**The Data Sessions: Prompt Engineering for Creatives**

Generative AI for Creatives

One of the greatest realizations that I think we're experiencing with regards to generative AI is that even though it's very powerful, it still requires a great deal of human involvement. As a person in a creative career, you might be asking, when is GenAI coming to replace me? After all, GenAI shows excellent promise with creative tasks in a few common mediums. Look, I can't tell you the future. But even though the pace of development here is intense, the more I interact with chatbots and GenAI products, the more I realize, yes, it will change how we create, but it's still so completely reliant upon us. As the person prompting the model or chatbot, you are indispensable. You prompt, edit, refine, and modify the outputs and responses of these models to truly get something worth using. GenAI is a worthy assistant. It starts tasks, improves outcomes, and almost completely removes some of the barriers in the creative process. Let's look at how you can leverage GenAI, what to be wary of, and how your final outputs can be enhanced by its use. You may have a creative job in an enterprise, or maybe you're self‑employed, but one thing you're not a stranger to is working in collaboration with others. I suggest we reframe that fear of GenAI taking over into a confidence that GenAI might be a strong collaborator. Still images, text, video, spoken audio, custom music, you name it. There's either already a product made, especially for that task that you can collaborate with, or someone's making it now, I promise. Creating in each one of these mediums is an interesting task to examine. As we jump into each one, evaluate the limitations and strengths of each of the chatbots and products. After all, these are our new GenAI collaborators. Midjourney is a popular product for still image generation. As of today, which is late in 2023, it's interesting because the bot that you interact with is actually housed inside of Discord, a product that allows you to interact and chat in dedicated spaces. Midjourney works when you start first with a command, /imagine. There are a couple of spaces in which you can type this prompt, and the differences are important to understand. First, there are public spaces. The simplest here are labeled with the name newbie on your left. These are completely public. Nothing in here is private. First, I type the command imagine. Then, I describe my desired image. A parade in New York City. You can see that other people are using this channel, and we're all requesting Midjourney to generate images in the same conversational thread. Midjourney begins to iterate, and you see the images begin to materialize. It eventually returns four options, and I have some choices. Those labeled with a U are my upscale options, and I can identify one picture of which to upscale the resolution for continued editing. The others labeled with a V allow me to ask for a variation on a specific image, 1 through 4. I'm just going to upscale image 1 so that we can see what further options I actually have. Now that Midjourney has upscaled my image, I can vary the image in a number of ways. Upscale the resolution further, or zoom in a variety of ways. These arrows allow me to pin the image in a direction. Think of this as if I were holding a camera taking a picture, and I move where I point the camera a little bit in any direction, capturing what's just to the side of the original frame. It's not exactly the same because unlike a camera, Midjourney is actually creating the picture that extends the original image, but this is one way of understanding the panning function here. I can also favorite an image and view it on the web. Every image here can be viewed on the web. Like I said, it's all public. What if I don't really want my images to be in these public channels, though? First, I can chat directly with the Midjourney bot. You just need to click anywhere you see the bot's name and type in the chat box. This will keep my images out of the main channels of Midjourney on Discord, but not off the web. Everything we create here can be placed on Midjourney's explore page, which is public‑facing. Next, as of the time of this recording, in late 2023, you need to upgrade to a more expensive plan to enable a feature called stealth. This is enabled with a command /stealth. I'm only on the basic plan here, so there's really not a way for my images to remain completely private from all audiences. Becoming a Midjourney power user allows you to know and use all the parameter expressions, like ‑ar, to limit your results to a specific aspect ratio, or other commands, like /blend, that allows you to blend two pictures together, or even modifying an existing image that you either upload or reference with its URL. Something you really must understand if you're planning to use Midjourney at work is that Midjourney does not claim ownership over the images that it creates. This means you can use these images in commercial endeavors, but you cannot copyright GenAI images at this point in the United States, so you also don't own the images. Next, you should be very careful to avoid biases that result from Midjourney's model being trained on images that exist already. Many racial, age‑related, and gender biases exist in occupational imagery and other images throughout history. These images are the training dataset at play here. To avoid this, you can prompt Midjourney actively to increase diversity. And you'll need to notice biases in Midjourney's responses. This is true for all GenAI that creates images or other visual mediums. It's one important reason for human involvement today, reviewing and editing the responses. Last, you want to take note of elements in the image that might already be copyright protected. It would be a bad idea for me to use a generated image of a popular well‑known cartoon character for my commercial