Using ChatGPT To Control And Leverage Other AI Apps Such As Hugging Face Gets You HuggingGPT, Prompting Eyebrow Raising By AI Ethics And AI Law

Lance Eliot Contributor

Dr. Lance B. Eliot is a world-renowned expert on Artificial Intelligence (AI) and Machine Learning…

Follow

0

Apr 4, 2023, 07:00am EDT

Listen to article 35 minutes

Emerging trend of using ChatGPT as the front-end for leveraging lots of other AI apps is either ... [+] GETTY

Consider the splendors of going to a concert that is showcasing the classics.

When you go to hear an orchestra play, you are witnessing the coordinated actions of a multitude of musicians that are each versed in a particular musical instrument. The conductor makes sure that the otherwise disparate players are working together in

harmony. At times, the conductor might even call upon one particular musician and focus attention on unique musical sounds capable of that specific musical instrument.

The notion of a conductor and coordinating the actions of various specialties is something that the field of Artificial Intelligence (AI) has long been enamored of.

The latest instance of leveraging a federated coordinated approach toward AI is getting some avid attention due to the unveiling of a ChatGPT conglomeration entailing a set of associated AI capabilities via the firm Hugging Face (a company known for having an extensive online library of various AI apps), resulting in a new AI system coined by researchers as HuggingGPT (a catchy name that is mashup of the company name, Hugging Face, mixed with the now ubiquitous GPT acronym that is also part of the naming ChatGPT).

I'm betting that you already have heard about ChatGPT, a generative AI app made by OpenAI.

ChatGPT is a headline-grabber that is widely known for being able to produce fluent essays and carry on interactive dialogues, almost as though being undertaken by human hands. A person enters a written prompt, ChatGPT responds with a few sentences or an entire essay, and the resulting encounter seems eerily as though another person is chatting with you rather than an AI application

To get more details about how ChatGPT works, see my explanation at the link here. If you are interested in the successor to ChatGPT, coined GPT-4, see the discussion at the link

here.

Generative AI is based on a complex computational algorithm that has been data trained on text from the Internet and admittedly can do some quite impressive pattern-matching to be able to perform a mathematical mimicry of human wording and natural language. Please realize that ChatGPT is not sentient. We don't have sentient AI. Do not fall for those zany headlines and social media rantings suggesting otherwise.

ChatGPT by itself is a bit limited. If you want to for example have ChatGPT generate a video for you or an audio output, by and large, you are out of luck. It isn't devised for those types of multimodal tasks. You can get tricky and kind of get ChatGPT to do those specialty tasks, but it isn't pretty or easy to do so.

Aha, you might be thinking, suppose we could have ChatGPT access other apps that would provide various additional capabilities. Sure enough, there is an entire and rapidly flourishing add-on market associated with ChatGPT, see my coverage at the link here.

There are four primary modes of being able to access or utilize ChatGPT:

- **1) Directly.** Direct use of ChatGPT by logging in and using the AI app on the web or soon on your smartphone as an app
- **2) Indirectly.** Indirect use of kind-of ChatGPT (actually, GPT-4) as embedded in Microsoft Bing search engine
- **3) App-to-ChatGPT.** Use of some other application that connects to ChatGPT via the API (application programming interface)
- **4) ChatGPT-to-App.** Now the latest or newest added use entails accessing other applications from within ChatGPT via plugins

The capability of being able to develop your own app and connect it to ChatGPT is quite significant. On top of that capability comes the addition of being able to craft plugins for ChatGPT. The use of plugins means that when people are using ChatGPT, they can

potentially invoke your app easily and seamlessly. See my discussion about the API at the link here. For my analysis of how the plugins will be a game changer, see the link here.

I and others are saying that this will give rise to *ChatGPT as a platform*.

Let's take this same overall notion and consider augmenting ChatGPT by connecting up with other AI apps. You could presumably get the best of both worlds. There is the extensive Natural Language Processing (NLP) and conversational elements of ChatGPT. If you want to have ChatGPT generate video, you could have ChatGPT access another AI app that does video generation. And so on.

