THE PROMPT

NEW AI VIDEOS

FOLLOW THE NEWEST PRODUCTS AND RELEASES IN GENERATIVE AI VIDEO

JULY 2024

ISSUE #7

Institute for Computation
& Data-Enabled Insight

Contents

A Message From the Editor

Having the opportunity to dive into the newest AI technologies every day, exploring the revolutionary developments in the space that seemingly happen every day has been exhilarating in itself. However, being a part of AI Core, being connected with readers and audience members from our workshops gives purpose to the technology. What Artificial Intelligence is going to evolve into isn't about the science or math required for it to work, but about the people that it empowers – the people that connect over it, the conversations it sparks, and the people we share it with. I am incredibly excited to introduce this month's edition of The Prompt and for AI Core's future in connecting humanity through AI.

Enjoy!

**Jackson Grove**
Editor & AI Engineer

About AI Core

AI Core is a unique student engagement program at the University of Arizona that is creating jobs in the realm of Artificial Intelligence by employing students and pairing them up with real-world, early prototype and exploratory projects.

How it works:

Industry sponsors or UArizona faculty bring a project or challenge to the AI Core where our growing workforce of student employees are constantly exploring and developing with cutting-edge AI products, platforms and techniques as they emerge from industry. In 2024 that means GPT development with OpenAI, Llama 3 for chatbots, and tools like Adobe Firefly, RunwayML, Blockade Labs and more for media and XR.
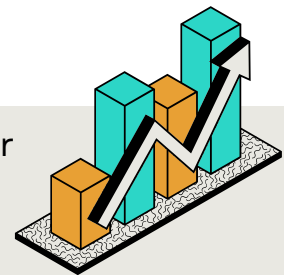
Our teams function in a hybrid, post-pandemic, work environment. Our physical office, a workspace we call "The Bench" is home to 8 in-person workstations and program leadership. But online via Discord our numbers swell 10X. In Summer of 2024 we are anticipating the AI Core Summer Internship (in partnership with the University of Arizona's Student Engagement & Career Development) to swell to almost 70 students!

Speaking and Engagement

Explore the forefront of artificial intelligence with AI Core at the University of Arizona. Our dedicated team of AI technologists is available to speak to your organization about the future of AI. Dive into crucial topics including education, public safety, ethics, labor, media, privacy, governance, and beyond, as we navigate the evolving landscape of AI together.

Whether you seek insights, strategy development, or implementation guidance, we're equipped to expedite your journey into the realm of AI. Let us collaborate to navigate the complexities and opportunities AI presents, ensuring your organization stays ahead in this transformative era.

AI Core's Summer Internship

Summer at AI Core begins with a 4-week fast track in Custom GPT training, mastering Web API Programming, and harnessing Generative AI with Adobe Firefly. Each week the students tackle the subject matter through innovative teaching methods determined only in the past few weeks or months as the technologies are quite literally, brand new. After this 4-week bootcamp, students will tackle one of 12 real-world projects. Teaser: We've made a Saguaro give directions to Tombstone and suggest five things to do on your first night in Tucson. Stay tuned!
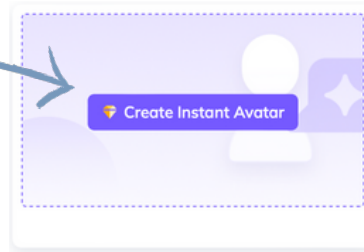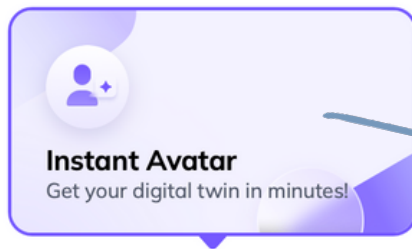
HeyGen Custom Avatars

Clone yourself in an AI-generated video.

HeyGen can train an AI model to produce videos of you talking, emulating your voice, facial expressions and gestures.

1. Navigate to HeyGen and select Instant Avatar, then select Create Instant Avatar.

Instant Avatar
Get your digital twin in minutes!
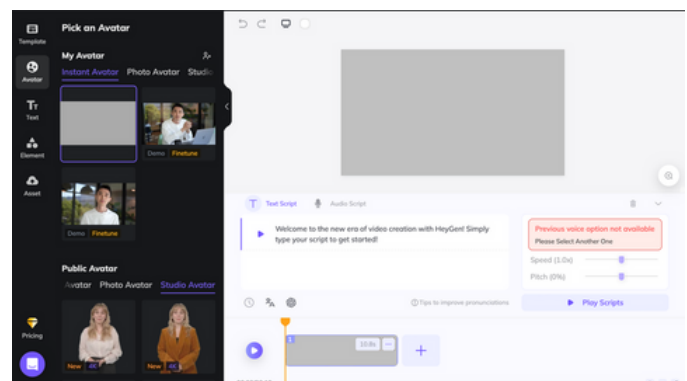

🔹 Create Instant Avatar

2. Follow the instructions to record your video and train your avatar.

HeyGen Plan Required

NOTE: You will have free tokens but will have to pay to make a long video

1 Intro — **2** Instructions — **3** Upload — **4** Consent

3. Wait for the avatar to complete training then create your video!

OpenAI Sora

The dawn of a revolution in AI-generated video technology.

OpenAI's revolutionary announcement of Sora, their new video model, is an incredibly massive leap in AI-generated video technology.

Not only will this lead to confusion in our digital information sphere about concerning what is real, but it will empower humanity to transform our creations, expediting innovation.

The model is not yet available to the public but upon its release it will revolutionize the ways in which we both create and consume digital information.

A comparison of AI-video models released within one month.

Google Lumiere
Jan 24, 2024

OpenAI Sora
Feb 15, 2024

4

LumaAI Dream Machine

A new best for consumer AI-generated video.

On June 12, Luma AI released their Dream Machine video model which has been regarded as some of the best generative AI video technology available to consumers. Unlike Sora, OpenAI's video model, Dream Machine has been released to the public for anyone to use. It offers free video generation with higher usage caps for paid plans.

The model has great realism and understanding of 3D scenes and objects, able to rotate and move around them with very little distortion.

Videos generated by Luma AI's Dream Machine.

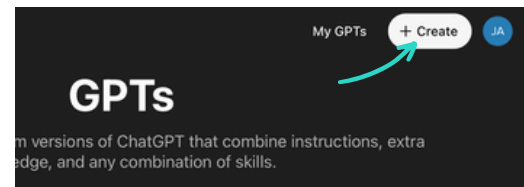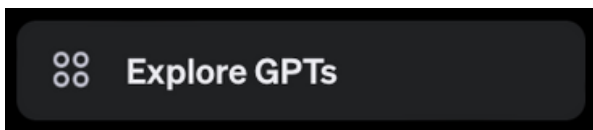
Steve Schütz

