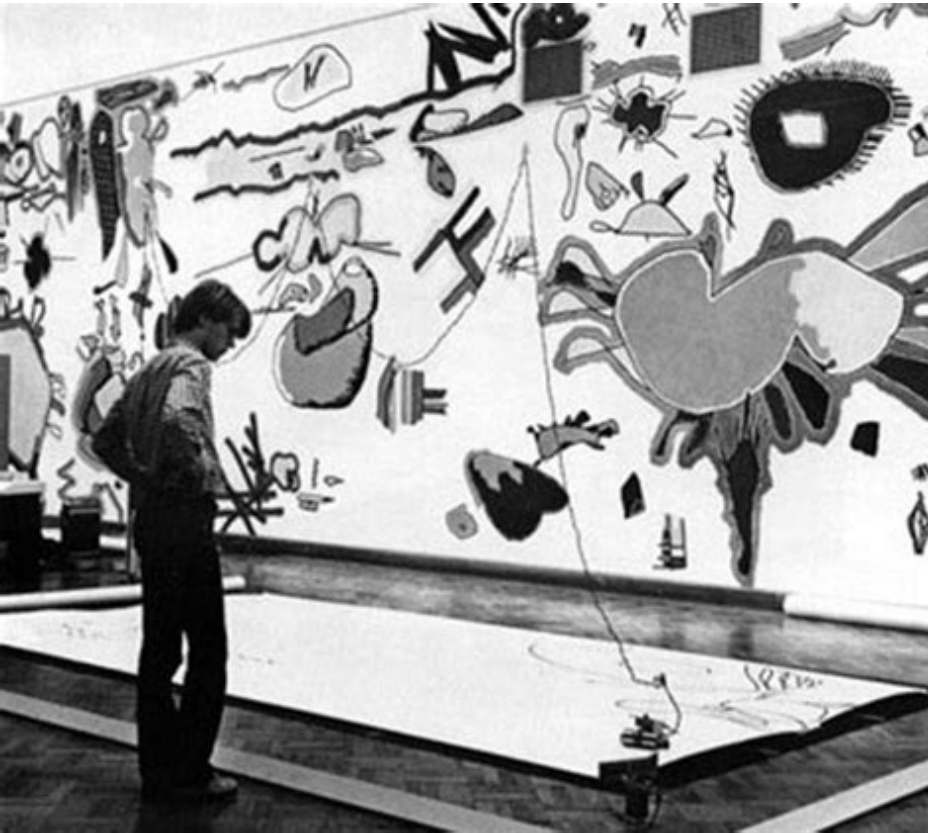


"Hommage à Paul Klee" Frieder Nake

# AARON



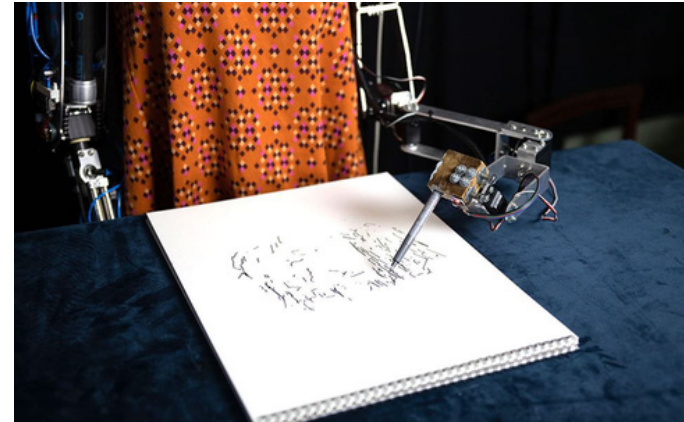
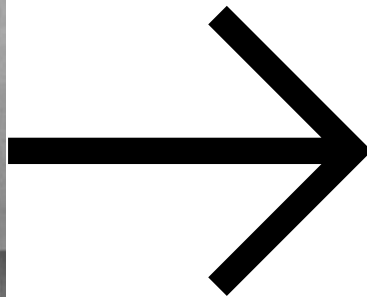
Following the work in Bell Labs, technology-backed art began to enter the fine art world. A foundational project was done in 1973 when Harold Cohen created AARON. AARON was one of the first programs made specifically for computer-generated art. This project was the first art-generated project that was designed by an artist instead of a scientist. Like the work of Nike and Nee, AARON would produce drawings that followed a predefined set of rules.



# Deterministic to Probabilistic



The computer art shown above is all deterministic and rule-based. Using pen plotters and algorithms, the artists would collaborate with computers to create rules that would generate art. The introduction of randomness in algorithms allows for there to be more uncertainty in the art-making process. These core ideas of probability, chance, and uncertainty had been explored before the use of computers by artists like Pollock. Computer art allowed the artist to make complex algorithms and insert randomness into their generative art.



With advancements in technology, we have traveled away from rule-based generation and moved towards more exploratory, probabilistic models. Because of increased accessibility to data we are able to use probabilistic analysis on data and generate art from it. This creates a huge foundation for conceptual art. Artists now have the ability to be surprised by the generated art.





Just in the past 8 years, the tools used in art generation have rapidly increased. It is expected that the advancements and accessibility of these tools will only improve faster and faster.

