

How to Use AI to Do Stuff: An Opinionated Guide

Covering the state of play as of Summer, 2023



ETHAN MOLLICK

JUL 15, 2023



542



61

Share

Increasingly powerful AI systems are being released at an increasingly rapid pace. This week saw the debut of Claude 2, likely the second most capable AI system available to the public. The week before, Open AI released Code Interpreter, the most sophisticated mode of AI yet available. The week before that, some AIs [got the ability to see images](#).

And yet not a single AI lab seems to have provided any user documentation. Instead, the only user guides out there appear to be Twitter influencer threads. Documentation-by-rumor is a weird choice for organizations claiming to be concerned about proper use of their technologies, but here we are.

I can't claim that this is going to be a complete user guide, but it will serve as a bit of orientation to the current state of AI. I have

been putting together a Getting Started Guide to AI for my students (and interested readers) every few months, and each time, it requires major modifications. The last couple of months have been particularly insane.

This guide is opinionated, based on my experience, and focused on how to pick the right tool to do things. I have written separately about [the kinds of tasks you may want AI to do](#), which might be useful to read first.

The Major Large Language Models

When we talk about AI right now, we are usually talking about Large Language Models, or LLMs. Most AI applications are powered by LLMs, of which there are just a few Foundation Models, created by a handful of organizations. Each company gives direct access to their models via a Chatbot: OpenAI makes GPT-3.5 and GPT-4, which power [ChatGPT](#) and Microsoft's [Bing](#) (access it on an Edge browser). Google has a variety of models under the label of [Bard](#). And Anthropic makes Claude and [Claude 2](#).

There are other LLMs I won't be discussing. The first is [Pi](#), a chatbot built by Inflection. Pi is optimized for conversation, and

really, really wants to be your friend (seriously, try it to see what I mean). It does not like to do much besides chat, and trying to get it to do work for you is an exercise in frustration. We also won't cover the variety of open source models that anyone can use and modify. They are generally not accessible or useful for the casual user today, but have real promise. Future guides may include them.

So here is your quick reference chart, summarizing the state of LLMs:

MODEL	RUNS CODE	SEE IMAGES	READS FILES	INTERNET CONNECTION	PERSONALITY	WHAT IS IT	WHEN TO USE IT
ChatGPT/ GPT-3.5				No	Neutral, bland	This is the free version of ChatGPT that came out in November. It is very fast and pretty solid at writing and coding tasks.	It is fast and cheap and capable, but other models are now better. Also, It is not connected to the internet, so don't use it like a search engine.
ChatGPT/ GPT-4		Soon		No	Helpful, a little preachy	Currently available to paying customers. The most powerful LLM for most purposes. Does a lot.	Better at everything (writing, coding, summarizing) than GPT-3.5 Still not connected to the internet.
ChatGPT/ Code Interpreter/ Plugins	Yes		Yes	Limited	Helpful sometimes witty	Adds new capabilities to ChatGPT. Plugins and web browsing still have rough edges. Code Interpreter lets the AI run Python code & work with files.	Code Interpreter is great for working with data of all sorts, and is the most generally impressive AI implementation so far. Plugins are generally not that useful, yet. The web browsing is being updated.
Bing (Uses GPT-4 in creative & precise modes)		Yes	Yes	Yes	Friendly but also somewhat insane	Connected to the internet. It is GPT-4 with a grab-bag of powerful features. The weirdest model. Creative mode lets you use GPT-4 and is free.	Does a remarkably wide range of stuff, from working as a browser companion to creating art. Is connected to the internet. Has a strong, weird personality.
Bard	Yes	Yes		Yes	None	Bard is a name for a variety of models used by Google. Current offerings are weak, but improving over time	This should be your last choice of AIs for now, but recent updates show promise. High levels of hallucination.
Claude 2			Yes	No	Pleasant when not scolding	Almost as powerful as GPT-4, but built to be "safer" and more pleasant to use. Has 100k context window - enough "memory" to hold a book.	Terrific on working with large documents, due to its extended context window and ability to upload content. It is very new so it is unclear how it stacks up.

