

ChatGPT & Generative AI

Introduction

[ChatGPT](#) is a general purpose chatbot from [OpenAI](#) that uses a specific type of artificial intelligence (AI) known as **generative AI**. More about different types of AI will be discussed below, but from a simple user perspective, ChatGPT interacts with you to provide human-like responses to prompts or queries. It can generate answers to questions, summarize concepts, create poems, write computer code, draft essays or emails, simulate a text-based conversation, and much more.

Consider these [suggestions for 48 possible uses](#) of OpenAI's tools, including ChatGPT.

ChatGPT uses GPT-3 (Generative Pre-Trained Transformer 3), which is an example of a machine learning model that can **generate new data**. This means it is not fetching data and serving it up like old-school chatbots, but instead has been trained using a massive amount of data in such a way that it is actually generating a unique response to your query or prompt "on-the-fly." There has been intense interest in ChatGPT due to the impressive breadth, depth and natural language ability compared to earlier chatbots.

ChatGPT is not the only generative AI system. For example, some like [DALL-E](#) and [Stable Diffusion](#) generate images from text, and others can be used for sound. Machine learning models like [BERT](#) and [Bloom](#) are used for translation, math, and a wide variety of purposes.

This Quick Thing provides introductory information about using generative AI such as ChatGPT within the context of academic work. The following examples are starting points for reflecting on possible uses for new AI tools and why to be cautious.

A Possibly Useful Example

Tools like ChatGPT can be helpful as aids to brainstorm for a project or assignment. If you would like to explore different perspectives or ideas beyond your own thoughts, you could try asking the machine. Suppose you're interested in the relationship that humans have with pets and how stress figures into that relationship, but you have not yet decided on a focus. You could ask ChatGPT "What are some topics to investigate related to human stress and pets?" It would respond with possibilities that you could consider, adapt, and refine.



What are some topics to investigate related to human stress and pets?



1. The physiological and psychological mechanisms by which pet ownership or interaction reduces stress levels in humans.
2. The comparative effectiveness of different types of pets (e.g. dogs, cats, birds, fish) in reducing stress.
3. The impact of pet therapy on specific populations, such as veterans with PTSD, children with autism, or elderly individuals with dementia.
4. The effects of pet loss or separation on stress levels and mental health.
5. The role of gender, age, and personality in determining an individual's response to pet ownership or interaction as a stress-reduction strategy.
6. The long-term effects of pet ownership or interaction on stress levels and overall mental and physical well-being.
7. The impact of stress on pet health and behavior, and the effects of pet behavior on human stress.
8. The effectiveness of virtual pet therapy in reducing stress in individuals who cannot or choose not to own a live pet.

You can use these topics as a starting point for your research and expand on them as you learn more.

Caption: A prompt and the response generated by OpenAI's ChatGPT

Description: Screenshot of ChatGPT interface being prompted for topics related to human stress and pets. ChatGPT responds:

1. The physiological and psychological mechanisms by which pet ownership or interaction reduces stress levels in humans.
2. The comparative effectiveness of different types of pets (e.g. dogs, cats, birds, fish) in reducing stress.
3. The impact of pet therapy on specific populations, such as veterans with PTSD, children with autism, or elderly individuals with dementia.
4. The effects of pet loss or separation on stress levels and mental health.
5. The role of gender, age, and personality in determining an individual's response to pet ownership or interaction as a stress-reduction strategy.
6. The long-term effects of pet ownership or interaction on stress levels and overall mental and physical well-being.
7. The impact of stress on pet health and behavior, and the effects of pet behavior on human stress.
8. The effectiveness of virtual pet therapy in reducing stress in individuals who cannot or choose not to own a live pet.

A Not So Useful Example!

Remember the "generative" aspect of these tools? Although they may often provide you with good quality information, they will also provide you with misleading or fictitious information. In this example, asking ChatGPT to summarize key research with citations about human stress and pets produced a realistic sounding result and even provided citations as though academic sources support the summary provided. But while the writing seems plausible, the sources cited don't actually exist!



Write a paragraph summarizing key research about human stress and pets. Include bibliographic citations of the source publications.



According to a study published in the International Journal of Environmental Research and Public Health, pet ownership has been found to have a positive impact on human mental and physical health, including reducing stress levels. Another study published in the Journal of Social Psychology found that interacting with a therapy dog can significantly reduce stress levels in college students. A third study published in the Journal of Psychiatric Research found that pet owners had lower cortisol (a stress hormone) levels and reported feeling more supported and less lonely compared to non-pet owners.