gain. Just running it through Midjourney doesn't wipe away any existing copyrights. Midjourney isn't the only visual GenAI in town. OpenAI's DALL‑E is another image generation option, and one thing that's really cool about it is that you can use it right in line with ChatGPT. Prompting DALL‑E doesn't require specific commands or syntax. It's very conversational, but it can be a little funky if you want to do any more editing of a generated image. Here, I'm prompting DALL‑E to create a photorealistic image of a fully decorated Christmas tree next to a fireplace in a cozy home all decorated for Christmas. It returns a couple of options, and they're good. I've definitely seen images like this in stores and on greeting cards. Let's convert the style to modern art and see what happens. DALL‑E is handling my requests pretty well. One thing I would caution you to avoid with DALL‑E is details in the human form. DALL‑E and a lot of GenAI struggles pretty hard with hands and faces today. Some of the generated images can be frankly terrifying. It's definitely improving over time, but just as an example, I'll ask for two hands holding. You can see GenAI lacks specific context. Sometimes this is just the number of fingers that are usually on a hand. Other times, it's the fact that hands are connected to arms. It's simply predicting based on your literal prompt and the images on which it was trained. Some opinions I've found online claim that DALL‑E has better photo realism, but Midjourney has more artistic flair. I actually think that's kind of hard to judge, as a good deal of the result you get is dependent on how good your prompt is in the first place. I have found Midjourney's workflow of editing an image to be a little more friendly because it's a tad better at remembering context within a thread. This means that I can reference previous responses well within a thread and the model responds. While we're in OpenAI's interface, though, I want to try a prompt in ChatGPT to generate some copy. There are so many enterprise use cases for creative copy, I can't even begin to list them all. But I think it's worth showing at least one. Pretend I'm a social media marketing specialist. How can I leverage ChatGPT to make my posts better? I'll start a new chat and prompt ChatGPT to create a post. You're a social media marketer at a men's clothing design company called STITXHED. You want to create an Instagram post that encourages men to wear tailored clothing because it makes them look better. Your company offers free measurements and 25% off of a customer's first tailored suit. Your brand is edgy and a little wild, but still for professionals climbing the corporate ladder. Create a post that makes coming in and getting measured sound fun and quick. Include five hashtags that target a rising professional with an edge, five more hashtags that pertain to tailored suits, and five more hashtags that describe how the customers will feel after they visit your store. Notice, I did a lot of specific prompting here to try and get a better result. I could stop there, but I want an accompanying image, so I'll ask for that, too. Also, create a prompt for DALL‑E that will generate a beautiful men's suit on an athletic model. Make the suit light gray and the background dark. Here's what I love. With how DALL‑E and ChatGPT are integrated, I can just tell ChatGPT to generate the image, and DALL‑E jumps in behind the scenes. Now you can see there are some edits to be made, and this again is where you, the human, become so important. If this were your post, you'd evaluate it for brand alignment. You'd modify it and maybe ask ChatGPT to make some changes in the conversation here. Usually at this point, though, I prefer to just copy the responses and start editing them myself on whatever social media platform I intend to use. The image looks a little intimidating, to be honest. I'll ask ChatGPT to modify the image into a more realistic male body type that's lean, but not so muscled. This is a little more inviting for the clientele I'm actually targeting. You can see how much faster initiating a creative visual work can be, but we still have to edit and watch out for biases, and especially for hallucinations in the images. ChatGPT also does a really fantastic job of helping you think differently about whatever task you're trying to accomplish. You can ask ChatGPT a series of related questions and get answers. You can also ask ChatGPT to ask you whatever questions it thinks you need to answer to get a better result. Video may be one of the most difficult of all the GenAI‑created assets to get perfect. There are a number of options for generating video, and in an enterprise setting, I think there are two main tasks to highlight here. The first is done well by Synthesia. They have a number of avatars and voices that you can animate easily to deliver a message. Let's test it out. I'm going to go to their site and generate a free video using text that I type in. After that, I'll need to buy a plan to continue, but this software is a leader in the market right now for animating avatars and delivering messages via video with GenAI. You get a good idea of what's possible with this sample video. They have an intense amount of voices that you can choose from today, and I'm going to animate one to deliver the next few sentences of this course. Is this perfect? No, but what I think this shows is that Synthesia and many other GenAI tools are developing so fast that in the very near future, this might become very normal to see online. It's obviously artificial, and watching it for more than a few seconds and maybe a few minutes becomes a little unnerving. For those of you who create a lot of video and would really like to see AI help in editing stages of video creation, there are definitely options for that. Descript is making waves by allowing your scripts to drive how you edit your film. This data session was actually recorded