as of Summer 2023, oneusefulthing.org

The first four (including Bing) are all OpenAI systems. There are basically two major OpenAI AIs today: 3.5 and 4. The 3.5 model kicked off the current AI craze in November, the 4 model

premiered in the Spring and is much more powerful. A new variation uses plugins to connect to the internet and other apps. There are a lot of plugins, most of which are not very useful, but you should feel free to explore them as needed. Code Interpreter as is an extremely powerful version of ChatGPT that can run Python programs. If you have never paid for OpenAI, you have only used 3.5. Aside from the plugins variation, and a temporarily suspended version of GPT-4 with browsing, none of these models are connected to the internet. Microsoft's Bing uses a mix of 4 and 3.5, and is usually the first model in the GPT-4 family to roll out new features. For example, it can both create and view images, and it can read documents in the web browser. It is connected to the internet. [Bing is a bit weird to use, but powerful.](#)

Google has been testing its own AI for consumer use, which they call Bard, but which is powered by a variety of Foundation Models, most recently one called PaLM 2. For the company that developed LLM technology, they have been pretty disappointing, although improvements announced yesterday show they are still working on the underlying technology, so I have hope. It has already gained the capability to run limited code and interpret images, but I would generally avoid it for now.

The final company, Anthropic has released Claude 2. Claude is most notable for having a very large context window - essentially

the memory of the LLM. Claude can hold almost an entire book, or many PDFs, in memory. It has been built to be less likely to act maliciously than other Large Language Models, which means, practically, that it tends to scold you a bit about stuff.

Now, on to some uses:

Write stuff

Best free options: [Bing](#) and [Claude 2](#)

Paid option: [ChatGPT 4.0](#)/ChatGPT with plugins

For right now, GPT-4 is still the most capable AI tool for writing, which you can access at Bing (select “creative mode”) for free or by purchasing a \$20/month subscription to ChatGPT. Claude, however, is a close second, and has a limited free option available.

These tools are also being integrated directly into common office applications. Microsoft Office will include a copilot powered by GPT and Google Docs will integrate suggestions from Bard. [The implications of what these new innovations mean for writing are pretty profound.](#)

Here are some ways to use AI to help you write.

- Writing drafts of anything. Blog posts, essays, promotional material, speeches, lectures, chose-you-own adventures, scripts, short stories - you name it, AI does it, and pretty well. All you have to do is prompt it. Prompt crafting is not magic, but basic prompts result in boring writing, [but getting better at prompting is not that hard, just work interactively with the system.](#) You will find AI systems to be much more capable as writers with a little practice.
- Make your writing better. Paste your text into an AI. Ask it to improve the content, or for suggestions about how to make it better for a particular audience. Ask it to create 10 drafts in radically different styles. Ask it to make things more vivid, or add examples. Use it to inspire you to do better work.
- Help you with tasks. AI can do things you don't have the time to do. Use it like an intern to write emails, create sales templates, give you next steps in a business plan, and a lot more. [Here is what I could accomplish with it in 30 minutes in supporting a product launch.](#)
- [Unblock yourself.](#) It is very easy to get distracted from a task by one difficult challenge. AI provides a way of giving yourself momentum.

The image displays three screenshots of AI chat interfaces, each showing a response to the prompt: "You are an excellent writer of nonfiction bestsellers. Write a paragraph clearly explaining how Large Language Models work using an analogy."

GPT-3.5: Large Language Models are like master storytellers with an encyclopedic knowledge of the world. Imagine sitting by a cozy fireplace, engrossed in a conversation with an exceptionally knowledgeable and creative individual. You can ask them any question or give them a prompt, and they will draw upon their vast knowledge, experiences, and understanding of language to craft a compelling and coherent response. Similarly, Large Language Models, such as the GPT-3.5 architecture, have been trained on an enormous amount of text data and can generate text that resembles human language. By analyzing patterns and context, they generate responses that aim to be informative, engaging, and contextually appropriate. Like a skilled storyteller, they can provide answers, generate ideas, or even assist in creative writing, making them invaluable tools for a wide range of nonfiction applications.