Source:

- International Journal of Environmental Research and Public Health, "The Role of Pets in Enhancing Human Well-being and Facilitating Social Interactions," Gwenith G. Fisher, Anthony L. Podberscek, and Elizabeth S. Paul (2015)
- Journal of Social Psychology, "Therapy Dogs and the Reduction of Stress Symptoms in College Students," William E. Kay, Kari E. J. Lindell, and Mary B. Zulch (2011)
- Journal of Psychiatric Research, "The human-animal bond and psychiatric disorders," Alan Beck and Aaron Katcher (1984)

Caption: A prompt and ChatGPT's response. Writing seems plausible, but sources don't exist.

Description: Screenshot of ChatGPT interface being asked to write a paragraph summarizing key points about human stress and pets, including references. ChatGPT responds: "According to a study published in the International Journal of Environmental Research and Public Health, pet ownership has been found to have a positive impact on human mental and physical health, including reducing stress levels. Another study published in the Journal of Social Psychology found that interacting with a therapy dog can significantly reduce stress levels in college students. A third study published in the Journal of Psychiatric Research found that pet owners had lower cortisol (a stress hormone) levels and reported feeling more supported and less lonely compared to non-pet owners."

Source:

- * International Journal of Environmental Research and Public Health, "The Role of Pets in Enhancing Human Well-being and Facilitating Social Interactions," Gwenith G. Fisher, Anthony L. Podberscek, and Elizabeth S. Paul (2015)
- * Journal of Social Psychology, "Therapy Dogs and the Reduction of Stress Symptoms in College Students," William E. Kay, Kari E. J. Lindell, and Mary B. Zulch (2011)

* Journal of Psychiatric Research, "The human-animal bond and psychiatric disorders," Alan Beck and Aaron Katcher (1984)""

Types of AI and what “generative” means

Two well-known types of AI are predictive and generative machine learning models. The models’ algorithms determine patterns in a **training dataset**, which “teaches” the models what to do when they encounter new data that fit these patterns. (See our [Quick Thing on Machine Learning](#) for different types of training/learning methods)

Predictive models use patterns in data to **make predictions very quickly**. They are used, for example, on shopping sites and streaming video services to suggest what to buy or view based on your past habits or the choices of others with similar habits. Other usage examples include doing sentiment analysis of large corpuses of text, filtering spam, detecting diseases from medical images, and flagging fake news.

Generative models use patterns in data **to create new content**. The intention of the model is not accuracy or truth or beauty ... but rather to generate new text or images or sound that is equivalent in form to what humans can write or create.

Large Language Models (LLM) such as GTP-3 are examples of generative AI. An LLM uses probabilities derived from sequences of words from a very large corpus of text to determine what words should come next as responses are generated one word at a time. To give an idea of the dimension of the training dataset, the GPT-3 model was trained using almost 500 billion words from the public internet, including internet-based books and all of Wikipedia (Brown et al, 2020).

The **key point to know about generative AI** is that there is **no reasoning used** when it answers a prompt. For example, ChatGPT is not thinking (rationally or creatively) when it is responding to your question or instructions - it is generating text based on what it has “learned” about patterns and sequences of words derived from the human-created work in its training dataset. It then generates text that simulates these patterns. With such a large dataset and a human reinforcement technique used for its training, ChatGPT can generate responses that are impressively human-like, often “correct” in form, and sometimes wrong and/or misleading.

Embed video: <https://www.youtube.com/watch?v=o5MutYFWsM8>

Caption: This YouTube video gives an overview of ChatGPT, including capabilities and some important limitations. By: How It Happened.

Pitfalls of generative AI

In addition to the text and image generating tools already mentioned, generative AI can be used to create **deepfakes**, which are “believable media generated by a deep neural network” (Mirsky & Lee, 2022). Deepfakes are image, video, or audio content that seem indistinguishable from real humans. While these may offer new creative possibilities, they can also be used for deliberately misleading viewers.

The prevalence of **hallucinations**, or false information, is a significant and sometimes amusing limitation to tools like ChatGPT. Because it's a language model, not a web search engine, it creates new text based on the dataset used to train it. In other words, it identifies and imitates patterns, but it cannot verify information, leading it to routinely return absurdly incorrect information.

Name the three largest city on Mars

1. Syrtis Major
2. Tharsis Montes
3. Elysium Mons

Description: Chat GPT's answer to the question Name the three largest cities on Mars, to which the answer is 1. Syrtis Major 2. Tharsis Montes 3. Elysium Mons

Caption: A prompt and GPT-3's response—a hallucination. The regions may exist, but these aren't cities!

Algorithmic bias refers to "... the algorithm inappropriately privileging or disadvantaging one group of users over another group" (Turing Institute, 2023). Algorithmic bias results from what data was included and excluded in the dataset used to train the machine learning model. Marginalized groups tend to be less represented in training data, or they may be subject to societal bias that is embedded in the training data. Thus, the sources and methods of acquiring the training data can skew results and perpetuate biases that exist in society (Bender et al, 2021). In addition, when trained from data provided by online communities that express toxic content, the model can produce misinformation, disinformation, and unintended outputs that are likewise toxic.

