

Created in DALL·E 2 to prove a point. Prompt: "A red robot making a blue robot. Insanely high quality, insanely intricate detail, studio photo, masterpiece, sharp focus, 8k artistic photography, ominous matte painting, dramatic studio lighting, trending on cgsociety, vibrant, epic, intricate, trending on artstation, volumetric cinematic perfect lighting."



**STEFANIA DRUGA, PH.D.**

is working on the future of AI in education, including research in the Personal Robots Group at MIT Media Lab.



# GENERATIVE AI FOR MAKERS

Written by  
Stefania Druga

FROM 2D ART TO 3D OBJECT DESIGN TO TEXT,  
AUDIO, AND CODE, AI HAS TRULY ARRIVED —  
AND IT'S HERE TO HELP YOU MAKE AND CRAFT

Imagine the *Star Trek* replicator exists, and you can use it to create or replicate any object. What would you ask it to make? Would you print ready-made things, or prefer instead to get parts you can assemble, polish, and customize?

**Generative AI** is the closest approximation of the replicator we currently have — powerful artificial intelligence models that can synthesize images, text, video, code, and even 3D objects. Their results are primarily digital and still require quite a bit of shepherding to materialize in the physical world, but already it can be done — and we're seeing a proliferation of new tools and improvements seemingly every day.

How can we use generative AI models for tinkering and making? What are some initial forays in real, physical crafting with generative AI that can inspire future potential directions? This article sheds light on how generative AI works (in particular diffusion models), how we might ideate, design, and make with it, and what this all means for creators.

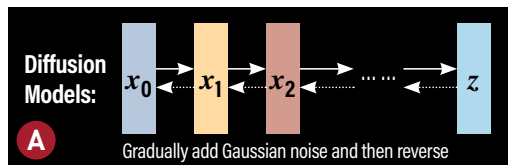
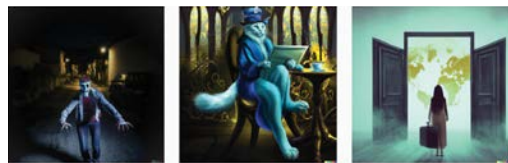
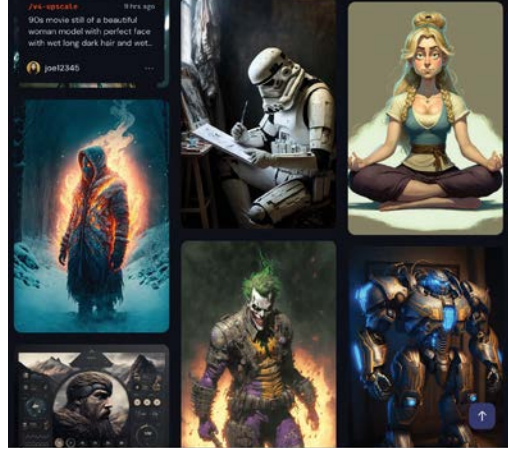
## The Current State of the Art

Recent applications for generative AI, such as DALL-E (1.5 million users) or Midjourney (4 million users), have taken the content creation world by storm and stimulated our collective imagination to consider AI a new medium for artistic expression.

Many of these applications use machine learning models that generate images based on a text description, also called a *prompt*. These large image-generation models are trained on an enormous amount of data, allowing users to create amazingly high-quality images with no graphics or design training. While many of you have seen examples of AI-generated images or videos, you may wonder how this technology works and why it has become so popular.

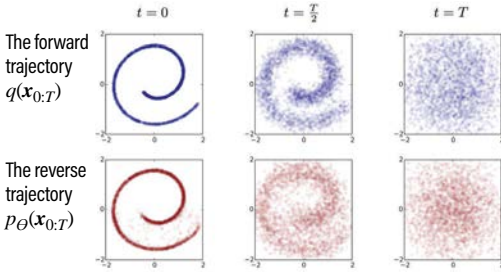
## HOW DO DIFFUSION MODELS WORK?