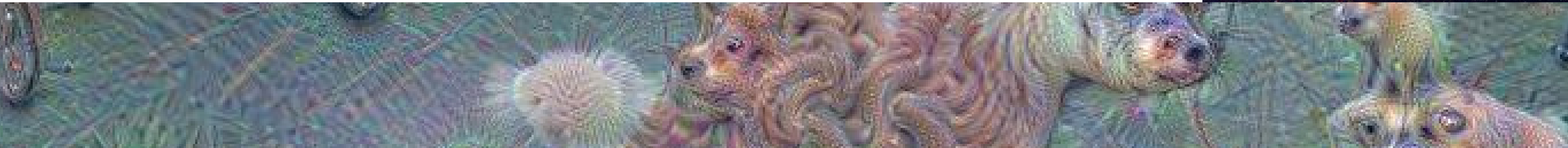
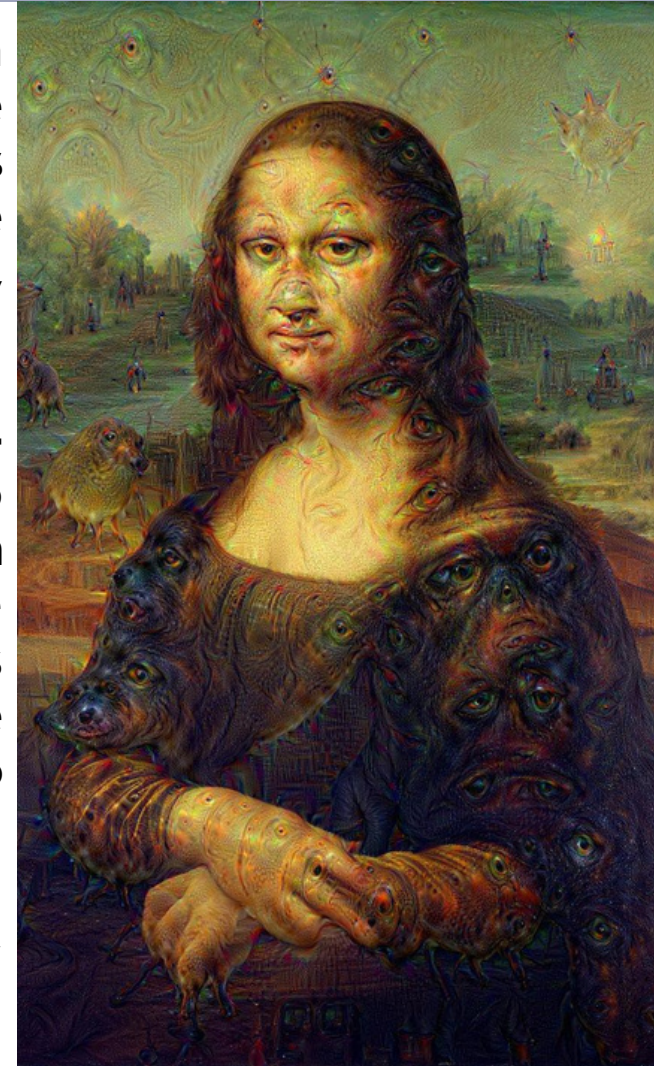
# Deep Dream



In 2015, Alex Mordvintsev and Google's Brain AI research team created Deep Dream. Initially, Deep Dream was made to advance the interpretability of deep convolutional neural networks. They did this by creating a visualization of the patterns that maximize the activation of their neurons. Because of the unique images created, they ended up accidentally creating a new AI art model.

Deep Dream is a neural network that was trained on images of real-world objects. These objects were sourced from ImageNet. Deep Dream learns to identify objects from their visual cues. Then when given a new image, Deep Dream works to find objects within the image, it updates the image each time, making the objects it finds within the image visible. As this cycle repeats over and over, the result ends up looking very trippy. From the images created by Deep Dream, a new genre of art was created: Inceptionism.

Deep Dream is a fundamental part of the timeline for AI Art, however, the style of Deep Dream gets repetitive quite quickly.