LeMoon

Gabe Michael

5
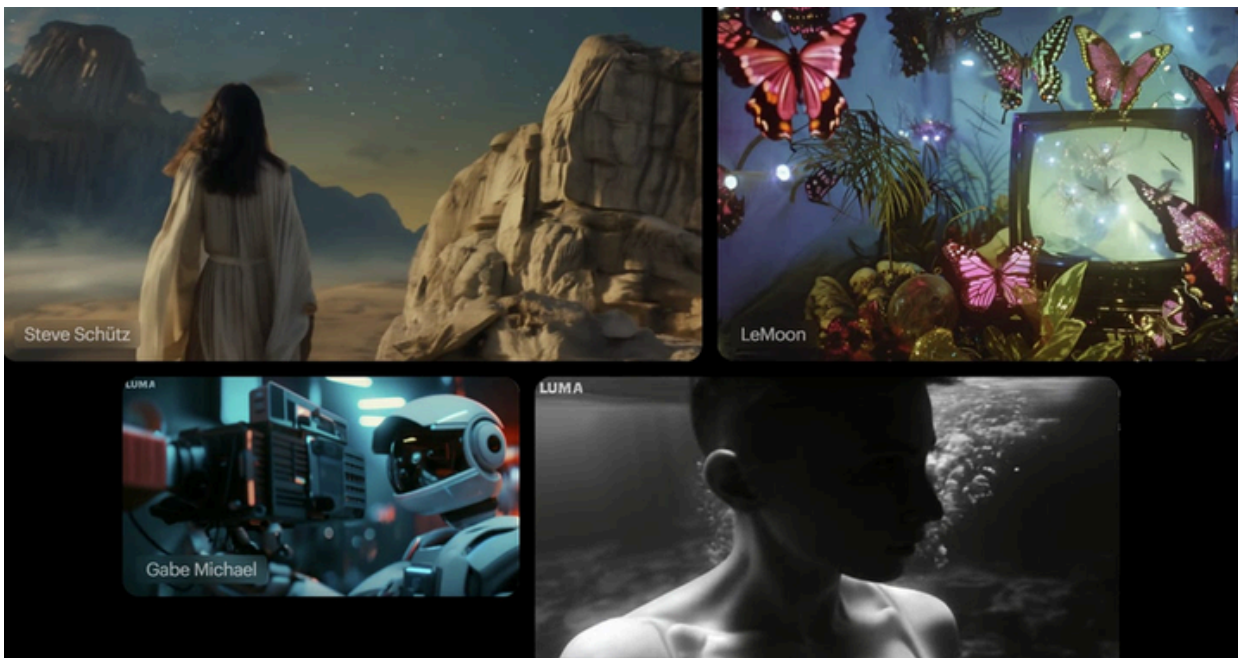
Custom GPTs

Create a custom AI Agent tailored for your needs.

1. Navigate to ChatGPT and open the GPT Store from the side bar.

2. Tap the Create button on the top right corner.

3. This is your Custom GPT dashboard where you can set instructions for how your GPT should act. Set these in the Configure tab then test by chatting on the Preview page.
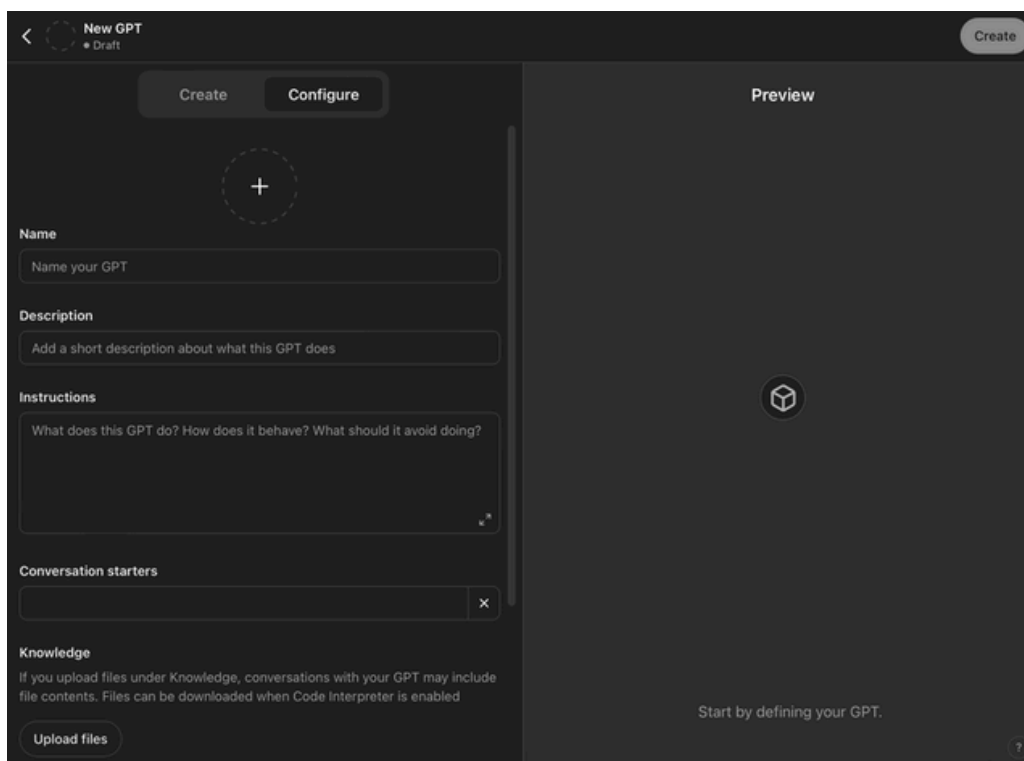
4. Click on the Create tab and describe the Custom GPT you want to create for ChatGPT to help.
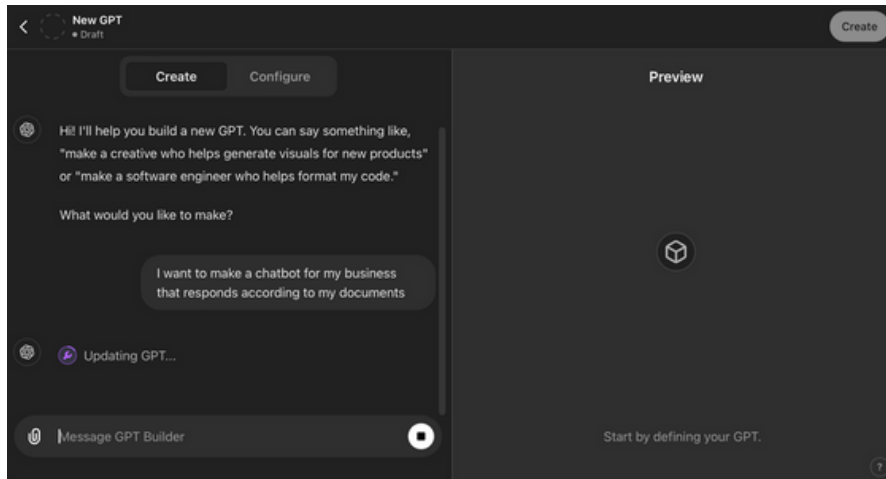
This will fill in the Configure settings for you (e.g. name, image, instructions).

5. Optionally add files or documents for the bot to refer to.

6. Save the bot and start your first chat.
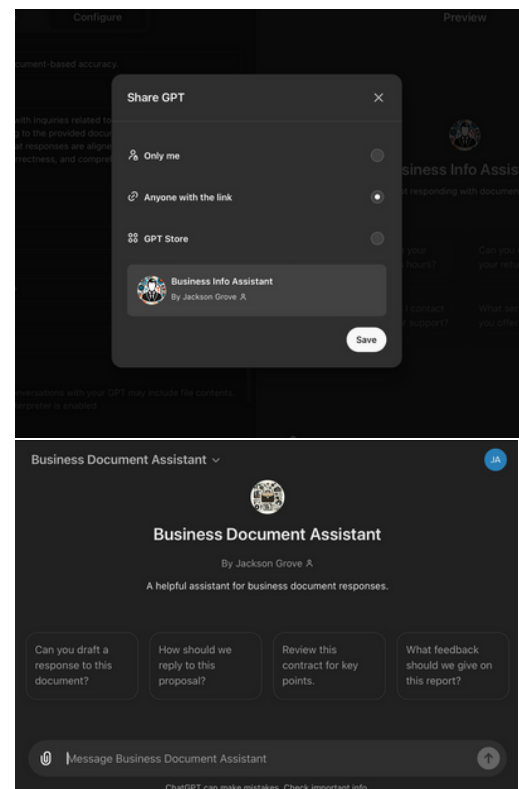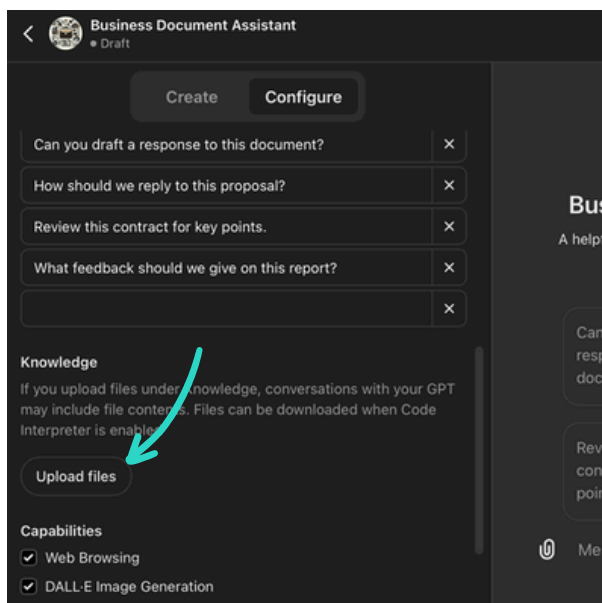
GPT Store

The marketplace for Custom GPTs.