Many generative AI applications use *diffusion model* architecture under the hood (see Figure A). Diffusion models are a type of AI algorithm inspired by non-equilibrium thermodynamics. They add random noise to an input image and

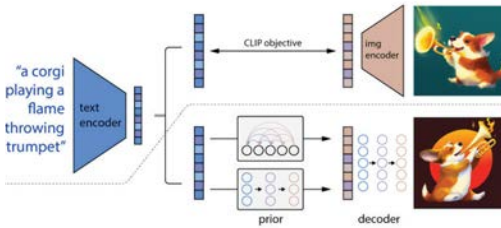


then learn to reconstruct a new, similar image from noise. As more noise is added to different samples of the original image ( $x_1$ ,  $x_2$ ), the image gets compressed into a low dimensional representation ( $z$ ) which is used to create a new image similar to the original one. The process of gradually adding noise is called a *forward*





**A** Training a diffusion model for modeling a 2D Swiss roll. From Sohl-Dickstein et al., 2015., [arxiv.org/abs/1503.03585v8](https://arxiv.org/abs/1503.03585v8).



**B** The architecture of unCLIP, from Ramesh et al., 2022, [arxiv.org/abs/2204.06125](https://arxiv.org/abs/2204.06125).



"A delicious steamed bun in the shape of a shiba inu. Studio lighting, high resolution, high quality, dark background"



"A very beautiful intricately shaped organic sculpture carved from steamed buns depicting a shiba inu. Studio lighting, high resolution, high quality, dark background"

**C** Examples of hybrid and unique artifacts created by Shai Noy and Daniel Smith using Google Imagen.



**D**



**E**

AI-generated design ideas for materials, fashion, and architecture from Shai Noy (D) and Oren Levantar (E).

*trajectory* or *forward pass*, and the process of reconstructing a new image progressively from noise is called a *reverse trajectory*.

The key insight is that a diffusion model needs to gradually learn the probability of the distribution of noise for different steps in the reverse trajectory (see  $p_{\theta}$  in Figure **A**).

Another way to think about this is to imagine that diffusion models work by destroying training data through the successive addition of noise, and then learning to recover the data by reversing this noising process. After training, we can use the diffusion model to generate unique new data by simply passing sampled noise through the learned denoising process.

To guide the reconstruction trajectory, more recent implementations of diffusion models use text, semantic maps, or other images to condition what possible image should be generated (reconstructed) from the space of all possible options with different probabilities, aka the *latent space* (see Figure **B**).

As mentioned above, diffusion models have exploded in popularity as they produce state-of-the-art image quality and enable people to create fantastic or photorealistic images that didn't exist before, such as hybrid creatures, intricate architectures, new materials, and unique artifacts (Figures **C**, **D**, and **E**).

## What Can I Make With Generative AI?

You can use generative AI to create images, text, music, games, avatars, UIs, videos, and even 3D models. Here are just a few platforms that have gained popularity:

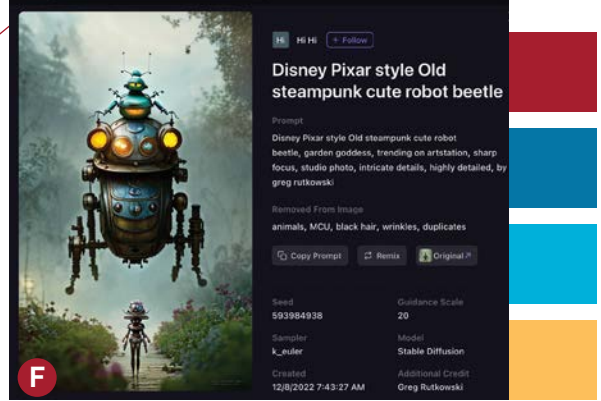
- **IMAGES:** DALL-E 2, Midjourney, Stability AI
- **TEXT:** GPT-3 Playground, Jasper, Google's AI Test Kitchen, Chat-GPT
- **VIDEO:** Meta's Make-A-Video, Google's Imagen Video
- **MUSIC:** Harmonai, Sony Flow Machines
- **AVATARS:** Character.AI, Lensa
- **USER INTERFACES (UI):** Figma plugins for Stable Diffusion
- **VIDEO GAMES:** NVIDIA's DLSS (Deep Learning Super Sampling)
- **VARIOUS DEMOS AND APPLICATIONS:** Hugging Face Spaces



Examples of popular Hugging Face Spaces ([huggingface.co/spaces](https://huggingface.co/spaces)), user-submitted generative AI apps of all kinds.