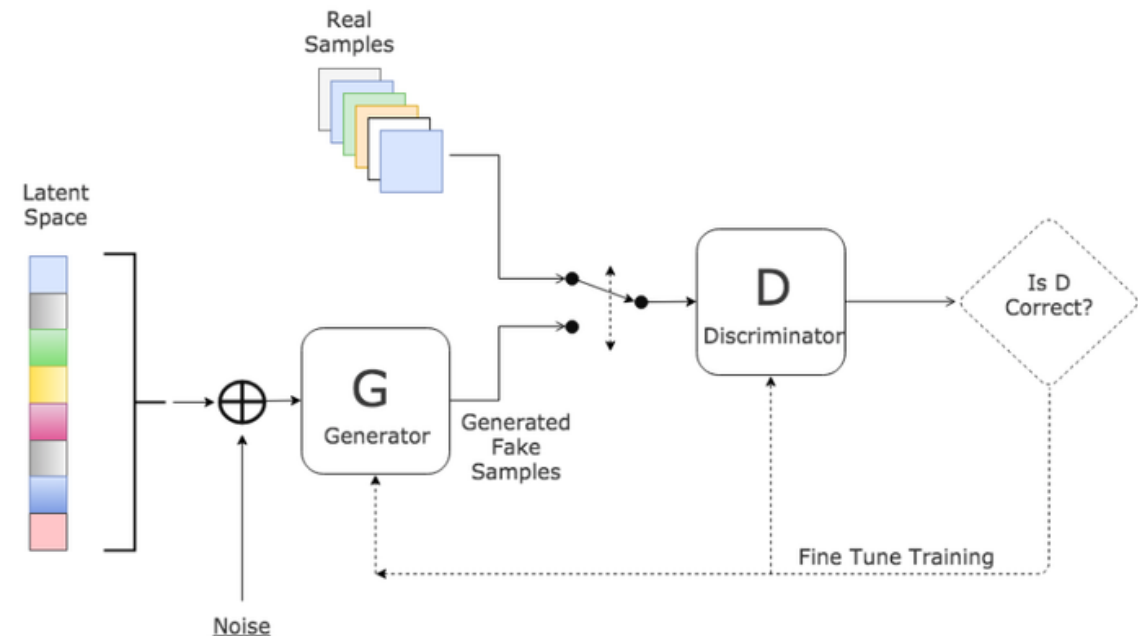
# GAN

The most talked-about tool for art generation is GAN, Generative Adversarial Network. Created by Ian Goodfellow in 2014, GAN is a neural network-based model that is used for image generation. In the model, there is a generator and a discriminator. The generator creates images and the discriminator goes into those images and decides if they are real or fake. It goes through this iterative process until the programmer tells it to stop or until the discriminator can no longer differentiate between the computer-generated images and the training set.

This model is what made "The Portrait of Edmond Belamy", the piece that sold at Christie's auction house. GAN holds many people's fascination because of the variety of results it can output. The generated images are dependent on the data that it is trained on and the parameters of the model.

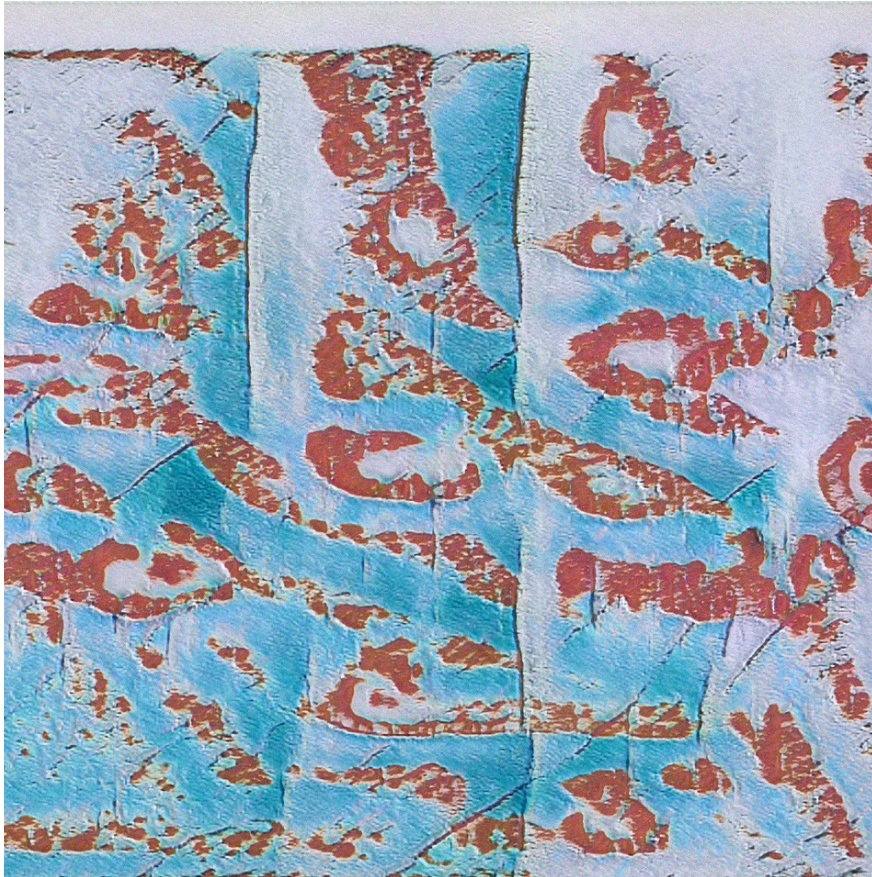


Generative Adversarial Network





**While training a GAN is quite a power-intensive task, it is relatively easy to generate images from a pre-trained model.**



**I used a GAN model that was trained on modern art images from WikiArt to generate the three images above. After generating these images, I began to ask questions about the ownership of this work. Who should get credit for the artificially generated art?**

More  
information  
about  
ownership



GANs are not just used to generate art-related images. A popular example is from a project called: This Person Does Not Exist. Using a GAN trained on images of real people, the model produces these never-before-seen, nonexistent faces.