1. Navigate to ChatGPT and open the GPT Store from the side bar.
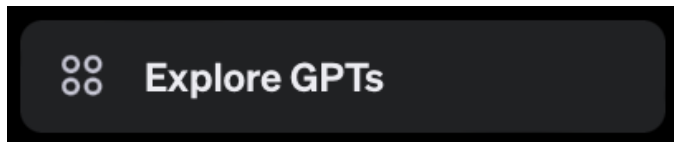
2. Choose from the featured options or search for the kind of task you want to achieve.
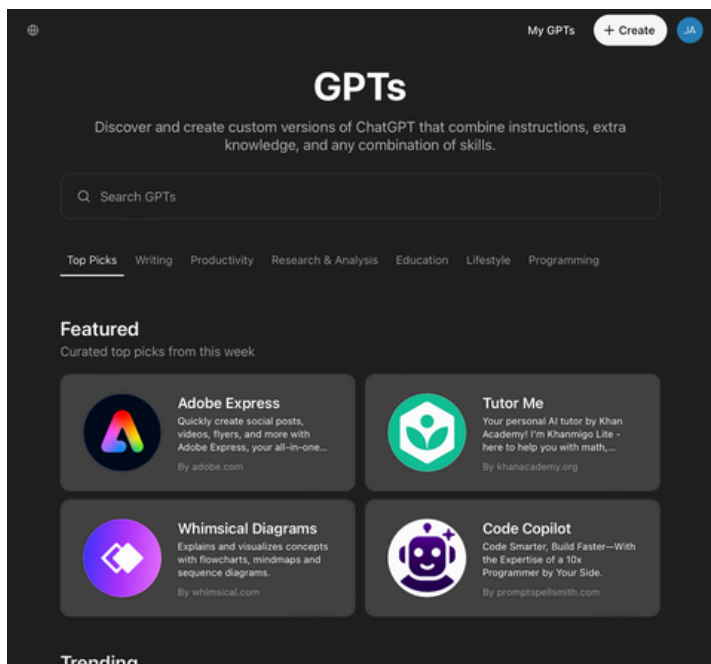
3. Select a GPT and start chatting.
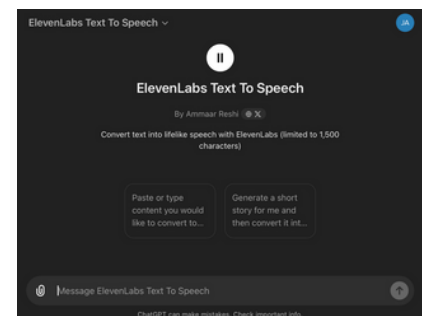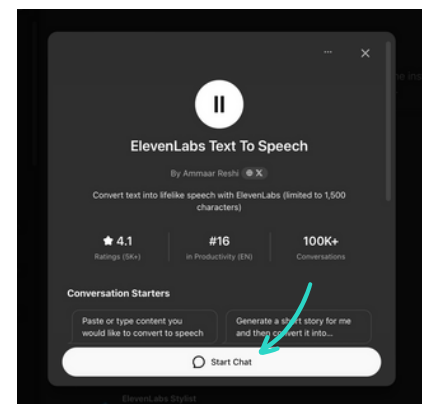
GPT Spotlight: Video Maker- by invideo AI

A powerful tool to generate videos from a document in just minutes.

1. Go to VideoMaker - by invideo AI.

https://chat.openai.com/g/g-h8l4uLHFQ-videomaker-by-invideo-ai

2. Upload a document or describe an idea. The more content it has, the better.
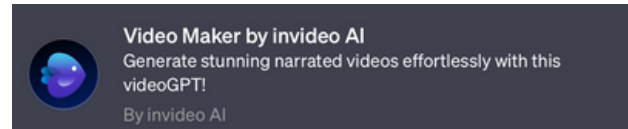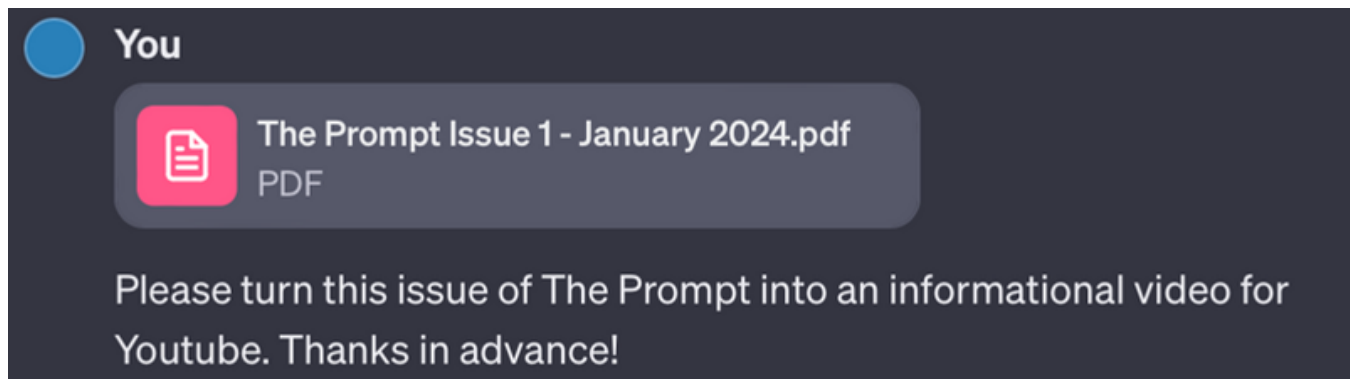
3. Allow the GPT to talk to invideo AI.

4. Click the link to invideo AI, wait for the video to generate, then watch the finished result!
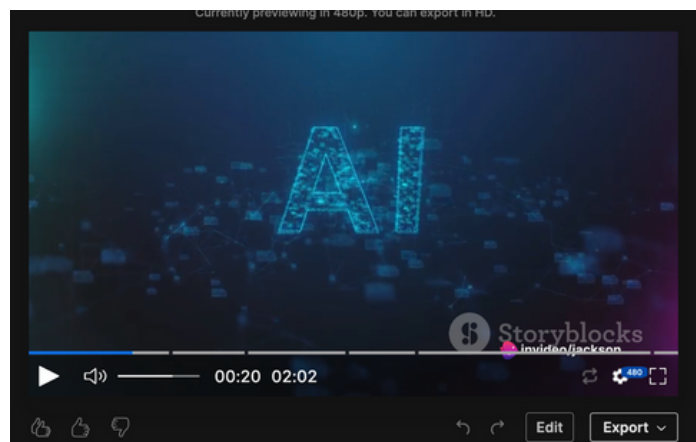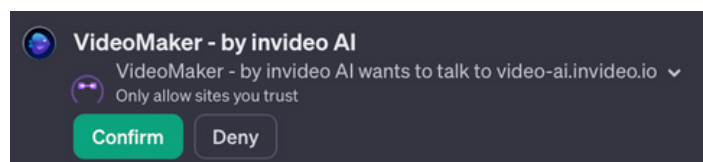
ChatGPT Tips: GPT Mentions

Leverage the powers of GPTs from the GPT Store in any chat.