In short, the widely popular ChatGPT is in a sense augmented and amplified by being able to leverage a slew of other AI apps that have capabilities not already built into ChatGPT. This in turn allows ChatGPT to take on various tasks that would otherwise be beyond the range and depth of what ChatGPT can do. You could also say that this tends to make access to those other AI apps easier and somewhat seamless because ChatGPT is taking on the chore of coordinating for you.

I trust that you see the two-way street in this.

Other AI apps that maybe aren't as readily accessed or well-known can see the light of day by being used from within ChatGPT. Meanwhile, ChatGPT is perceived by users as more capable because it can undertake tasks outside its usual range via making use of various other AI apps.

A match made in heaven, perhaps.

Just to let you know, not everyone necessarily is overjoyed with this kind of AI-combo or conglomeration construction.

One viewpoint is that the means toward attaining Artificial General Intelligence (AGI) will be best undertaken by combining all manner of AI systems. AGI is the nowadays catchphrase referring to having AI become on par with human intelligence, including potentially having AI enter into sentience. It could be that the path toward AGI will involve mixing together a plethora of otherwise disparate AI apps. By bringing together the narrower AI apps, you could conceivably maybe end up with a totality that goes far beyond what the individual components alone could achieve.

For AI researchers seeking to design and devise AGI, the federated idea of stitching together other AI apps is one avenue that seems worthy of pursuit. You might not have to build one AI app that can do everything. Instead, you build lots of AI apps that do all kinds of things and bring them together in a unified manner. The unification could simply be that one of the AI apps is chosen as the conductor or coordinator of all the rest of them (this is not the only method and other approaches exist of a shared nature or other coordinating mechanisms).

Wonderful, some proclaim, we might be able to arrive at sentient AI by fitting all the pieces together into a magnanimous whole.

Yikes, some respond harshly, you are possibly putting together a modern-day Frankenstein. This oddball conglomeration might morph into a sentient AI that opts to wipe out humankind. For my analysis of the oft-cited existential risks of AI, see the link

here. Even if the combination doesn't reach sentience, you could still have crafted something that will turn ugly, the critics say. Imagine that the AI conglomeration goes awry and launches nuclear missiles or takes other adverse actions. This might not be a result of the AI venturing into human-like capacities and simply due to the automation and algorithms getting ahead of existing safeguards and guardrails.

Ponder this conundrum for a moment.

Does the combining and coordinating of AI apps with other AI apps foretell a better future as a result of having AI systems that seemingly will be larger, more useful, and increasingly capable to produce beneficial results, or does this all maybe spell a doomsday scenario and unseemly foul results for humanity?

That is sometimes referred to as the *AI dual-use* problem, namely that as AI gets bigger and more adept, we can possibly accrue the good aspects of AI, though we also simultaneously are placing a larger and more ominous sword over our own heads, see my discussion at the link here.

In today's column, I will take a close look at the HuggingGPT release and take you through what it does and explain how this works in conjunction with ChatGPT. You might be quite excited and enthralled at this newly unveiled conglomeration. On the other hand, you might be hesitant about embracing these types of amplified AI assemblies. We need to keep our wits about us and not blindly paint our way into a corner that humankind cannot extricate itself from.

Into all of this comes a slew of AI Ethics and AI Law considerations.

There are ongoing efforts to imbue Ethical AI principles into the development and fielding of AI apps. A growing contingent of concerned and erstwhile AI ethicists are trying to ensure that efforts to devise and adopt AI takes into account a view of doing *AI For Good* and averting *AI For Bad*. Likewise, there are proposed new AI laws that are being bandied around as potential solutions to keep AI endeavors from going amok on human rights and the like. For my ongoing and extensive coverage of AI Ethics and AI Law, see the link here and the link here, just to name a few.

The development and promulgation of Ethical AI precepts are being pursued to hopefully prevent society from falling into a myriad of AI-inducing traps. For my coverage of the UN AI Ethics principles as devised and supported by nearly 200 countries via the efforts of UNESCO, see the link here. In a similar vein, new AI laws are being explored to try and keep AI on an even keel. One of the latest takes consists of a set of proposed *AI Bill of Rights* that the U.S. White House recently released to identify human rights in an age of AI, see the link here. It takes a village to keep AI and AI developers on a rightful path and deter the purposeful or accidental underhanded efforts that might undercut society.

I'll be interweaving AI Ethics and AI Law related considerations into this discussion.