2014



2015



2016



2017



2018

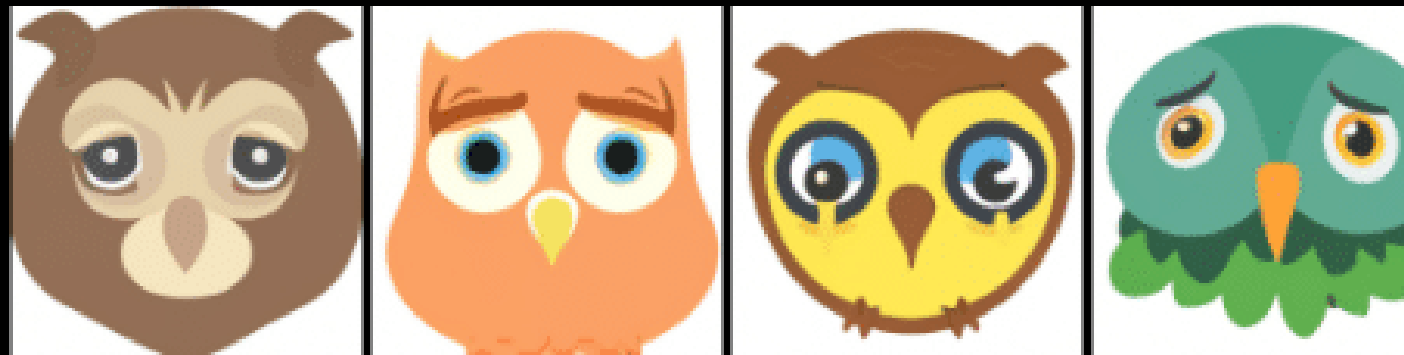


# DALL·E



GAN is an impressive tool, but, it pales in comparison to the most recent technology, Dall·E. Dall·E is a text-to-image model that was released by OpenAI in January 2021. The name is a combination of the legendary artist Salvador Dalí and the iconic Pixar animation robot Wall·E. The reason that this model is truly revolutionary is because of its ability to express creativity. Natural language text is input into the model and then it outputs an image translation.

a professional high quality emoji of a depressed owl



an armchair in the shape of an avocado. an armchair imitating an avocado.





# DALL·E



Dall·E not only understands objects but also expresses a temporal, geographical, and emotional understanding of the world. Dall·E is built upon a transformer model. Due to the expense of running the model and the power of what the model can do, OpenAI has not released its code.

To the right are the images generated by the prompt "a phone from the..."



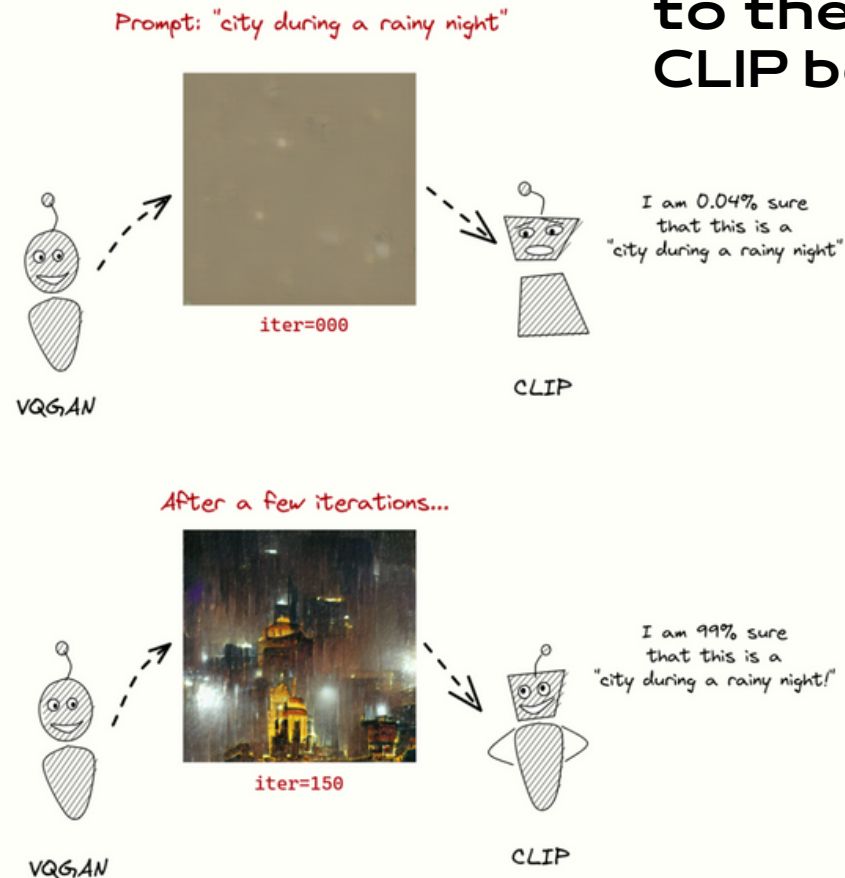
Above are the images generated simply by the prompt "a photo of San Fransisco's golden gate bridge"



[More information about Dall·E](#)

# VQGAN+CLIP and Clip Diffusion

Because OpenAI has not released Dall-E's code, many programmers have taken it upon themselves to create powerful text-to-image generation models. They did this by using another Open AI tool, CLIP. The two newest models are VQGAN CLIP and CLIP Diffusion. VQGAN + CLIP is built upon a GAN architecture. Basically, VQGAN generates an image and CLIP matches the image to the text prompt. After many iterations of this cycle, CLIP becomes confident in the text-to-image pairing.