## Digital Crafting With Generative AI

Most generative AI models use text prompts as input, creating unique opportunities for creators and designers to iterate on their ideas quickly or to collaborate with others. As a result, large communities of practitioners have emerged around these technologies, with people sharing



An AI-generated image shared with the prompt that was used to generate it, referencing the styles of Pixar and artist Greg Rutkowski, on the Playground AI community. [playgroundai.com/post/clbf910ht00bcs601q5xf2u2d](https://playgroundai.com/post/clbf910ht00bcs601q5xf2u2d)

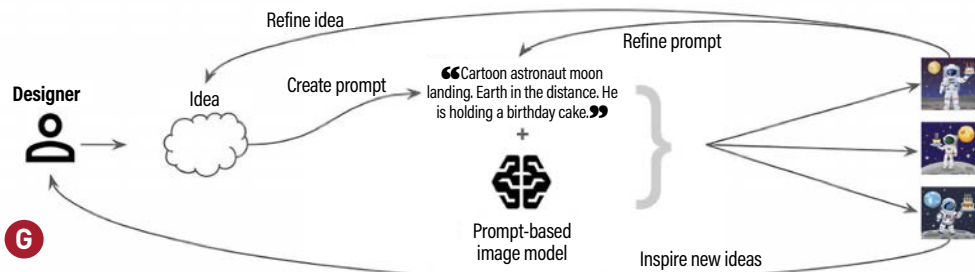
images, prompts, or tricks to achieve specific effects or styles. For example, Midjourney has more than 5 million users on their Discord, using the platform for fun and for professional projects.

## Prompts as a Craft Material

Figure F shows an example of a text prompt, shared by user “Hi Hi” on the Playground AI platform — “Disney Pixar style Old steampunk cute robot beetle, garden goddess, trending on artstation, sharp focus, studio photo, intricate details, highly detailed, by greg rutkowski” — alongside the image that it generated. Other creators can remix whichever prompts inspire them, hoping to achieve similar effects. There are even secondary markets, such as PromptBase, where creators sell their successful prompts.

In my research with the PAIR team at Google Research, I found that designers working in pairs to create specific artistic artifacts prefer using generative AI over working without it, and that they collaborate more effectively when using it (Figure G).

In our observations of designers' work, the



The process used by designers to generate specific design concepts with AI, from Kulkami, Druga, et al., 2023

indirect nature of prompting both supported the design process (by augmenting creative freedom) and made it more challenging (working to rephrase prompts to match their intent).

In some ways, prompts now occupy a similar role in visual design as HTML did in early web design. By seeing how a webpage was constructed, designers could rapidly adopt good ideas, remix them, and popularize them widely. The role of web browsers was also key — by making View Source a universal feature, browsers likely transformed millions of people from web “readers” to web “writers.” Sharing AI prompts alongside the generated artifacts could catalyze visual design in a similar way.

For example, platforms such as Playground AI support more straightforward iteration and remixing, by allowing users to share images with all the metadata required to reproduce them (prompt, model ID, etc.) (Figures **H** and **I**). These features are making prompt-based image generation even more accessible and more craftable. Moreover, many of these features for generative AI are becoming available directly in design tools such as Photoshop or Figma, enabling designers to integrate them into their workflow.