Claude 2: Here is a paragraph explaining how large language models work using an analogy: Large language models are like extremely skilled chefs that have tasted and memorized millions of recipes. When you give them a writing prompt, it's like telling the chef what ingredients you have on hand. The model looks through its vast memory to find relevant ideas and themes that would go well with those ingredients. It then strings together sentences and paragraphs the way a chef combines ingredients into a coherent dish, drawing on its knowledge of language structure and patterns. The model strings ideas together into a smooth flow just as a chef layers flavors, working to create something engaging yet easy to digest. While not infallible, large language models can deliver thoughtful and fluent text on a wide range of topics through their expansive knowledge and deft linguistic processing.

Bard: Sure, here is a paragraph explaining how Large Language Models work using an analogy: Analogy: Imagine a giant library filled with books. Each book represents a different piece of text, such as a news article, a blog post, or a book chapter. A Large Language Model (LLM) is like a librarian who has read every book in the library. The librarian can use their knowledge of the books to answer your questions about the library's contents. For example, if you ask the librarian "What is the capital of France?", the librarian can use their knowledge of the books to find the answer. How it works: LLMs are trained on massive datasets of text. The training process is similar to how a human learns a language. The LLM is presented with a large number of examples of text, and it learns to identify the patterns and relationships between words. Once the LLM is trained, it can be used to answer questions, generate text, and translate languages. Example: For example, if you ask an LLM to write a poem about a cat, the LLM will use its knowledge of the books in the library to generate a poem that is similar to the poems that it has read. The LLM will also use its knowledge of the English language to ensure that the poem is grammatically correct and that the words flow together in a natural way. I hope this analogy helps you to understand how Large Language Models work.

Some things to worry about: In a bid to respond to your answers, it is very easy for the AI to “hallucinate” and generate plausible facts. It can generate entirely false content that is utterly convincing. Let me emphasize that: ***AI lies continuously and well.*** Every fact or piece of information it tells you may be incorrect. You will need to check it all. Particularly dangerous is asking it for references, quotes, citations, and information for the internet (for the models that are not connected to the internet). Bing will usually hallucinate less than other models, because GPT-4 is generally more grounded and because Bing’s internet connection means it can actually pull in relevant facts. [Here is a guide to avoiding hallucinations](#), but they are impossible to completely eliminate.

And also note that AI doesn't explain itself, it only makes you think it does. If you ask it to explain why it wrote something, it will give you a plausible answer that is completely made up. When you ask it for its thought process, is not interrogating its own actions, it is just generating text that sounds like it is doing so. This makes understanding biases in the system very challenging, even though those biases almost certainly exist.

It also can be used unethically to manipulate or cheat. You are responsible for the output of these tools.

Make images

Most transparent option: [Adobe Firefly](#)

Open Source Option: [Stable Diffusion](#)

Best free option: Bing or [Bing Image Creator](#) (which uses DALL-E), [Playground](#) (which lets you use multiple models)

Best quality images: [Midjourney](#)

There are four big image generators available for most people:

1. Stable Diffusion, which is open source and you can run from any high-end computer. It takes effort to get started, since you have to learn to craft prompts properly, but once you do it can produce great results. It is especially good for

combining AI with images from other sources. [Here is a nice guide to Stable Diffusion if you go that route \(be sure to read both parts 1 and part 2\).](#)

2. DALL-E, from OpenAI, which is incorporated into Bing (you have to use creative mode) and Bing image creator. This system is solid, but worse than Midjourney.
3. Midjourney, which is the best system in mid-2023. It has the lowest learning-curve of any system: just type in "thing-you-want-to-see --v 5.2" (the --v 5.2 at the end is important, it uses the latest model) and you get a great result. Midjourney requires Discord. [Here is a guide to](#) using Discord.
4. Adobe Firefly, built into a variety of Adobe products, but it lags DALL-E and Midjourney in terms of quality. However, while the other two models have been unclear about the source images that they used to train their AIs, Adobe has declared that it is only using images it has the right to use.