The Orchestra Of AI

In a recent research paper posted on March 30, 2023, entitled "HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in HuggingFace", AI researchers Yongliang Shen,

Kaitao Song, Xu Tan, Dongsheng Li, Weiming Lu, and Yueting Zhuang, presented a new AI system that they've opted to call HuggingGPT.

The problem that they say they are trying to solve consists of this identified predicament:

- "Solving complicated AI tasks with different domains and modalities is a key step toward artificial general intelligence (AGI). While there are abundant AI models available for different domains and modalities, they cannot handle complicated AI tasks. Considering large language models (LLMs) have exhibited exceptional ability in language understanding, generation, interaction, and reasoning, we advocate that LLMs could act as a controller to manage existing AI models to solve complicated AI tasks and language could be a generic interface to empower this."

As per my earlier discussion herein, there is a general belief within the AI field that combining together various specialty AI apps might be handy, including that doing so might provide progress toward AGI. No one can say for sure that this in fact is the path or even a proper possibility for attaining AGI. It might not be. In any case, the assumption is that the approach might lend itself to interesting insights about how to devise AI, and therefore one way or another the approach should be pursued.

You might also find it useful to know that the reference to LLMs is customary vernacular within the AI field. ChatGPT can be described as being a Large Language Model or LLM. You see, it is a large model in size and entails being able to pattern-match on natural language. Social media and everyday reporting tend to say that ChatGPT is *generative AI*,

rather than using the more formalized verbiage of LLM. Either usage is acceptable in this context.

Here is what these AI researchers opted to craft as a means of seeking to solve or approach the aforementioned identified AI problem:

- "Based on this philosophy, we present HuggingGPT, a system that leverages LLMs (e.g., ChatGPT) to connect various AI models in machine learning communities (e.g., HuggingFace) to solve AI tasks. Specifically, we use ChatGPT to conduct task planning when receiving a user request, select models according to their function descriptions available in HuggingFace, execute each subtask with the selected AI model, and summarize the response according to the execution results. By leveraging the strong language capability of ChatGPT and abundant AI models in HuggingFace, HuggingGPT is able to cover numerous sophisticated AI tasks in different modalities and domains and achieve impressive results in language, vision, speech, and other challenging tasks, which paves a new way towards AGI."

I'll take a moment to explain this.

Envision that ChatGPT is serving as an orchestra conductor (minus of course any sentience).

ChatGPT is in this case specially set up to access a variety of other AI apps that can augment or supplement what ChatGPT currently is able to do (in a sense, other musical instruments in my analogy to an orchestra). When a user is making use of ChatGPT, this

augmented version of ChatGPT will call upon those other AI apps when needed. The user doesn't necessarily have to prod or urge ChatGPT to do so. Presumably, ChatGPT can computationally identify when it would be useful to invoke one or more of those other AI apps.

For example, imagine that you are using ChatGPT and you decide to enter a prompt telling ChatGPT to create a video depicting horses running in an open field. The likely usual response from ChatGPT would be that it does not have video creation capabilities at this time. If you are using the HuggingGPT app, presumably the prompt would be parsed computationally and the algorithm would ascertain that you want a video to be generated. This in turn would cause the AI app to invoke a Hugging Face suite tool that can generate video. Etc.

I'll discuss such an example in a moment.

Note that the supplemental or augmented AI apps in this circumstance are all part of the Hugging Face suite of AI apps. This is essentially a matter of convenience or choice by these particular AI researchers. You can anticipate that other AI apps will undoubtedly be utilized in a similar fashion by other AI researchers. Overall, realize that you can generalize from this approach and that this research instance is an exemplar of how we will likely and gradually see more such AI-related conglomerations.

Furthermore, these researchers opted to use ChatGPT as the conductor, but it is readily feasible to use some other AI app in that role. Selecting ChatGPT is prudent due to its

immense popularity and ease of use as a result of the NLP capabilities. Nonetheless, you could use some other NLP in that same role.

ChatGPT sits at the front end of the augmentation. A user enters a prompt into ChatGPT. This prompt is computationally examined. If needed, one or more AI apps in the Hugging Face suite are made use of. The result comes back to ChatGPT. ChatGPT then presents the results to the user. All in all, the user only interacts with ChatGPT. They might not realize what is taking place behind the scenes, nor would they likely normally care to know (well, it depends on whether knowing might make a difference).