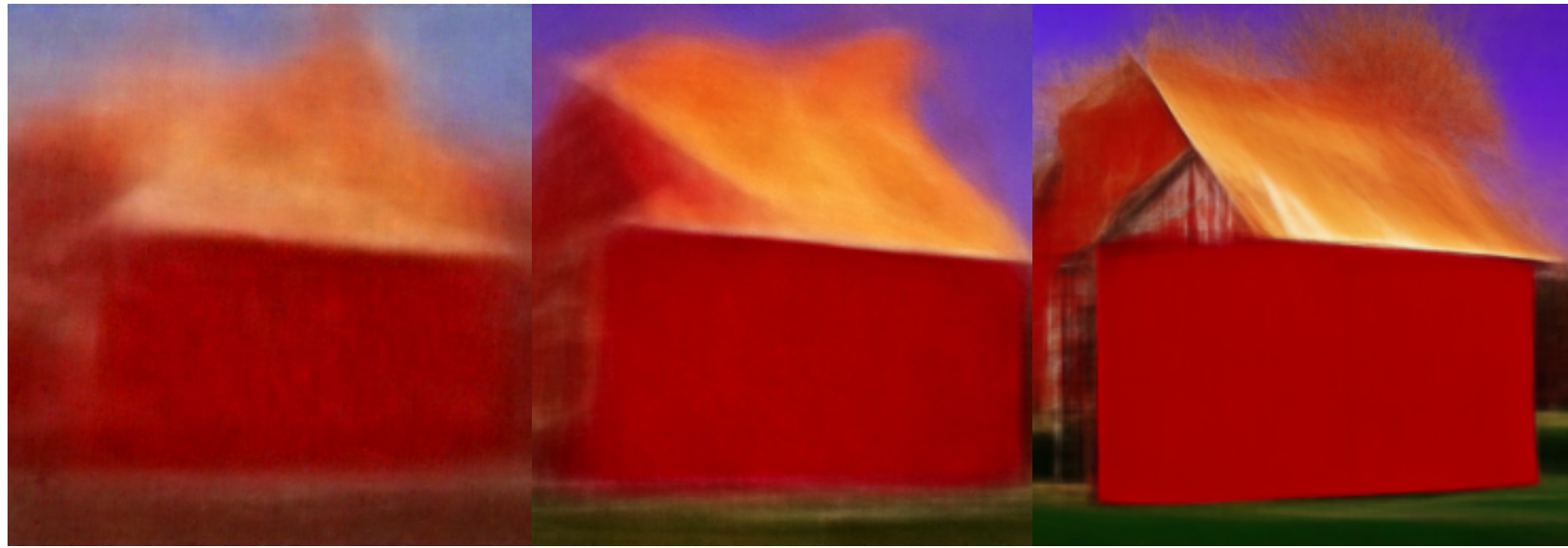


From the prompt "Calm" VQGAN CLIP generated the image above



From the prompt "'a small hut in a blizzard near the top of a mountain with one light turned on at dusk trending on artstation | unreal engine" VQGAN CLIP generated the image above

Clip Diffusion is built on a diffusion architecture. Generating images from these pre-trained models takes about an hour per text prompt. The three images below were generated using a CLIP Diffusion model. The images from left to right show the stages of development that the image went through in the generation process. The prompt I used was "A painting of a red barn on fire in the style of David Hockney." I was surprised to see the beautiful array of colors and the similar Hockney geometric style.



Within the last year, text-to-image generation has become increasingly popular in the AI Art community. Within the past six months, there have been a series of different no-code AI art generation programs like Dream by Wombo, Night Cafe, and StarryAI. These sites allow people to generate entirely new images with a click of a button. Many of these sites have direct ways for people to sell their digital art as NFTs. It is great that tools like these are becoming accessible to the public, but are the people that input the text into the apps really the artists?



I explored these themes of ownership in an art project recently. Using a pre-trained GANs model, I generated the image on the left. Then using oil paints I reproduced the image by hand. At the beginning of the art process and the presentation to my class, it felt unnatural to call this my art. I didn't decide the colors or composition.



However, because I created the concept behind the work, produced it, and followed through with the idea, I do feel some ownership over this work.

This piece reminds me of the art being made by the computer artists at Bell Labs in the 1960s. Those artists were giving the computer instructions and the computer would output a physical image. In this case, about sixty years later, the computer is giving me the "instructions" and I am producing the output. I cannot begin to imagine how much technology will change in another sixty years.

# Resources for getting involved



## Create.

- Runway ML
- ArtBreeder
- Magenta
- Processing
- ML5.js
- P5.js
- Night Café
- Dream WOMBO
- StarryAI