Here is how they compare (each image is labelled with the model):



Adobe



DALL-E



Midjourney



Stable
Diffusion 2.1

Prompt: "Fashion photoshoot of sneakers inspired by Van Gogh" - the first images that were created by each model

Some things to worry about: These systems are built around models that have built-in biases due to their training on Internet data (if you ask it to create a picture of an entrepreneur, for example, you will likely see more pictures featuring men than women, unless

you specify “female entrepreneur”), you can use [this explorer](#) to see these biases at work.

These systems are also trained on existing art on the internet in ways that are not transparent and [potentially legally and ethically questionable](#). Though technically you own copyright of the images created, legal rules are still hazy.

Also, right now, they don't create text, just a bunch of stuff that looks like text. But Midjourney has nailed hands.

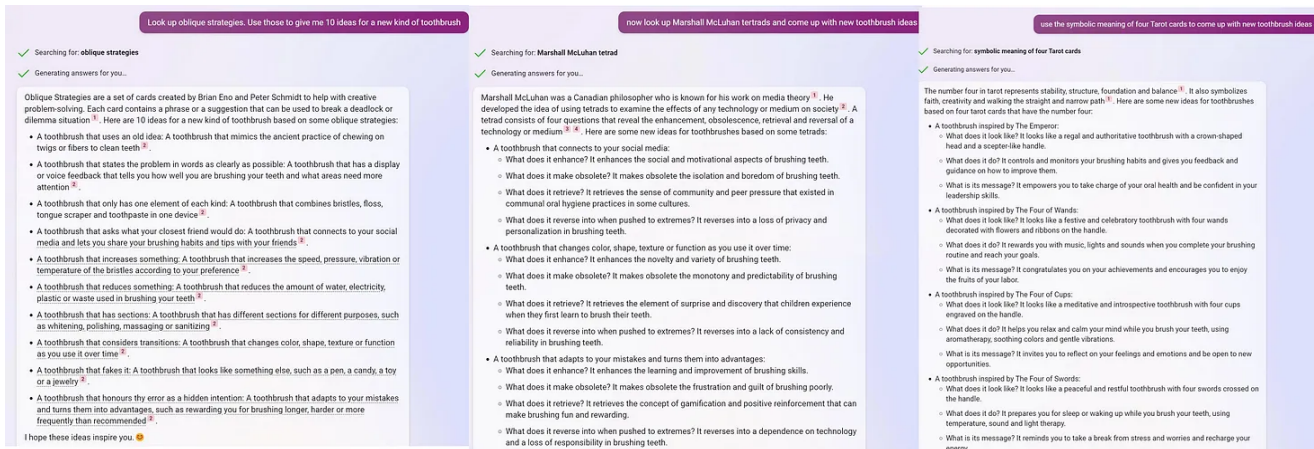
Come up with ideas

Best free option: [Bing](#)

Paid option: [ChatGPT](#) 4.0, but Bing is likely better because of its internet connections

Despite of (or in fact, because of) all its constraints and weirdness, AI is perfect for idea generation. You often need to have a lot of ideas to have good ideas, and AI is good at volume. With the right prompting, you can also force it to be very creative. Ask Bing in creative mode to look up your favorite unusual idea generation techniques, like Brian Eno's oblique strategies or Mashall McLuhan's tetrads, and apply them. Or ask

for something weird, like ideas inspired by a random patent, or your favorite superhero...



Make videos

Best animation tool: [D-iD](#) for animating faces in videos. [Runway v2](#) for creating videos from text

Best voice cloning: [ElevenLabs](#)

It is now trivial to generate a video with a completely AI generated character, reading a completely AI-written script, talking in an AI-made voice, animated by AI. [It can also deepfake people, as you can see in this link where I deepfaked myself. Instructions and more information here.](#) Use with caution, but this can be great for explainer videos and introductions.

The first commercially available text-to-video tool was also recently released, Runway v2. It creates short 4-second clips, and is more of a demonstration of what is to come, but is worth taking a look at if you want a sense of the future development in this space.

Some things to worry about: Deep fakes are a huge concern, and these systems need to be used ethically.