Here's how the researchers describe this in their research paper:

- “The LLM first plan a list of tasks based on the user request and then assigns expert models to each task. After the experts execute the tasks, the LLM collects the results and responds to the user.”

For this, they suggest that a new concept arises:

- “Therefore, we introduce a concept: ‘Language is a generic interface for LLMs to connect AI models’. In other words, by incorporating these model descriptions into prompts, LLMs can be considered as the brain to manage AI models such as planning, scheduling, and cooperation. As a result, this strategy enables LLMs to invoke external models for solving AI tasks.”

I’d like to take a moment and proffer a comment about some of the wording that from time to time is being used by those within the AI field. Anytime there is a reference to AI

as being a “brain” please be cautious. These kinds of references seem innocuous. All that the researcher is probably trying to do is suggest that the AI is doing something that we associate with the overarching abilities of the human brain. Unfortunately, this kind of referencing tends to anthropomorphize AI. It leads people to assume that today’s AI is or is soon to be sentient. In my columns, I have repeatedly and respectively urged that such misleading wording not be used, see my exploration at the link here.

Moving on, here are the four major tasks that the HuggingGPT is said to undertake, according to the research paper:

- “Task Planning: Using ChatGPT to analyze the requests of users to understand their intention, and disassemble them into possible solvable sub-tasks via prompts.”
- “Model Selection: Based on the sub-tasks, ChatGPT will invoke the corresponding models hosted on HuggingFace.”
- “Task Execution: Executing each invoked model and returning the results to ChatGPT.”
- “Response Generation: Finally, using ChatGPT to integrate the prediction of all models, and generate answers for users.”

Those tasks are akin to what I have mentioned. First, there is task planning whereby ChatGPT computationally attempts to figure out what series of tasks might be needed to carry out the prompt or request by the user. Next, ChatGPT computationally identifies

which of the Hugging Face apps to make use of, including the appropriate sequencing or maybe concurrency involved. Then the selected AI apps are utilized and the results are brought back into ChatGPT. Finally, ChatGPT derives a response for the user and provides that response.

Pretty much a straightforward and logical way to arrange things.

The AI researchers indicated that they have experimented with this AI conglomeration in a variety of ways:

- "Up to now, our HuggingGPT has integrated hundreds of models on HuggingFace around ChatGPT, covering 24 tasks such as text classification, object detection, semantic segmentation, image generation, question answering, text-to-speech, and text-to-video. Experimental results demonstrate the capabilities of HuggingGPT in processing multimodal information and complicated AI tasks."

An especially handy feature is that they opted to use ChatGPT to produce an explanation of what it has done when making use of the Hugging Face suite. A step-by-step explanation can be generated. This allows the user of ChatGPT to inspect the explanation, perhaps being able to ergo find something amiss or at least be aware that there was all this other processing taking place under the hood.

Consider an example that was showcased in the research paper and pertains to asking ChatGPT to generate a video. There are many other examples in the research paper and I

am simply for brevity herein choosing to highlight this one example about generating video.

Here is purportedly the entered prompt:

- “Query: generate a video with the title ‘An astronaut is walking in space’ and dub it.”

ChatGPT would presumably computationally examine the prompt and arrive at the somewhat unstated or implied directive to generate a video about an astronaut walking in space. The video might normally by default be silent. The prompt though indicates that the video is to be dubbed and thus is to contain an audio track.

According to the research paper, the HuggingGPT app proceeded accordingly, and this is part of the generated explanation about what was undertaken:

- “Response: Sure. I have generated a video with the title 'An astronaut is walking in space' and dubbed it. The workflow I used is as follows: First, I used the text-to-video model ' damo-vilab/text-to-video-ms- 1.7b' to generate a video from the text. The inference result is the generated video ' /videos/5696.mp4'. Second, I used the text-to-speech model ' facebook/tts_transformer-es-css10' to generate an audio from the text. The inference result is the generated audio ' /audios/3f9e.flac '. Finally, I combined the video and audio to generate the dubbed video. I hope this answers your question. Is there anything else I can help you with?”