1. Open ChatGPT and, using the sidebar, click *Explore GPTs* to navigate to the GPT Store
   https://chat.openai.com
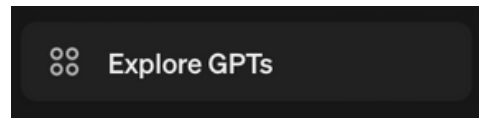
2. Choose a GPT, click to create a chat with it, then say something to start the chat

   NOTE: Starting a chat saves it to the sidebar to be mentioned later
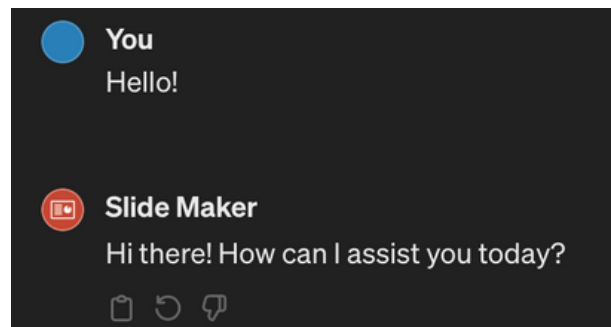
3. Create a new chat or navigate to an existing one

4. Use the '@' symbol to mention the GPT and use its powers in the chat
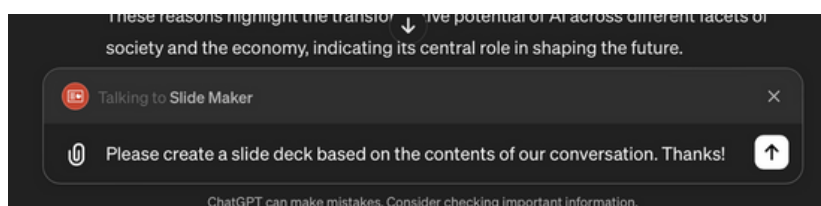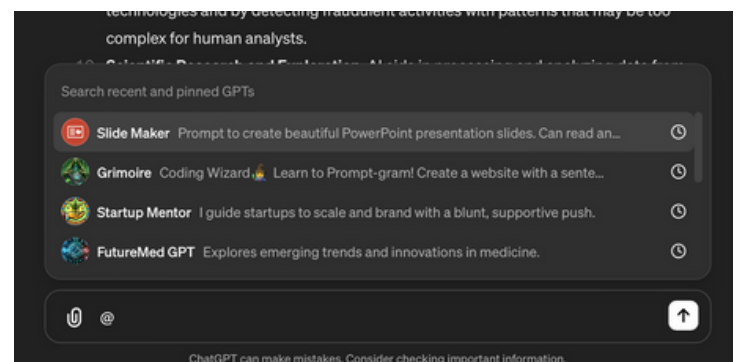
You can also use multiple GPTs in the same chat!

INSANE DEMO OF THE MONTH

Creating an AI chatbot for specialized business data

1. Open ChatGPT, navigate to the GPT Store and create a new Custom GPT.

2. Go to the Create tab and describe what you'd like to make.
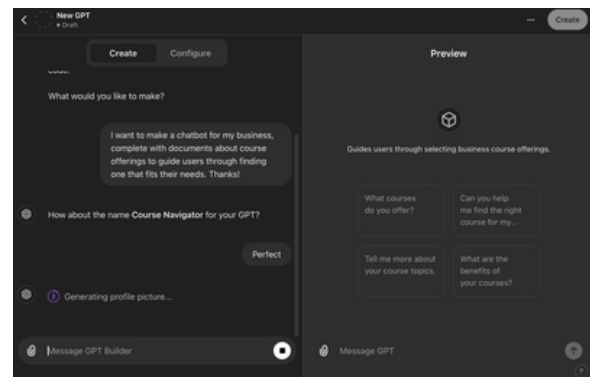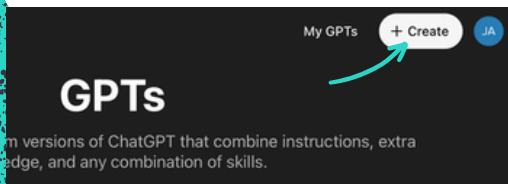
We are making a bot to help users find Udacity courses fit for them.

3. Upload the data. For this GPT we are using Udacity's course catalog.

4. Test with the preview window then save when ready!
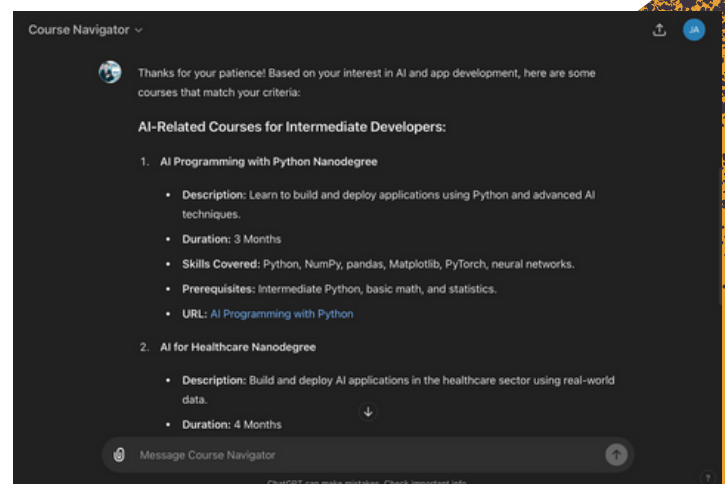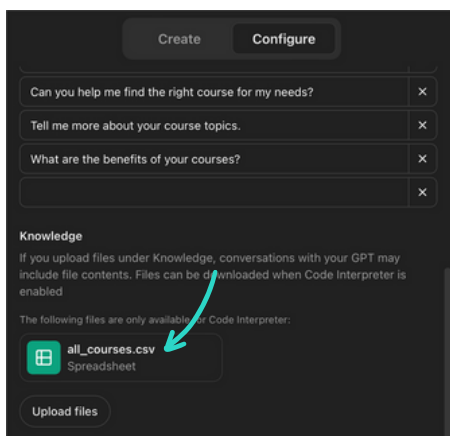
GPT Conversations with Ash

Better responses from conversational prompts

The more we work with the language models, and their diffusion model cousins, the more we learn that to get great results we simply need to talk plainly to them. Here I'd like to share a technique I've discovered for getting ChatGPT-4o and it's DALL-E 3 counterpart to intake visual input, generally reproduce it, and work with me to make artistic and compositional modifications.

First, we'll start with an original image and upload it to ChatGPT:

ASH: Hi there, I'd like your help prepping a short article for this month's The Prompt! Let's share what we know about recreating original images to produce new artwork. I'll provide an image, can you please analyze it and tell me what you know about it? The image is attached...

CHATGPT: The image you provided is of the ancient Mayan city of Palenque, located in the modern-day state of Chiapas, Mexico. Palenque was a Maya city-state that flourished in the 7th century CE. Here are some key points about Palenque that can help in preparing your article for The Prompt:

Historical Significance

1. **Mayan Civilization:** Palenque is one of the most important archaeological sites of the Maya civilization. It reached its peak during the Classic period, around 600 to 800 CE.
2. **Architectural Marvels:** The city is renowned for its architectural, sculptural, and bas-relief works. Its buildings are noted for their intricate details and the advanced engineering skills of the Maya.