and edited using Descript. I can upload a video, it transcribes the narration, and then I can easily make cuts and edits by working with the script itself, much like I would in a word processor. Can I make cinematic art films with Descript? Hmm, maybe, if I wanted to really stretch the program. But I wouldn't. It's not the purpose of this software. There really is no prompting involved with Descript, and, to be honest, text‑to‑video conversion has a long way to go, but I will iterate, it's likely the hardest of all generation tasks. At the time of this recording, the text‑to‑video products I could find simply used stock footage paired with text generation, and they arranged them all together on screen with captions, like building with blocks. VEED.IO does add one valuable element in that it will deliver your video to you in a non‑linear editing window, which is great because you're going to have to edit this anyway. Depending upon the GenAI software you're using, you might be able to just pick a royalty‑free track and sync it to the video and you're done. However, text‑to‑audio generation is one area that I am watching very closely. My background originally was in music, specifically music for film. What is completely amazing to me is how the production music library industry is reacting to products like Stable Audio. Royalty‑free music libraries like Pond5 and production music libraries like Warner Chappell have been around for a very long time, and just saying those names makes me hear their audio watermark in my mind, and I have to chuckle a little. Stable Audio is positioning to disrupt the market, though. Just sampling some of the music on the home page is pretty impressive. I love that they show you the text prompt that created the music right there in front of you. You can basically see the prompts with results in Midjourney as well because that product is not private, remember. I feel like reading these helps me to understand the skill of prompt engineering in the context of the media I want as an output. It's pretty great. (Music playing) Let's look at what you can get out of this software. I've previously created four prompts as an illustration for you. The first was an attempt to get a dramatic underscore from Stable Audio. I asked for a pretty granular level of musical elements, and I needed to score a short excerpt of a film. You can see I asked for a specific tempo or speed, and for some shifts in the music along the way. I also asked for a sustained sound at the end. I think GenAI will get to the point where it could handle this, but just listening here, I'm not sure it's a custom dramatic score solution. (Music playing) Next, I wanted to see what would happen if I asked for some orchestral music. Orchestral music that's acoustically recorded has incredible nuance in it, so I didn't expect great results. As I listen to this, what I'm seeing is that the model here doesn't really have a typical western music theory knowledge. That probably sounds super picky, but if you listen closely, the music seems to just move around to any chord regardless of what may be customary or common in practice. This is an example of GenAI not really having the context that we may know from studying music. Instead, it's just predicting what could possibly come next based on a training set. (Music playing) I'm not convinced that GenAI will be writing a symphony on its own that anyone's going to pay to hear over and over anytime soon. But maybe film score and orchestral works aren't a strong part of the training dataset for Stable Audio. What if I tried a more popular music style? I left this prompt vague, intentionally giving the model more freedom to experiment. This time, it's a little more convincing and shows some promise. Once again, the model doesn't really have a great deal of context for harmony and chord progressions. You can hear that we've been listening to the exact same chord the entire time. It's fairly uncommon in music to have zeo variation in harmony, especially in this style of music, but we do have a good beat appropriate to the genre, and other patterns that we hear in house music are present, too. (Music playing) I gave Stable Audio one more contrasting genre. Spa music is pretty simple and usually contains very few noticeable musical elements. It should blend in and create more of a feeling. I think maybe we might have a sweet spot here for GenAI music. (Music playing). As with all GenAI products, if you're using it in a commercial setting, you must read and understand the Terms of Service. These Terms of Service agreements are subject to change, so please understand the terms when you sign up for a service that you're planning on using commercially, and read the updates they send out as they change their terms. More than anything, I hope what you've noticed are the strengths and limitations of GenAI when used with your creative work. These products are great collaborators for creators of content. They can spark creative ideas and kickstart some of your work. But if you're afraid of it replacing you today, let me point out that none of the outputs in my experimental work here were publication‑ready on their own. GenAI relies on you to learn how to prompt, edit, and interact to get an output that you can use. The person in this formula is indispensable today in order to end up with something of high quality. There are still news articles that spread the fear of GenAI replacing us here and there, and, frankly, I hope they don't stop because I appreciate the reminder to keep AI in check. I surely hope we are intelligent enough to know, or at least find, the correct boundaries for GenAI. The flood of AI‑generated ad copy, job listings, commercials, etc., that we're already seeing becomes recognizable the more that you embrace and interact with these products. So don't be afraid. It will be you and your specific touch that makes it art instead of mere data.