The generated explanation might seem somewhat hard to read because it contains the internal program names and file names. Those details would be helpful to the developers. Undoubtedly, you could use ChatGPT to summarize or omit the technical details if you wanted to do so.

Wondering About Wonderment

We shall now mull over this handy-dandy AI advancement and do so from a big-picture perspective.

First, some wonderment.

There is a lot of complexity underneath the surface to get all of this to work together appropriately. I mention this salient point to emphasize that it isn't as easy to pull off as might otherwise seem to be the case. Imagine trying to use a plethora of AI apps that each have its idiosyncratic facets. It could be a nightmare to figure out how to use each one of them (a technical topic that is often referred to as inter-model cooperation protocols).

For users of AI, the beauty of this kind of setup is that you can do one-stop-shopping. You make use of a handy interface such as ChatGPT. The chosen conductor or coordinating AI app does the rest of the heavy lifting for you. There aren't any added actions you need to take. If all is done well, the other components are fruitfully utilized and all you care about is that the results are valuable to you.

Score one point for AI apps.

The downsides though are aplenty.

We will start by exploring the more innocent adverse aspects.

Suppose that the conductor-designated AI app, in this case, ChatGPT, misinterprets your request. Maybe you did not intend for ChatGPT to generate a video of astronauts walking in space. Instead, you wanted a video of flowers and merely wanted a title for the video that said it was of astronauts walking in space (if this seems farfetched, perhaps it is an art project, and the user is trying to make some kind of esoteric and artsy statement).

ChatGPT opts to computationally proceed to invoke several of the Hugging Face apps. This chews up expensive computer processing cycles that perhaps you are paying for. Upon seeing the resulting video, you realize that your request was improperly undertaken. Yikes, you have to pay the computer processing bill anyway.

That's a relatively small inconvenience and presumably can be dealt with.

Envision a different scenario.

Here's where the existential risk proclaimers come into the picture.

ChatGPT or some other AI interactive conversational app has been hooked up with other AI apps that can control a factory that makes ball bearings. Via a user entering a prompt, the conductor-oriented AI app essentially controls and activates those other AI apps. This allows a factory manager to easily enter prompts in natural language and have the factory operate according to their entered instructions.

But suppose that the conductor-oriented AI app misinterprets the user-entered prompt. Perhaps the prompt is computationally analyzed to instruct the other AI apps to start flinging ball bearings throughout the factory. The machines in the factory seem to have gone crazy. You can extrapolate this type of scenario to all manner of doomsday possibilities. For example, see my coverage about the use of AI for warfare and the controlling of weapons of mass destruction, at the link here.

The gist is that there is a rising amount of handwringing that we are going to inexorably become dependent upon these front-end AI apps that do our bidding, yet the results might not be what we intend.

Some would insist that any AI app that controls or activate other AI apps ought to first ask the user whether this is what they intended to have occurred. A retort is that the user might not know what those other AI apps do, or the person might mindlessly say yes to something bad, or overall, not be in a posture to properly ascertain whether the conductor-oriented AI app is doing the right thing or the wrong thing.

Another twist is that perhaps an evildoer might utilize the AI conglomeration in nefarious ways.

Even if the factory manager was careful and thoughtful, suppose that a cyberhacker was able to steal the login of the factory manager or in some manner break into the conductor-oriented AI app. This wrongdoer intentionally instructs the AI to control and activate the other AI apps to do berserk actions. Thus, just because the controlling AI might first let

the user know what the AI is going to do, this doesn't matter since the devious person purposefully wants something bad to happen anyway.

You might be tempted to say that all of this can be dealt with by ensuring that the conductor-oriented AI has sufficient capabilities to assess the user prompts. If a user enters a prompt that one way or another appears to be dangerous, the mainstay AI ought to refuse to abide by the request.

The problem with this considered solution is that trying to constrain a natural language dialogue and detect when the user is being sneaky is a lot harder than you might assume. In the case of ChatGPT, an elaborate effort was made to detect when user prompts are questionable. Nonetheless, users have been able to find ways around those precautions, see my coverage at the link here of how people are able to get generated hate speech and other undesirable outputs from ChatGPT.