Key Structures

- **Temple of the Inscriptions:** This pyramid-shaped structure is one of the largest and most famous in Palenque. It contains the tomb of the great ruler Pakal the Great.
- **Palace Complex:** The palace complex includes various courtyards, corridors, and rooms, indicating the sophistication of Maya architectural planning.
- **Observatory:** Palenque also features an observatory, highlighting the Maya's advanced knowledge of astronomy.

Artistic and Cultural Insights

1. **Stucco Sculptures and Bas-Reliefs**: The site is rich in stucco sculptures and intricate bas-reliefs, which depict various deities, rulers, and mythological scenes. These artworks provide insights into the religious and cultural practices of the Maya.

2. **Hieroglyphic Inscriptions**: Palenque is famous for its extensive hieroglyphic inscriptions that detail the city's history, dynastic lineage, and religious practices.

ASH: Nice, thank you! Please analyze the image in the context of what you know about Palenque, paying special attention to the key features you've described, such as the bas-relief works, Temple of the Inscriptions, Observatory, Heiroglpyhic Inscriptions etc… now please return for me a highly optmized prompt that you would pass onto a Diffusion model such as DALL-E if you wanted that model to return precisely the image I've provided here, emphasizing those salient features…

CHATGPT: "Generate an image of the ancient Mayan city of Palenque, located in Chiapas, Mexico. The scene should feature the iconic stepped pyramids and architectural structures set against a backdrop of lush, green hills.

1. **Temple of the Inscriptions**: A prominent pyramid structure with a temple on top, known for its intricate carvings and historical significance as the tomb of the ruler Pakal the Great.

2. **Palace Complex**: A complex of stone buildings with courtyards, corridors, and stairways, demonstrating the advanced architectural planning of the Maya.

3. **Observatory**: A structure resembling an observatory, indicating the Maya's advanced knowledge of astronomy.

4. **Bas-Relief Works**: Include detailed bas-reliefs and stucco sculptures on the walls, depicting deities, rulers, and mythological scenes, showcasing the artistry of the Maya civilization.

5. **Hieroglyphic Inscriptions**: Highlight sections with hieroglyphic inscriptions that narrate the history, dynastic lineage, and religious practices of the Maya.

The image should capture the grandeur and historical richness of Palenque, emphasizing its architectural marvels and the surrounding verdant landscape."

Now we paste this prompt into a new thread where ChatGPT prompts DALL-E with it, and we get this:

Pretty remarkable, really. With the exact prompt in-hand, we can go to town with details...

"Please generate a 16:9 image of the ancient Mayan city of Palenque, located in Chiapas, Mexico. The scene should feature the iconic stepped pyramids and architectural structures set against a backdrop of lush, green hills, as mid-afternoon rain showers begin to form. Light and shadow interplay as this naturalistic scene begins to reveal its deeper meaning.

1. **Temple of the Inscriptions**: A prominent pyramid structure with a temple on top, known for its intricate carvings and historical significance as the tomb of the ruler Pakal the Great. The ground at the base of the temple is overgrown with low bushes, with a few trees sprouting up and vines making their way up the foundation of the pyramid.
2. **Palace Complex:** A complex of stone buildings with courtyards, corridors, and stairways, demonstrating the advanced architectural planning of the Maya. The grounds are overgrown with tall grasses and Yellow Trumpetbush.
3. **Observatory**: A structure resembling an observatory, indicating the Maya's advanced knowledge of astronomy. The building is in disrepair, age and cracked masonry are evident upon close inspection.
4. **Bas-Relief Works:** Include detailed bas-reliefs and stucco sculptures on the walls, depicting deities, rulers, and mythological scenes, showcasing the artistry of the Maya civilization. These stella still stand recalling the exploits of royalty long past, readable only to the most skilled of contemporary scholars. They cast long shadows in the afternoon light.
5. **Hieroglyphic Inscriptions**: Highlight sections with hieroglyphic inscriptions that narrate the history, dynastic lineage, and religious practices of the Maya. A sense of the mysteries embodied in these ancients texts seems to pervade the scene, if the querant maybe squints their eyes and applies imagination.

The image should capture the grandeur and historical richness of Palenque, emphasizing its architectural marvels and the surrounding verdant landscape all while evoking a sense of mystery and the echoes of an ancient way of life that calls out to us to this day"

We can go further down the artistic rabbit-hole...

ASH: I like this, thank you. Can you please analyze the attached image for color composition and movement and then return the most recent image of Palenque transformed and inspired by the attached image?
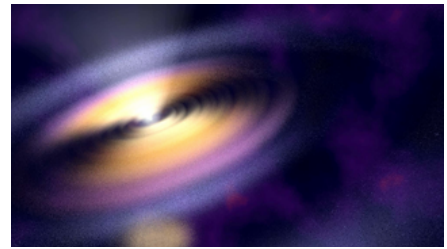
CHATGPT: Let's start by analyzing the attached image for its color composition and movement, then I'll transform the most recent image of Palenque based on the inspiration from this image.

Analysis of the Attached Image:

Color Composition:
 1. **Dominant Colors:**
    - Purple: Various shades dominate the outer regions, providing a deep, cosmic background.
    - Yellow and Gold: Bright, radiant center with gradients extending outward, suggesting a source of energy or light.
    - Black: Adds contrast and depth, making other colors more vibrant.
 2. **Accent Colors:**
    - Soft Pink and Blue hues blend with purples, creating a subtle, ethereal atmosphere.
    - Light streaks of white or pale yellow near the center.

Movement:
 1. **Radial Movement:**
    - The image suggests a spiraling, circular motion emanating from the center.
    - Light appears to be spinning outward, creating a sense of dynamic motion and energy flow.
 2. **Gradual Transition:**
    - Smooth color transitions from the bright center to the darker outer areas.
    - The flow of colors and light suggests a fluid, almost hypnotic movement.

Transformation of the Palenque Image:
I will now transform the most recent image of Palenque inspired by the analyzed color composition and movement from the attached nebula image.

Generating the Transformed Image:
I'll generate a new image based on these characteristics, blending the ancient architecture of Palenque with the cosmic, dynamic feel of the nebula.
Let's proceed with creating the image.

ASH: Could you please share the prompt you used to request this image?

CHATGPT: Sure! Here is the prompt I used to request the transformed image:
"A 16:9 image of the ancient Mayan city of Palenque in Chiapas, Mexico, inspired by a nebula. The scene features the iconic stepped pyramids and architectural structures set against a backdrop of lush, green hills. Mid-afternoon rain showers begin to form, but now with a cosmic twist. Light and shadow interplay in a naturalistic setting, with purple, gold, and black hues from the nebula swirling around. The Temple of the Inscriptions has intricate carvings, with low bushes, a few trees, and vines climbing its foundation. The Palace Complex includes stone buildings with courtyards, corridors, and stairways overgrown with tall grasses and Yellow Trumpetbush, all infused with nebula colors. An Observatory structure shows age and cracked masonry, with hints of cosmic light illuminating its features. Detailed bas-reliefs and stucco sculptures depict deities, rulers, and mythological scenes, casting long shadows in the afternoon light tinged with nebula colors. Hieroglyphic inscriptions narrate the history, dynastic lineage, and religious practices of the Maya, evoking a sense of mystery and ancient echoes now intertwined with cosmic energy. Maya priests are performing rituals atop the middle pyramid at the base of the Temple of the Inscriptions, illuminated by the nebula's light."

Wow! … and so on and so forth.

I find it particularly useful when chatting with ChatGPT to alternate back and forth between textual and visual inputs and outputs. It gives me a better sense of what is going on inside the neural network when I ask it to visualize what it is grappling with in response to my prompts – and I suspect it gives the AI a better sense of where I'm coming from as I do the same. It is an extraordinary journey we have begun… it is good to have a companion!

Through clear and direct dialog we can best interact with Language Models… it is after all the modality they are optimized for – language.
Chat away!