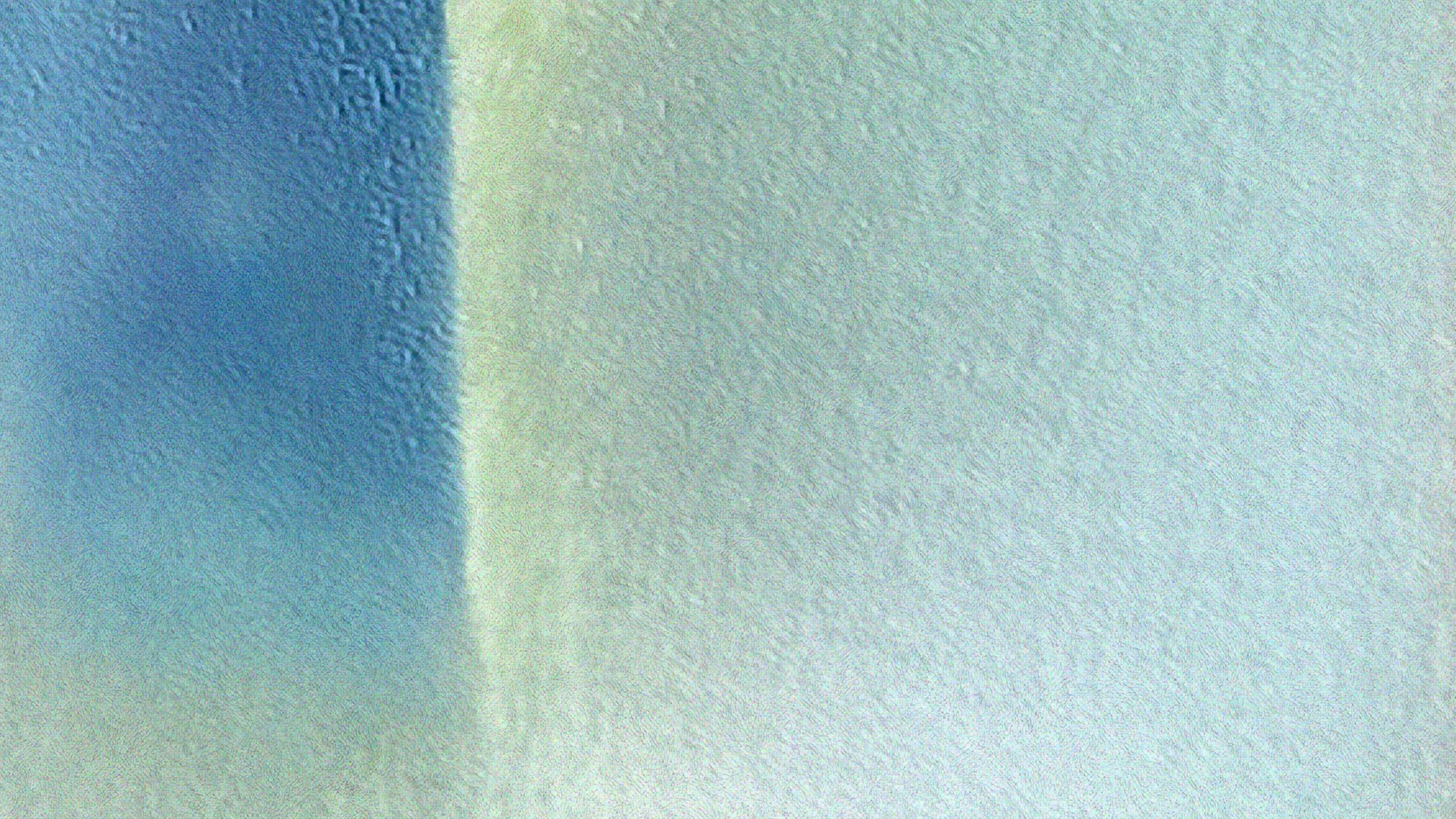
## Educate.

- Coding Train
- Nerdy Rodent
- Artnome
- The Gradient
- AIArtists

## Appreciate.

Lynn Hershman Leeson  
Anders Hoff  
Mark J. Stock  
Manolo Gamboa Naon  
INSA  
Joel Simons  
Alexander Reben  
Katia Vega  
Mario Klingemann  
Robbie Barrat  
Ahmed Elgammal  
Georg Muncey  
Obvious Collective  
Uncanny  
Stephanie Dinkins  
Joy Buolamwini  
Sofia Crespo  
Mimi Onuoha  
Memo Akten  
Tom White  
Mike Tyka  
Sarah Meyohas  
Jake Elwes  
David Young  
Dariusz Gross  
Refik Anadol  
Christian "Mio" Loclair  
Daniel Ambrosi  
Scott Eaton  
Anna Ridler  
Katherine Crowson







# More on..."The Portrait of Belamy"

## **Exaggerated the Advancement of AI**

The press, upon seeing Obvious's initial statement, spread fear-mongering articles that implied that art was being taken over by machines. Obvious now regret their initial exaggerated statement that read, "creativity isn't only for humans".

Claiming that the human is not in the loop of AI art generation not only spread inaccurate information but also discredited the work that many AI artists put into their creations.

## **Unoriginal**

This was the first moment that AI-generated art was taken seriously in the fine art world. Many AI artists were disappointed that this piece was what was being highlighted because they saw it as technically rudimentary. Also, the concept behind the work was uncomfortably similar to the project done by Robbie Barrat, in which he trained a GAN on historical portraits. The code Obvious used was directly borrowed from Barrat.

Obvious art collective has acknowledged their lack of originality and their lack of innovation. It is a shame that this piece has become the most iconic AI art piece because there are so many other fine artists making more compelling work. This can exist as an example of how marketing and the presentation of art are sometimes just as impactful as the art itself.

The inner workings of AI art are foreign to many people. What made "Portrait of Edmond Belamy" so successful was its accessibility.

One thing that is very clear is that this piece made big waves in the art scene, amplifying conversations about the intersections of technology and art. The misunderstandings that surround this piece of art perfectly parallel the confusion that we are currently at with the integration of AI into society as a whole. Because of this massive sale, there are more conversations happening about ownership, metrics, and creativity in this new era of art.

ImageNet: In 2009 Fei-Fei Li realized that the improvements of algorithms were based on the data it was fed. In order to ensure that the data reflected the real world, she decided to map out the entire world of objects.