Realize that today's AI lacks any semblance of common sense. When using ChatGPT, users are apt to fall into the mental trap of believing that they are conversing with a sentient being. They are not. That's also why it is so important to devise generative AI to not falsely suggest to the user that it is perhaps sentient. For example, ChatGPT and other such AI apps tend to respond to the user with the word "I" as though the AI app is a human being. A lousy and outright bad practice. The AI developers can easily stop that kind of wording, and when they don't, I refer to this as *anthropomorphizing by design* (a practice that ought to be curtailed).

Let's add more fuel to the fire about qualms concerning these emerging AI conglomerations.

You might vaguely know that generative AI such as ChatGPT has many flaws. Besides the possibility of producing offensively worded essays and interactions, there are many additional and extremely disconcerting issues about today's generative AI.

Four concerns about generative AI that I have extensively covered include:

- **1) Errors**. Generates wording and essays that have errors of fact or miscalculations, etc.
- **2) Falsehoods**. Generates false assertions and other insidious falsehoods.
- **3) Biases.** Generates wording and essays that contain biases of nearly any and all kinds.
- **4) AI Hallucinations**. Generates what appears to be factual but is made-up and not at all factually based (I don't like the term "AI hallucinations" due to the anthropomorphizing of AI, but it seems to be a catchphrase that has regrettably gained acceptance, see my discussion at the link here).

Lest you shrug off those pitfalls, realize that people using generative AI are bound to fall into the trap of accepting the outputted essays as truthful and factual. Doing so is easy-peasy. You see lots of essays and interactions that seem on par with human levels of

fluency and confidence. You get lulled into assuming that everything uttered is of the utmost correctness.

Even the most ardent supporters of generative AI would acknowledge that we have severe problems associated with the generation of errors, falsehoods, biases, and AI hallucinations. No reasonable AI researcher or AI developer could disagree with that contention.

We can tie this back to the AI conglomerations. You are using ChatGPT and it is connected to other AI apps. Upon your entered prompt, ChatGPT generates a falsehood that is then utilized as a part of the invoking of the other AI apps. Assume that these other AI apps aren't programmed to detect when falsehoods are fed into them. Thus, they do their thing and provide back perhaps false results. This in turn is then presented to the user by ChatGPT.

The same concerns can go in the other direction too. ChatGPT correctly interprets a prompt, but somehow one or more of the other AI apps that are invoked opts to produce a result that contains errors, falsehoods, biases, or AI hallucinations. Presumably, ChatGPT is going to take this as a correct result and present it as such to the user.

Crucial checks and balances are needed throughout these AI conglomerations.

Whether those checks and balances receive sufficient attention from AI developers is an open question. There is a race right now to make advances in AI. Those in the AI Ethics field are worried that the rapt desire to make AI advances is outpacing any similar desire

to ensure that checks and balances are also being incorporated. A rising interest in drafting and enacting new AI laws is gaining steam, partially due to these and other concerns, see my analysis at the link here.

Conclusion

Some final remarks for now on this topic.

Eugene Ormandy, the famous music composer, said this about orchestras: “Watch me closely - only one can spoil it.”

Suppose that an AI conducting a symphony of other AI apps gets ensnared by one bad apple in the bunch. If the AI at the front of the pack isn’t able to anticipate beforehand the potential impact, one might say that the entire opus could be ruined. In the case of music, the cringe that results might be mild. When the AI is undertaking more life-involving efforts, the offbeat note might be deadly.

Is the pursuit of federated AI involving AI apps that collectively work together and as led by one anointed AI a considered viable approach that we should herald as advancing us flourishingly towards AGI?

It might work out okay, but on the other hand, it might not be music to our ears and we could find ourselves in a chaotic ear-shattering morass. Keep listening. Help out in whatever way you can.

Let’s make sure that AI aligns with and is harmonious with humankind.

Follow me on Twitter.

Lance Eliot

Follow

Dr. Lance B. Eliot is a Stanford Fellow and a world-renowned expert on Artificial Intelligence (AI) with over 6.8+ million amassed... **Read More**

Editorial Standards

Reprints & Permissions

Join Our Conversation

One Community. Many Voices. Create a free account to share your thoughts. Read our community guidelines here

Commenting as **Guest**

Log in Sign up

Be the first to comment...

No one seems to have shared their thoughts on this topic yet

Leave a comment so your voice will be heard first.