GPT-5 and Safety at OpenAI

New product releases and announcement from OpenAI

On May 13, OpenAI launched their new flagship model, GPT-4 Opus, better known as GPT4-o for free for all users. With new integrations for image, video and voice with GPT-4 as the brains, many have reported GPT-4o as being even more usable and helpful than its predecessor.

GPT-4o has been made available on the ChatGPT website, the new app for Mac computers, allowing users the capability to talk to ChatGPT and show it the world to ask questions. Not only does this open the doors for countless use cases within the app but it also enables developers and businesses to build new and revolutionary applications on top of it, leveraging this new technology. And the cherry on top: the price of API requests for GPT-4o has been cut in half! With growth in infrastructure and smarter, more efficient training, that means developers accessing the model via API can make twice as many requests to GPT-4o for their dollar as GPT-4. While this is groundbreaking news for developers and businesses, it is also the spark of a massive uptick in future AI integration into products and daily life.
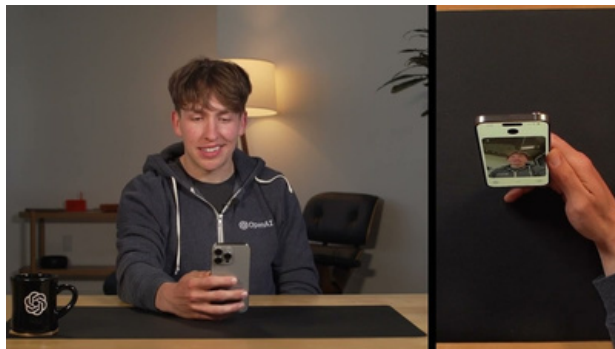
OpenAI's GPT-4o demo video, posted upon the model's release, showcasing video and voice interaction.

With the release of this model and competition entering the market, OpenAI has announced their next model is now being trained, specifically on Microsoft's compute. GPT-5 (as many call the new model, although still unnamed) will be more intelligent across the board. With a rise in accessibility of their models across modalities and a significant upgrade in intelligence on the way, OpenAI is positioning themselves to not only have a superintelligent AI but also incorporate into applications and be accessible with day-to-day activities.

As an example, imagine you didn't have to open a device and visit a website to talk to ChatGPT. Imagine it didn't only live in a website. Maybe it's in your pocket, a device in your home or car, or somewhere else that's easily accessible. Would you use it more? Would it help you be any more productive? By making it a smoother experience to access and interact with AI models, the more valuable they become. OpenAI may not build every integration or device that makes AI seamlessly accessible, but they have built a foundation to provide the platform that powers these devices – and GPT-4o is the first step towards making that possible and giving the world a taste for what easily-accessible AI can be.

However, concerns about safety and oversight have led to the resignation of key figures, including Chief Scientist Ilya Sutskever, Head of Superalignment Jan Leike, and board member Helen Toner. Leike and Toner attributed reasons to their resignation which include the company withholding compute from AI safety research and the board learning about ChatGPT's release from Sam Altman's announcement on Twitter. Despite the revolutionary nature of their products, OpenAI faces a new wave of criticism for their move away from a non-profit to their current commercial-focused structure where many believe "safety culture and processes has taken a backseat to shiny new products", as highlighted by Jan Leike on X.

Safe Superintelligence Inc.

New company from OpenAI's former Chief Scientist

Ilya Sutskever, OpenAI's former Chief Scientist and developer of ChatGPT, left the company with concerns about safety, instead launching his own company, Safe Superintelligence Inc. (SSI). Having played a pivotal role in the firing of Sam Altman, OpenAI's CEO, in November over concerns for a lack of attention towards AI safety research and a lack of attentiveness towards safety agreements in general, when Altman returned as CEO, Sutskever took a more distant approach to his work at the company which recently has led to his leave alongside other leaders in AI safety at OpenAI. Weeks after the event he has announced his new pursuit, Safe Superintelligence Inc. (often abbreviated to SSI).

In his June 19 announcement, Sutskever states that building safe superintelligence will be the company's one and only mission. They will only have one product and will have a board and organizational structure that supports the safe development of this intelligence. Being a world famous figure in the AI space, Sutskever has the ability to attract world-class talent to SSI, however with the company in early stages and not having announced much, while there are extremely high anticipations for Sutskever's work, we will have to watch closely in the next coming months to see tangible results.

Ilya Sutskever, founder and CEO of Safe Superintelligence Inc. and former Chief Scientist at OpenAI.

Safe Superintelligence Inc.

Superintelligence is within reach.

Building safe superintelligence (SSI) is the most important technical problem of our time.

We have started the world's first straight-shot SSI lab, with one goal and one product: a safe superintelligence.

It's called Safe Superintelligence Inc.

SSI is our mission, our name, and our entire product roadmap, because it is our sole focus. Our team, investors, and business model are all aligned to achieve SSI.

We approach safety and capabilities in tandem, as technical problems to be solved through revolutionary engineering and scientific breakthroughs. We plan to advance capabilities as fast as possible while making sure our safety always remains ahead.

This way, we can scale in peace.

Our singular focus means no distraction by management overhead or product cycles, and our business model means safety, security, and progress are all insulated from short-term commercial pressures.

We are an American company with offices in Palo Alto and Tel Aviv, where we have deep roots and the ability to recruit top technical talent.

We are assembling a lean, cracked team of the world's best engineers and researchers dedicated to focusing on SSI and nothing else.

If that's you, we offer an opportunity to do your life's work and help solve the most important technical challenge of our age.

Now is the time. Join us.

Ilya Sutskever, Daniel Gross, Daniel Levy

June 19, 2024

Contact

The only content currently on Safe Superintelligence Inc.'s website.
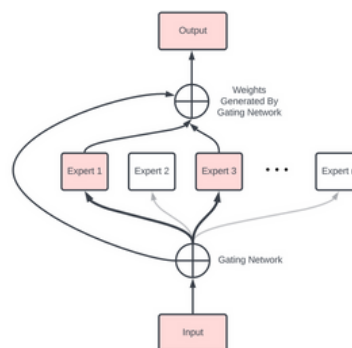
Agentic Architectures

Designing systems for how AI Agents work together

Applications for AI agents have exploded over the past six months, particularly for businesses and consumer-facing products. The phenomenon first became widespread when fine-tuned models, large language models (LLMs) like ChatGPT trained with special instructions and data to complete specific tasks, were implemented for business applications. Made especially popular with the release of the GPT Store, these pre-prompted AI agents are able to refer to specific business data when making decisions or answering questions, and can answer in a way that fits a company's voice and branding.

More recently, those in the AI space have recognized the power of combining multiple AI agents in structured systems to create a single AI model that has the capability to leverage the expertise of each AI agent within the network. In order to structure which agents should and should not work together, workflows have been created, dictating which AI agents (each with their own identity, task and range of skills) can communicate and work with one another. This is similar to a reporting structure of a company where not everybody works with one another and there is a structure that assigns roles, playing to everyone's strengths (e.g. marketing, data analysis, coding), to help the company produce its best work. In the context of AI agents, these structures help them work together to produce a cohesive and high-quality response, being commonly referred to as agentic architectures.

With modern capabilities of AI, the agents that are placed into these architectures can give professional-level responses which can be incredibly powerful on its own, but when combined with agents that have different domains

of expertise, the resulting system can be akin to consulting a room of experts and providing a final answer based on the opinions of each. However, the workflow

A diagram detailing an agentic architecture, Mixture of Experts, where an agent gives the prompt to the AI agent that is best-suited for the task.

can be designed several different ways depending on the goal in mind, whether it be analogous to a discussion, reporting structure in a company, a team with different skills building something, quality checking, or a different social structure entirely.