# OpenAI's Dall-E and The Future of Text-to-Image Generation

Jill Noorily with Professor Chun and Professor Elkins

IPHS 300 AI for the Humanities: A Multidisciplinary Approach to Artificial Intelligence



## Abstract

The release of OpenAI's text-to-image generation model, DALL-E, has brought image generation to a new era of sophistication. OpenAI is a research and deployment company with a mission of ensuring that AGI (artificial general intelligence) benefits all of humanity. Founded in 2015, OpenAI has been a leader in producing powerful transformers such as GPT-2 and GPT-3. In January 2021, OpenAI released a blog post and a controlled interactive pilot of Dall-E: a 12-billion-parameter neural network built upon the 175 billion parameter transformer, GPT-3. Trained on text and pixel tokens, Dall-E's capabilities range from generating illustrations and combining unrelated concepts to applying a transformation to images and visualizing perspective. With this comes new possibilities for design, architecture, and art. Because OpenAI has not released the code yet, many independent programmers have pieced together their own open-sourced copy of DALL-E. Being overly impressed by the results, I wanted to explore how practical and usable the image-to-text capabilities are to the artist, creative and average person. To do this I ran three of the best open-source DALL-E models, compared their results, and gathered insight on DALL-E's future implications for art and design.

## Introduction

Image generation models have been rapidly growing in sophistication since 2014 when GANs (generative adversarial networks) were introduced. GANs are made up of a generator network, which creates training data, and a discriminator network, which detects and classifies the data as real or fake. By pairing both of these models against each other they become adversaries. A healthy competition between them grows. GANs was a huge advancement in the deep learning field because of their ability to "generating 'data' when you don't have enough to start with, and 'training' models with no manual intervention."

The development of GANs enabled the work done in text-to-image generation. GANs were used in many of the early text-to-image models, however, with the introduction of transformers in 2017, GANs have become less useful. Using a 12 billion parameter version of GPT3, OpenAI's Dall-E has advanced text-to-image generation forward with its transformer architecture and natural language processing.



## Materials

Google Colaboratory, Python, OpenAI DALL-E Blogpost, Big Sleep, Deep Daze, Aleph2Image.

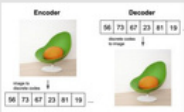
## How Dall-E Works

Although OpenAI has not released DALL-E's code, they did release a research paper in February 2021 that outlined DALL-E's zero-shot text-to-image generation process. Dall-E can be split into two parts: the discrete variational auto-encoder (dVAE), which compresses images, and the transformer, which learns the relationship between text and image.

First, the input data must be reorganized so it can be understood by the transformer. To do this the text tokens must be organized into a single stream of data and the image's pixel must be tokenized by the dVAE. Using the pixels directly from the input training set would take way too much memory. The dVAE learns discrete representation for the images, reducing the memory footprint.



In the figure shown above, the top image is the original input image and the bottom image is the image after it has been encoded and decoded. The dVAE process does not have a large impact on the visual quality of the image.



The image shown above maps the process of the dVAE encoding and decoding an image.

Then, the dVAE image tokens and the single stream of text are sent as input into the transformer model. The transformer is the most important part of DALL-E's architecture. It allows the model to generate images that fit the text prompt. The transformer has an understanding of basic physics, accurate and creative designs, and global knowledge (geography/time). Because the transformer models in the latent space it focuses more on the big picture of the image, letting the dVAE render the small details after.

DALL-E's transformer is incredibly complex and because of that fairly unexplainable. Currently, there is no solid theory for deep learning that can fully explain these massive networks

Then it is sent to CLIP (Contrastive Language-Image Pre-training), a pre-trained contrastive model that is used to optimize the pairing between the image and text. CLIP is a neural network produced by OpenAI in January 2021. It was trained on 400 million text-image pairs and has zero-shot capabilities. The image and caption are fed into CLIP. CLIP then gives a score for how well the image matches the caption. Then it reranks the top 32 of the 512 sample images for each caption.

## Methodology

To compare DALL-E to the available text-to-image models, I ran three open-sourced Google-Colab notebooks: Deep Daze, Big Sleep, and Aleph2Image: CLIP DALL-E. In OpenAI's blog, they fed unique text prompts, such as "avocado chair", into DALL-E to prove that the model wasn't overfitting. I decided to do the same with the open-sourced models and fed them the prompt, "dog driving a car". Each model ran for around 5 hours with 1050 iterations. The results were elementary in comparison to that of Dall-E; however, this was understandable considering it was free and relatively fast to run these models.

## Results

Out of the three open-sourced models, I found Deep Daze to be most successful. As seen below Deep Daze most accurately depicts a dog and steering wheel. Also, the color used in the background closely resembles that of a road.

Model	Results	Prompt	Architecture	Cost to Train	Time to Train
Dall-E		"An illustration of a baby fox in a helmet riding a car"	dVAE transformer and CLIP	Millions	Weeks to Months
Big Sleep		"Dog driving a car"	OpenAI's CLIP and Sites (Implicit Neural Representation Network)	Free	5 hours
Deep Daze		"Dog driving a car"	OpenAI's CLIP and the generator from a BigGAN	Free	6 hours
Aleph2Image CLIP DALL-E		"Dog driving a car"	OpenAI's CLIP and DALL-E's dVAE	Free	5 hours

Both Big Sleep and Aleph2Image were highly abstracted, leaving the result to be uninterpretable. While all three models have different architectures all of them used OpenAI's CLIP.

In Deep Daze the image of a dog's face is decipherable in the center and part of a steering wheel can be seen above the dog. Even though Deep Daze is still quite abstract, I am confident that with more development it has the most potential. I was curious to see how creative and diverse Deep Daze could be. To test this I ran the model once more with the same text prompt and the result was quite repetitive:



As shown on right, the image has a similar color palette and composition to the first result. There is one main abstraction (possibly depicting a car) and various other shapes surrounding it. This leads me to believe that Deep Daze is not very creative.