Work with documents and data

For data (And also any weird ideas you have with code): Code Interpreter

For documents: Claude 2 for large documents or many documents at once, Bing Sidebar for smaller documents and webpages (the sidebar, part of the Edge browsers can “see” what is in your browser, letting Bing work with that information, though the size of the context window is limited)

[I wrote about Code Interpreter last week.](#) It is a mode of GPT-4 that lets you upload files to the AI, allows the AI to write and run code, and lets you download the results provided by the AI. It can be used to execute programs, run data analysis (though you will need to know enough about statistics and data to check

its work), and create all sorts of files, [web pages](#), and even [games](#). Though there has been a lot of debate since its release about the risks associated with untrained people using it for analysis, many experts testing Code Interpreter are pretty impressed, [to the degree that one paper suggests it will require changing the way we train data scientists](#). Go to my previous post if you want more details on how to use it. I also made an initial prompt to set up Code Interpreter to create useful data visualizations. It gives it some basic principles of good chart design & also reminds it that it can output many kinds of files. You can find that [here](#).

For working with text, and especially PDFs, Claude 2 is excellent so far. I have pasted in entire books into the previous version of Claude, with impressive results, and the new model is much stronger. You can see my previous experience, and some prompts that might be interesting to use, [here](#). I also gave it numerous complex academic articles and asked it to summarize the results, and it does a good job! Even better, you can then interrogate the material by asking follow-up questions: what is the evidence for that approach? What do the authors conclude? And so on...

Claude 2

PDF A+Review+of... 211.35 KB

PDF SSRN-id43752... 524.69 KB

PDF Falling+Asleep... 391.55 KB

summarize these three documents and the connections between them

Here is a summary of the main points and connections between the three articles:

The first article is a research paper examining the effects of AI assistance on human productivity and performance in an HR recruiting task. The study finds that high-performing AI leads to worse performance compared to lower-performing AI, as humans are more likely to "fall asleep at the wheel" and overly rely on the high-performing AI's recommendations.

The second article is about ChatGPT, a new generative AI system. It describes results from an experiment showing that ChatGPT increases productivity in professional writing tasks by allowing workers to generate content more quickly while maintaining quality. ChatGPT appears to mostly substitute for human effort rather than complementing human skills.

The third article reviews applications of ChatGPT across sectors like e-commerce, finance, healthcare and education. It highlights use cases where ChatGPT streamlines processes and provides benefits through natural language interaction.

Connections:

- Both the first and second articles find that AI systems can substitute for human effort and lead to productivity gains, though too high-performing AI may have downsides (human disengagement)
- The second and third articles are about ChatGPT specifically. The second focuses on an experimental evaluation while the third reviews real-world applications.
- All three suggest AI systems like ChatGPT will have significant impacts on tasks and sectors involving generating text, conversing with customers/patients, etc.
- Both the first and second papers point to risks of overreliance on AI and the need to keep humans engaged. The third article is more optimistic about ChatGPT's potential.
- The findings on productivity impacts in the first two articles likely explain some of the promising applications covered in the third article.

Copy Retry

Code Interpreter

As an AI trained by OpenAI, I can help create a wide range of visualizations, including but not limited to:

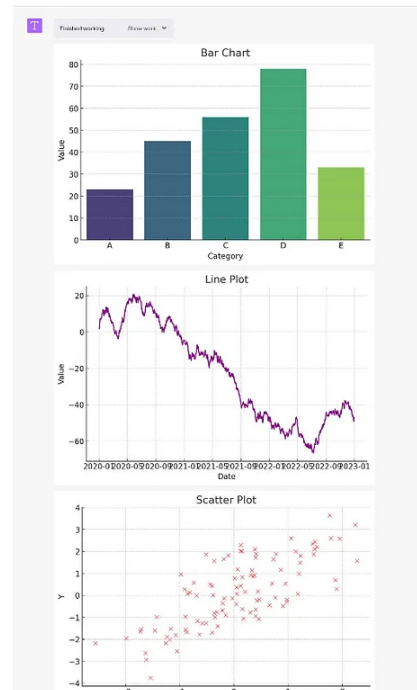
1. Bar charts: These are effective for comparing quantities across different categories.
2. Line plots: These are useful for showing trends over time.
3. Scatter plots: These are used to visualize relationships between two variables.
4. Box plots: These are useful for displaying the statistical distribution of data.
5. Heatmaps: These can be used to visualize patterns or correlations in matrix data.
6. Histograms: These are useful for visualizing the distribution of numerical data.
7. Pie charts: These are used to show proportions of a whole.
8. Area charts: These are useful for showing cumulative totals over time.
9. Treemaps: These are used to visualize hierarchical data.