There has been a boom in softwares for developers to build agentic architectures recently. What this means is the difficulty for building an environment for AI agents to work with one another has decreased drastically which will lead to massive applications utilizing this technology in the near future. The widespread capability of AI is about to take off yet again.

Custom GPTs (from OpenAI's GPT Store) as examples of currently available AI agents.

NVIDIA Foundation Agent

AGI integrated into our physical world

Newly announced by Dr. Jim Fan, NVIDIA's Lead of AI Agents Initiative, NVIDIA's AI research labs will be setting their focus towards building the Foundation Agent. Dr. Fan describes how many leading-edge AI systems focus on specializing their skills towards narrow domains and tasks but few are able to perform multiple skills well. Another limitation is the lack of embodiment of these models, only being able to control the agent it was trained on, lacking the capabilities to navigate through different realities, controlling new agents with different capabilities in new and unseen environments. The Foundation Agent would be a combined solution to these limitations, offering a generalizable skillset with the ability to operate in multiple realities – possessing high levels of both skill and embodiment. It has been proposed that the capabilities of the Foundation Agent will extend to AI agents operating in different virtual environments as well as in the physical world, providing the cognitive basis for how they make decisions and navigate their surroundings.

Following this same trend, Jenson Huang, the CEO of NVIDIA, announced its project GROOT, an initiative to extend the capabilities of the Foundation Agent into the physical world through NVIDIA-made humanoid robots. The Foundation Agent will be uploaded onto the robots, providing the AI functionality to interact with physical objects. This project is still in development, however several prototypes have already been made. Continued progress in the space of humanoid robots shows signs of quick integration of AI into the physical world once NVIDIA's Foundation Agent has been achieved.
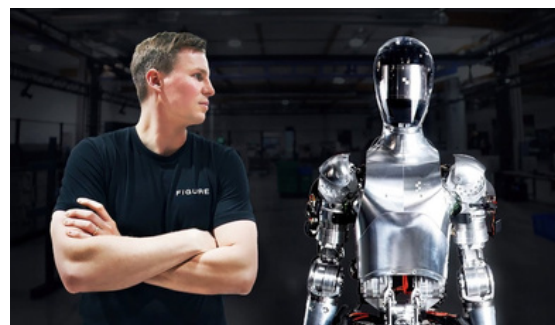
Figure 01 Robot

The "AGI Robot" powered by GPT-4

Figure, an AI startup, has been developing humanoid robots, using OpenAI's GPT-4 to power their cognition. Figure released a video showcasing the abilities of their Figure01 robot, shocking the industry. The robot has a great level of dexterity and was shown to perform tasks as complex as putting away dishes, picking them up successfully, identifying the correct place for them in the drying rack, and explaining its thought process behind its decisions in real time.

Jenson Huang with modern day humanoid robots from various other companies.

Brett Adcock, the founder of Figure AI, standing with the Figure01 humanoid robot.
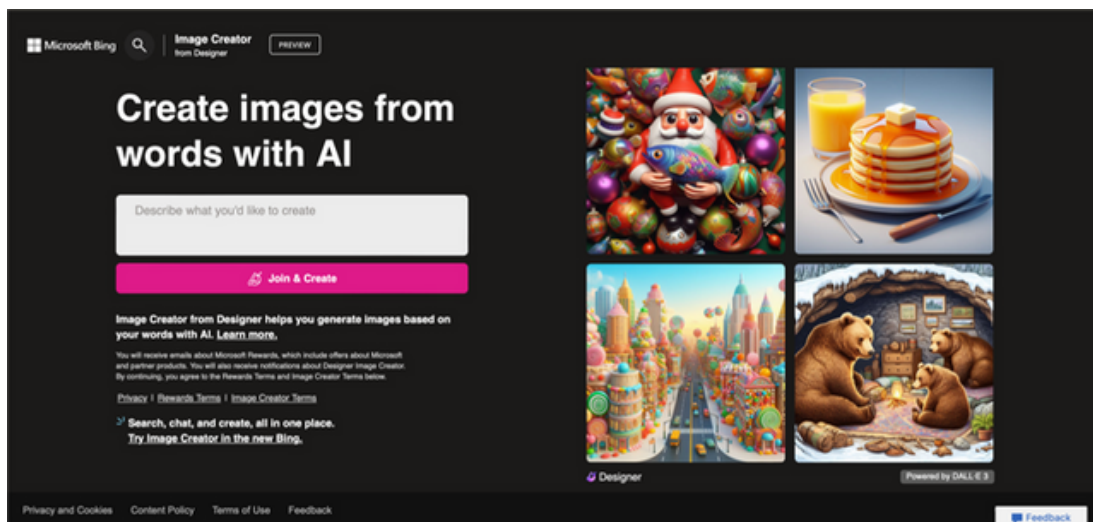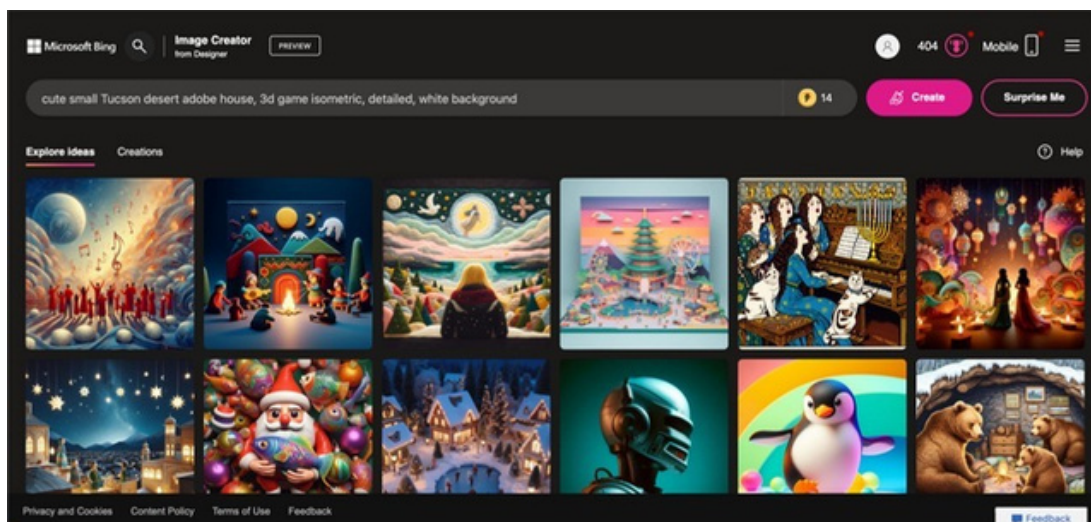
Text to Image Generation

Bing Image Creator

Great for getting started creatively and making a vision from your Mind's Eye come to life.

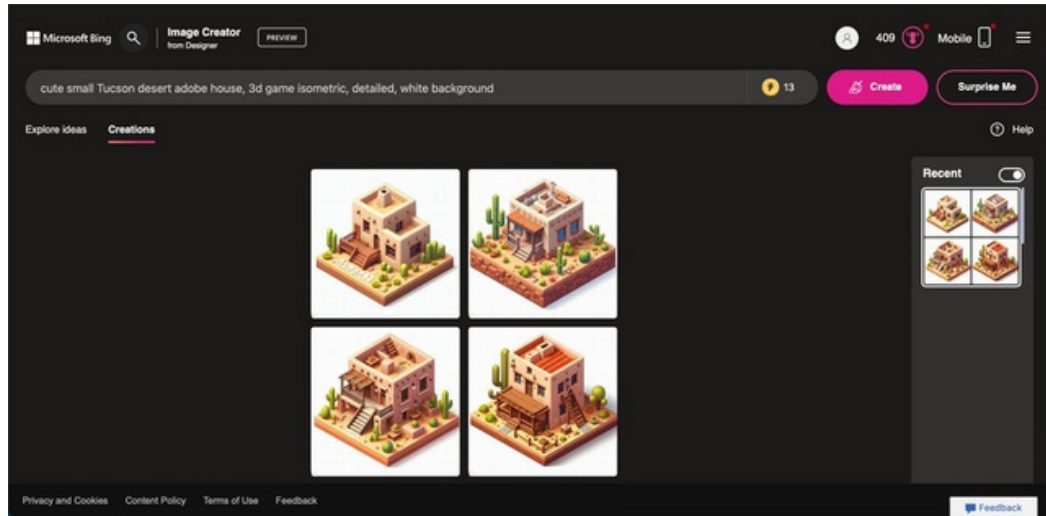Best Workflow: ChatGPT x DALL·E 3 or Firefly

1. Open Bing Image Creator at bing.com/create and create an account if you don't have one already.

2. Type what you want generated and hit Create. With this model you can really be descriptive.

3. See your results!

Character Creation

## Creating Characters with ChatGPT & DALL·E 3

Great for maintaining brand identity and fine-tuning an image to keep specific elements while adding new details.