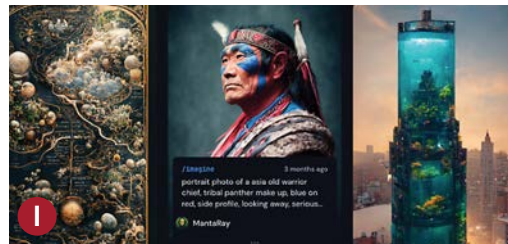
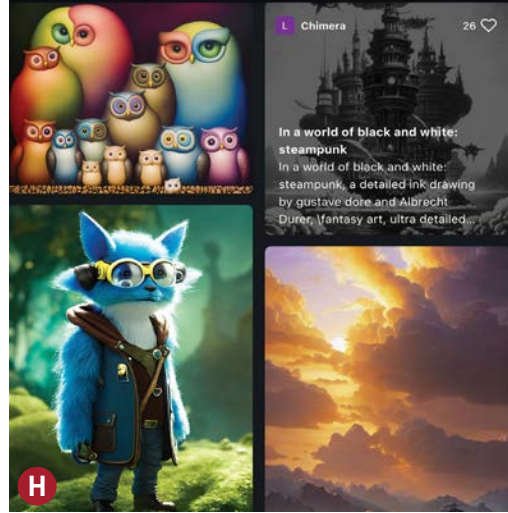
### PHYSICAL CRAFTING WITH GENERATIVE AI

In a variety of maker communities, generative AI is starting to be integrated into fabrication and crafting projects. These examples show that generative AI models are primarily used for two purposes: **ideation** or **generative design**.

#### Ideation

Many makers are already using generative AI for ideation. For example, they use Midjourney to generate concept boards starting with an object or a concept they like, such as *shell earrings* (Figure **J**), *Birds of Paradise fashion* (Figure **K**), or *Rambutan dress* (Figure **L**).

Then they select an intriguing initial composition and use AI models to generate many revised iterations based on the original image. With each one, the AI learns more about your end goal and sometimes suggests its quirky take on



Artifact and prompt sharing at [playgroundai.com](https://playgroundai.com) (H) and [midjourney.com/showcase](https://midjourney.com/showcase) (I).

### PROMPT ENGINEERING

Becoming a good “promptist” is key to getting good AI images. Writing effective prompts is a black art almost as mysterious as what’s going on inside the AI; many users compulsively include “greg rutkowksi” and “trending on artstation” in every prompt for reasons that seem unrelated, even talismanic.

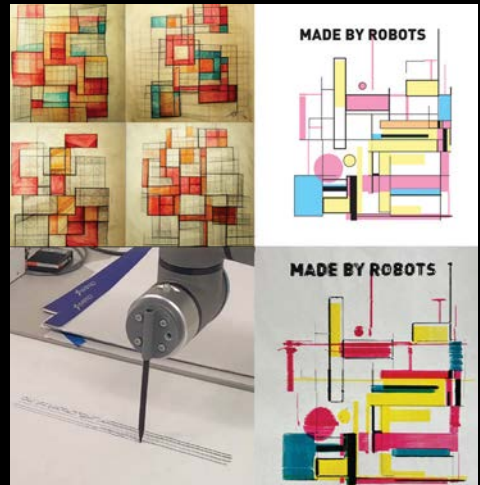
You can go a long way just by modifying existing prompts, but there are sites where you can learn the science of how they work, including your desired subject and style of course but also *negative prompts* (what you don’t want in the image), *seed number* (random by default, but reusing the same one lets you control your experiments), and *guidance scale* (how closely the image must adhere to the prompt).

Some sites even use AI to help you write prompts for the image AI! Check out Lexica, PromptoMania, Phraser, PromptHero, and Krea.ai, and learn more about how prompts work at [makezine.com/go/trending-on-artstation](https://makezine.com/go/trending-on-artstation). —Keith Hammond