I did find it interesting how in this image, the model incorporated the input text. Above the large white "car" one can decipher letters spelling "driving car".

It becomes clear that accessible text-to-model generation is not at a level that would be a useful tool.

## Conclusion

OpenAI's Dall-E was simultaneously met with amazement and anxiety. The amazement came from people that saw DALL-E as a tool that could be used in design, e-commerce, and art fields. The anxiety came from those who saw this as the "end of human creativity". Exploring DALL-E in conjunction with the accessible text-to-image models reveal the gap between plausible and practical text-to-image generation. DALL-E has shown the world that a very high level of creativity from AI is possible. The exploration of the open-sourced models has shown that there is still much work to be done for text-to-image generation. The hype of Dall-E is well deserved; however, the fears of text-to-image programs fully taking over the jobs of designers and artists is not realistic for the very near future.

DALL-E loses confidence with poorly worded prompts and frequently has trouble positioning objects relative to one another. The biggest limitation of DALL-E is the fact that all its creative attempts are restricted to the data sets it is trained on. Even as the data sets increase in size they will still be a limiting factor in the creativity of the model. However, one could argue that the same limitation could be applied to human creativity. Much human creativity comes from inspiration from other humans/nature, it is rare that an idea is fully original.

Despite the importance of transparency in AI, I do find comfort in OpenAI's strict regulation of DALL-E. They claim to be analyzing the societal impact and potential biases in models like DALL-E before they release the source code. Companies like OpenAI should continue to take measures to ensure their social impacts are ethical and responsible.

DALL-E has proven itself as a promising technology for the future and I am confident that with further advancements we will be seeing much more impressive and accessible text-to-image generation. The race for a creative AI is happening and DALL-E is definitely at the forefront.

## Acknowledgements

A. Ramesh et al., "Zero-shot text-to-image generation, 2021, arXiv:2102.12092 [cs.CV]"  
"How is it so Good ? (DALL-E Explained Pt. 2)." *How is it so Good ? (DALL-E Explained Pt. 2)* - ML@B Blog, ml.berkeley.edu/blog/posts/dalle2/  
Macaulay, Thomas. "Say Hello to OpenAI's DALL-E, a GPT-3-Powered Bot That Creates 'Weird Images from Text.'" *TWJ | Neural*, 27 Apr. 2021, thenextweb.com/heres/say-hello-to-openais-dall-e-a-gpt-3-powered-bot-that-creates-weird-images-from-text.  
OpenAI. "About OpenAI." OpenAI, OpenAI, 2 Sept. 2020, openai.com/about/.  
OpenAI. "DALL-E: Creating Images from Text." OpenAI, OpenAI, 27 Feb. 2021, openai.com/blog/dall-e/.  
"OpenAI's DALL-E: Text-to-Image Generation Explained [With Code Available!]." YouTube, YouTube, 25 Feb. 2021, www.youtube.com/watch?v=DJTODLBPovg&t=97s.  
Pollock, Griselda. *Vision and Difference: Feminism, Femininity and the Histories of Art*. Routledge, 2003.  
Ryan Murdock, Phil Wang. (2021) Deep Daze (version 1) <https://github.com/lucidrains/deep-daze>  
Ryan Murdock, Phil Wang (2021) Big Sleep (version 1) <https://github.com/lucidrains/big-sleep>  
Ryan Murdock, (2021) Aleph2Image (version 1) <https://github.com/lucidrains/aleph2-image>  
Source, "Here's How OpenAI's Magical DALL-E Image Generator Works." *PressbotNews*, 10 Jan. 2021, pressbotnews.com/heres-how-openais-magical-dall-e-image-generator-works/36570/.  
Team —, et al. "What Does DALL-E Mean for the Future of Design?" *The Hard Copy*, 6 Mar. 2021, thehardcopy.co/what-does-dall-e-mean-for-the-future-of-design/.  
Verma, Aishwarya. "Comprehensive Guide to DALL-E By OpenAI: Creating Images from Text -". *Analytics India Magazine*, 2 Mar. 2021, analyticsindiamag.com/comprehensive-guide-to-dall-e-by-openai-creating-images-from-text/.

RETURN



# Who Gets Credit for AI-Generated Art



Currently, the authorship rights are attributed to the artist who produced the artwork using AI techniques. AI technology is a tool for the artist just like the piano is a tool for a musician.

A study that measured how the public intuitively perceived and assigned credit to an AI involved in making art concluded that the allocation of responsibility for individuals that create AI Art heavily depends on the language framing it. When an AI is considered the agent(creator) of the art, they assign more value to the technologist that wrote the code instead of the artist, compared to when the AI is described as a non-agent.

Increasing the anthropomorphic language surrounding AI systems decreases responsibility towards the artist. As we move into a world where AI systems are becoming more anthropomorphized, it is important to emphasize that AI is a tool, not a being.

# What is CLIP?

CLIP (Contrastive Language-Image Pre-training), is a pre-trained contrastive model that is used to optimize the pairing between an image and text. CLIP is a neural network produced by OpenAI in January 2021. It was trained on 400 million text-image pairs and has zero-shot capabilities. The image and caption are fed into CLIP. CLIP then gives a score for how well the image matches the caption.

CLIP is used in most text-to-image generation models



[Return](#)