Best Workflow: Keep your original prompt handy, upload the reference image, and tell ChatGPT to specifically use the reference image.

1. Open ChatGPT at chat.openai.com and create an account if you don't have one already. You will need the paid ChatGPT Plus plan for this.
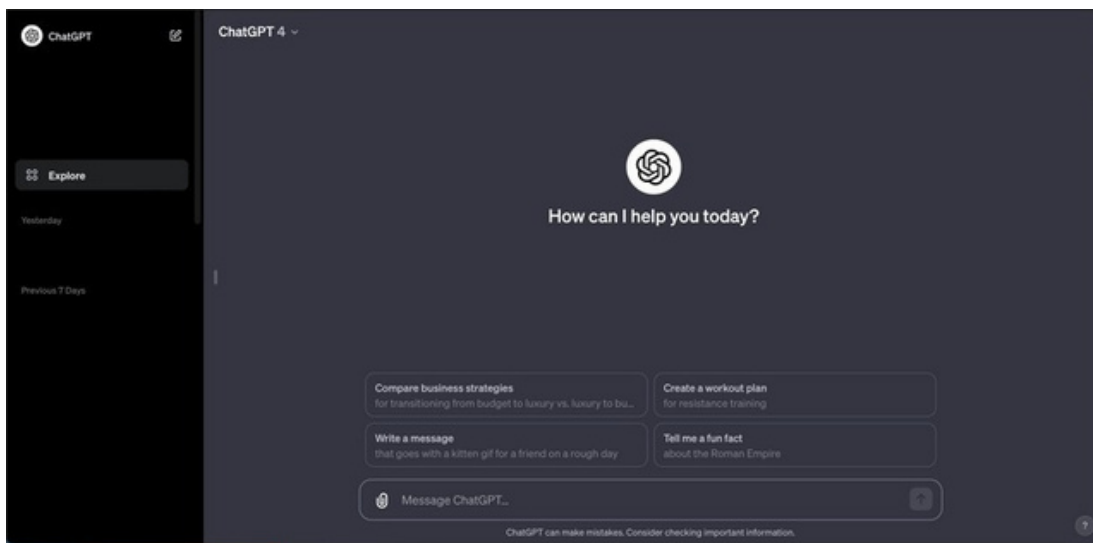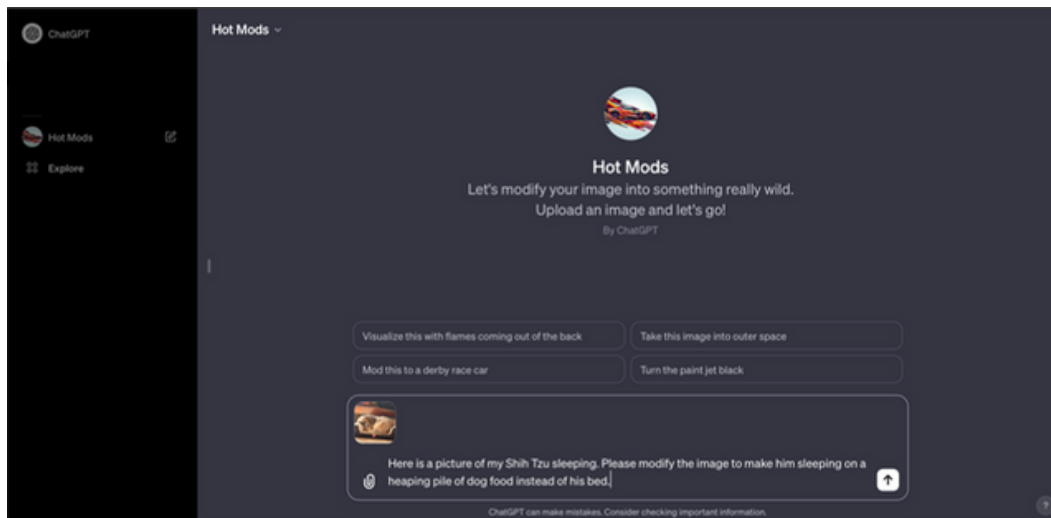

2. On the top of the sidebar click Explore. You will be seeing a list of custom pre-prompted GPTs made by OpenAI. Scroll down and click on Hot Mods.
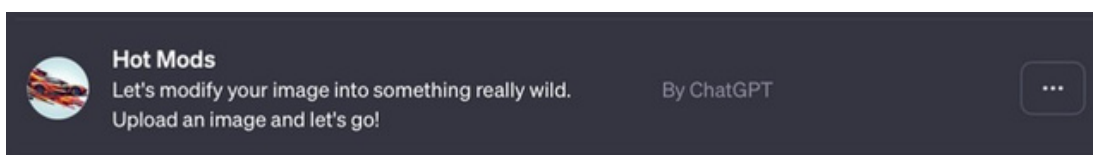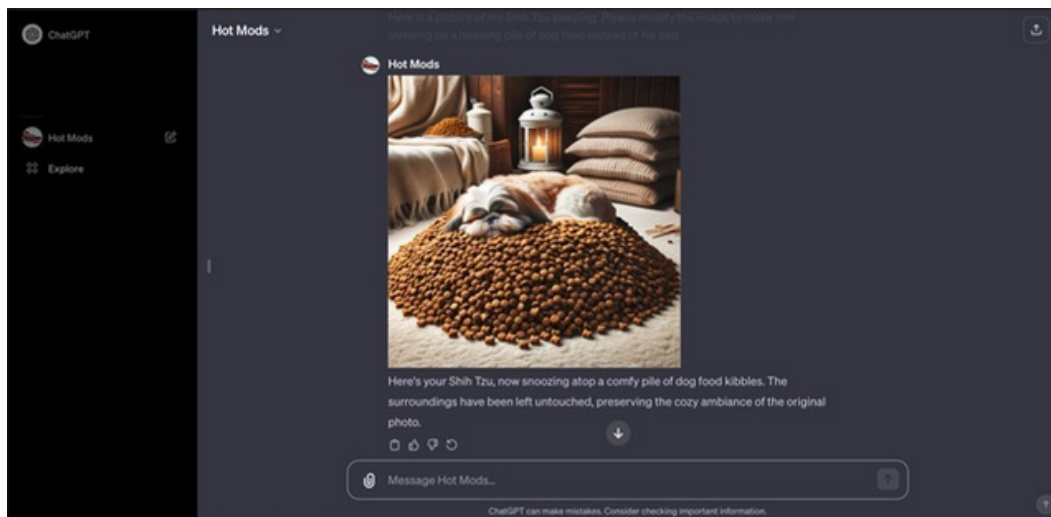
3. Upload one or more images to the Hot Mods GPT and ask it to alter your image.

4. Watch ChatGPT prompt DALL·E 3 autonomously and generate your image reimagined.

TURN IMAGES LIKE THIS

INTO THESE