The output formats I can generate include:

1. Static images (like .jpg, .png, .svg).
2. Interactive visualizations (like HTML files, which can be embedded in web pages).
3. Animated graphics (like .gif).
4. Interactive maps (using libraries like Folium).
5. Data tables.

Now that we have established this, I'd be happy to help you create a visualization! Could you please tell me what kind of data visualization you are interested in? If you have a specific dataset you'd like to visualize, please feel free to upload it and provide a brief description of what you're hoping to achieve.

Can you show me some samples of very good graphics using these approaches. Feel free to make up data.



Some things to worry about: These systems still hallucinate, though in more limited ways. You need to check over their results if you want to ensure accuracy.

Get information and learn stuff

Best free option: [Bing](#)

Paid option: Usually Bing is best. For children, [Khanmigo](#) from Khan Academy offers good AI-driven tutoring powered by GPT-4.

If you are going to use AI as a search engine, probably don't do that. The risk of hallucination is high and most AIs are not connected to the Internet, anyway (which is why I suggest you use Bing. Bard, Google's AI, hallucinates much more). However, there is some evidence that AI can often provide more useful answers than search when used carefully, [according to a recent pilot study](#). Especially in cases where search engines aren't very good, [like tech support, deciding where to eat, or getting advice](#), Bing is often better than Google as a starting point. This is an area that is evolving rapidly, but you should be careful about these uses for now. [You don't want to get in trouble](#).

But more exciting is the possibility of using AIs to help education, including helping us learn. [I have written about how AI can be used for teaching](#) and to [help make teachers' lives easier and their lessons more effective](#), but it can also work for self-guided learning as well. You can ask the AI to explain concepts and get very good results. This [prompt is a good automated tutor](#), and you can find a [direct link to activate the tutor in ChatGPT here](#). Because we know the AI could be

hallucinating, you would be wise to (carefully!) double-check any critical data against another source.

And more?

Thanks to rapid advances in technology, these are likely the worst AI tools you will ever use, as the past few months of development have shown. I have no doubt I will need to make a new guide soon. But remember two key points that remain true about AI:

- AI is a tool. It is not always the right tool. Consider carefully whether, given its weaknesses, it is right for the purpose to which you are planning to apply it.
- There are many ethical concerns you need to be aware of. AI can be used to infringe on copyright, or to cheat, or to steal the work of others, or to manipulate. And how a particular AI model is built and who benefits from its use are often complex issues, and not particularly clear at this stage. Ultimately, you are responsible for using these tools in an ethical manner.

We are in the early days of a very rapidly advancing revolution. Are there other uses you want to share? Let me know in the comments.



This post is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).



Type your email...	Subscribe
--------------------	-----------

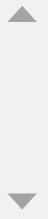


542 Likes · 47 Restacks

61 Comments



Write a comment...



Ross Stevenson Writes The Career Game Jul 15

I really like this part: "AI is a tool. It is not always the right tool. Consider carefully whether, given its weaknesses, it is right for the purpose to which you are planning to apply it."

On the video front, have you tried Descript @Ethan? I've delved into this recently and find it's AI features very helpful in my daily flow.

LIKE (13) REPLY ...

4 replies



Tim Brunelle Writes Curiosity+Courage Jul 15

This is brilliant. Thank you for the cogent utility! As you note: "Documentation-by-rumor is a weird choice for organizations claiming to be concerned about proper use of their technologies, but here we are." I hope Altman, et al remember to send you a check.

LIKE (4) REPLY ...

59 more comments...

© 2023 Ethan Mollick · [Privacy](#) · [Terms](#) · [Collection notice](#)
[Substack](#) is the home for great writing