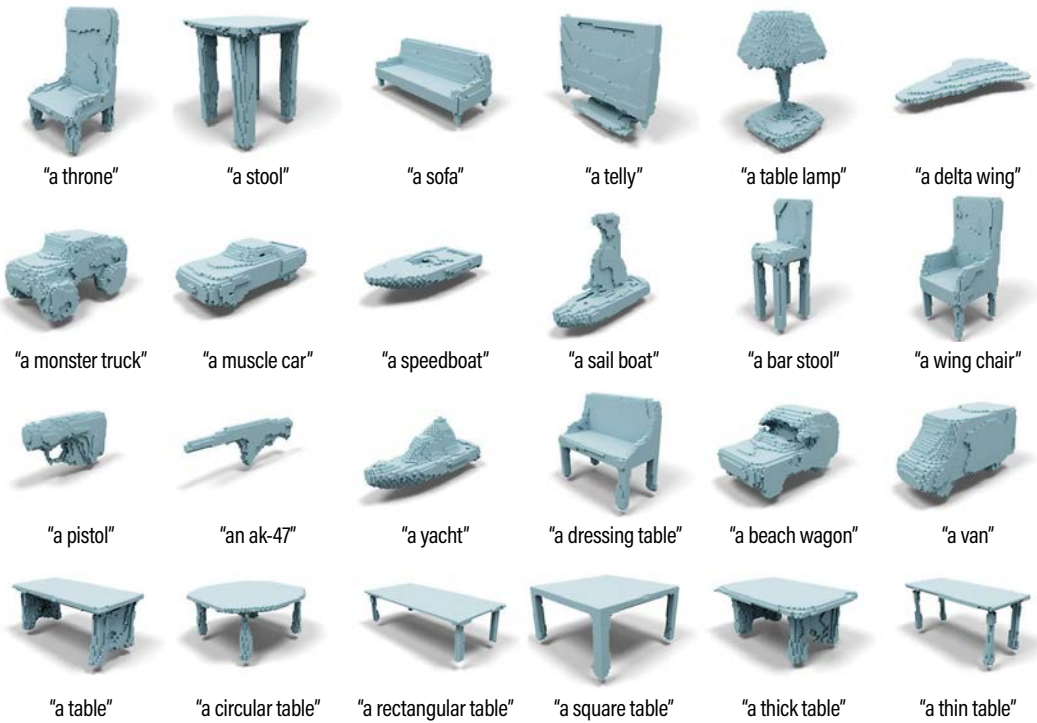
## Art by AI, Drawing by Robot

Design students at CSU-Long Beach are using diffusion models to generate art that a robot can draw. The tricky part is picking a drawing style that could be successfully painted by machines — in this case a **Universal Robots UR5E robotic arm** wielding Tombo brush pens. The CAD tool of choice for this operation is Rhino's Grasshopper, which generates a topographical model of the image input into this definition based on the light/dark values of the image; lighter areas will cause the robot to lift itself upward, away from the page. Adobe Creative Cloud's Illustrator and Photoshop tools are used to adjust the outcomes.



the initial prompt along the way. Makers can then use the Upscale and Remaster features of the AI several times to get a very polished composition before moving onward with their fabrication process. Once they achieve a design they like,

they either generate a 3D model in CAD tools or — amazingly — use the successful prompt to directly generate 3D renderings in CLIP-Forge or other text-to-3D diffusion models.



**M** 3D objects generated with Autodesk's CLIP-Forge, Sanghi et al., 2021, [arxiv.org/pdf/2110.02624.pdf](https://arxiv.org/pdf/2110.02624.pdf).

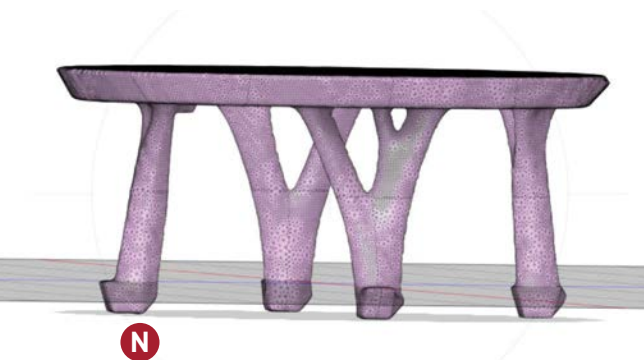
## GENERATIVE DESIGN

Makers also use generative AI when they want to explore quickly a design space or various form factors for the same object type. Suppose you want to build a table; you could use a **text-to-3D AI model** like Autodesk's CLIP-Forge to generate 3D models of various types of tables directly from a text prompt (Figure **M**).

Once you pick a table model you like, you could go further and use **generative design tools in CAD programs** to generate various design

options for the legs or the top, like this project done in Fusion 360 (Figure **N**).

Many of the text-to-3D rendering AI models can also export 3D meshes. The newest DreamFusion model ([dreamfusion3d.github.io](https://dreamfusion3d.github.io)) adds additional optimization strategies to improve geometry, allowing the final rendered models to have high-quality normals, surface geometry, and depth which could easily be exported to CAD for 3D printing (Figure **O**).



Autodesk/Sanghi et al., Jacqueline Mai, Poole et al., ArtStation, Spawning.ai



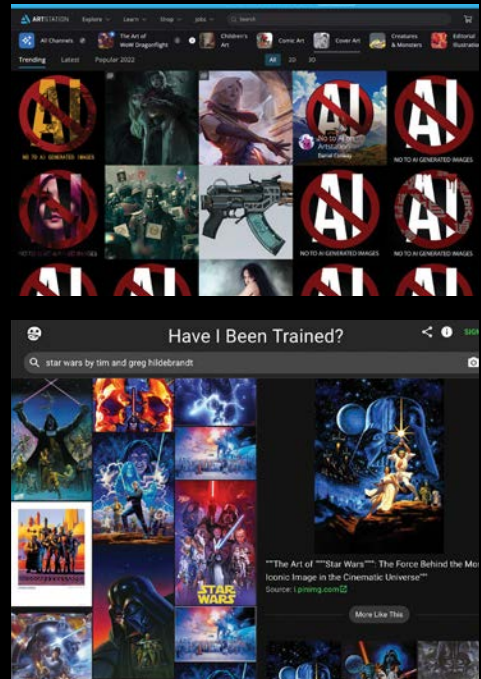
**O** DreamFusion, by AI researchers at Google and UC Berkeley, generates 3D printable meshes from a text prompt, Poole et al., 2022.

## WHAT DOES THIS MEAN FOR CREATORS?

While these generative AI models allow anyone to express themselves with images, videos, music, or 3D models, they've been received with mixed reactions in the creators' communities. When an image generated by AI won an art competition, the artist community reacted strongly against allowing such submissions.

Art historians argue that generative models like DALL-E do not themselves create art but that the artists and technologists who apply them as tools are the ones creating art. Art communities such as Getty Images/iStock/Unsplash, Newgrounds, PurplePort, and [reddit/r/DigitalPainting](#) have banned AI-generated art on their platforms. However, design firms such as Ideo confirm that they are currently using generative AI in their practice to generate more inclusive personas or unique concept boards.

I think the examples of imagery we see emerging in the existing communities, such as Midjourney, really call us to revisit the famous quote from Alan Kay, "The music is not in the piano," and maybe create alternative metaphors. Rather than thinking of these models as paintbrushes or musical instruments — or as robots replacing us — maybe we can think of them as an opinionated design partner that sometimes will inspire us to diverge our creative process in surprising and whimsical ways. 🗳️



## ARTISTS AGAINST AI

*How about we scrape all the images off the internet, then train a computer to copy the style of every artist and photographer, living or dead? What's the worst that could happen?*

A machine that steals your style is a new ethical, legal, and economic problem — especially if you're a working artist who's being mimicked. Do you get attribution? Get paid? Or just get ripped off and lose work to the machines? One thing seems clear: Existing copyright law can't keep up with AI art technology. And AI music and video can't be far behind.

Some artists are trying to opt out by explicitly denying companies permission to use their images for training AIs ([haveibeentrained.com](#)). In 2022 anti-AI protests broke out on ArtStation and other art portals; DeviantArt responded by letting artists flag their works "noai" to opt out of third-party training, and launched their own image AI, DreamUp, trained only on images whose creators give permission. Learn more at [spawning.ai](#). —Keith Hammond