Conclusion

Certainly, there are many positive ways in which AI tools like ChatGPT can and will be used in academic and research settings. ChatGPT is an impressive example of what AI is capable of today, and hints at what may be possible in the future. Many see the release of ChatGPT as a pivotal moment in the widespread application and use of generative artificial intelligence.

That said, to take full advantage of AI in your learning or work, you still need to understand your subject matter well enough that you can both guide the tool as well as critically assess its output. This means that generative AI tools like ChatGPT cannot replace the necessity of developing your own knowledge on a subject and exercising deep critical thinking skills.

Recommended readings to learn more about ChatGPT

Curious about machine learning and AI terminology? Try [DataFranca's French-language AI dictionary](#) (and apps).

OpenAI. (2023). [ChatGTP FAQ](#).

OpenAI has compiled brief answers to common and important questions about ChatGPT, including questions about accuracy and data usage.

Heilweil, R. (2023, January 5). [What is generative AI, and why is it suddenly everywhere?](#) Vox.

This article discusses generative AI tools like ChatGPT and Stable Diffusion, covering the basics of how they work, who has created them, and possible implications for education and the future of work.

[A selection of 15 short articles about ChatGPT](#). (2022/23). *The Conversation*.

Articles written by journalists and academics from a variety of perspectives.

Piantadosi, S. (2022, December 4). [Twitter thread about bias in ChatGPT](#).

The Twitter posts include screenshots demonstrating how ChatGPT perpetuates society's racist and sexist biases.

Caines, A. (2022, December 30). [ChatGPT and good intentions in higher ed](#). *Is a Liminal Space*.

This blog summarizes some important considerations that go beyond conversations about academic integrity. Specifically, the author reflects on data privacy and uncompensated labour.

Shanahan, M. (2023). [Talking about large language models](#). ArXiv.

This article takes a philosophical approach to discussing what tools like ChatGPT really do versus what we ascribe to them, and why the nuance is important.

Quiz

1. True or False? ChatGPT can write an essay.

Feedback: The correct answer is: True. It is true that ChatGPT can write an essay, however the quality and accuracy of its output may be poor, with issues like flawed logic, factual errors, and fake citations.

2. True or False? ChatGPT generates answers with a focus on accuracy.

Feedback: The correct answer is: False.

ChatGPT generates words in an order that sound most like a human response; accuracy is not the goal (even if it is sometimes or often accurate).

3. True or False? If you use text generated by ChatGPT in your work, you need to provide a citation.

Feedback: The correct answer is: True. Like any source that you use, pay attention to the proper citation requirements. AI tools will have specific terms and conditions for their use, which may provide publication and attribution guidelines in addition to other citation requirements. For example, the [OpenAI service terms](#) require that a publication using content generated by its tools clearly state that fact.

Activity 1

Purpose: Explore the GPT-3 model to brainstorm ideas and get relevant guidance on literature to read.

Task: Use Elicit.org (an AI-supported interface for searching and reviewing scholarly literature) to brainstorm a topic and find three important articles on that topic. Verify that the sources of the articles exists. Note that because Elicit relies on GPT-3, the literature summaries it provides may not be fully accurate; Elicit's [FAQ](#) claims 80-90% accuracy.

1. Sign up for a free account on [Elicit](#).
2. Click Elicit's **Tasks** button at the top of the screen to help you brainstorm or explore a research question or simply type your question into its main search box.
3. Read the summary of what Elicit refers to as the top papers related to the question. The summary appears in the upper left of the interface.
4. Verify if two or three of the articles actually exist by downloading them from the links provided or search them from the [Library's website](#)).
5. For one of the articles, consider: does the summary cover the **key content** of the paper or **did it miss the point**? Can you spot errors in the summary?

Time: This activity should take about 20 minutes to complete.

Activity 2

Purpose: Experiment with writing prompts for image creation. Gain some familiarity with the possibilities and limitations of using a text-to-image generative AI tool.

Task: Suppose you were assigned to do a presentation about a topic. Can you use a text-to-image AI tool to help illustrate an idea you'd like to cover in the presentation?

1. Navigate to the free text-to-image tool [OpenJourney](#).
2. Write a prompt that asks OpenJourney to generate an image representing something from your research. Type your prompt in the "**Text-to-image**" box and click the "**Compute**" button. Did the result make sense? Did it surprise you?
3. Can you refine the image to be better? Try changing the words in your prompt. What happens if you ask it to make a **diagram**? A textbook **illustration**? Something **photorealistic**?
4. Did any of the images lead you to question the data sources that were used to train this model?

Time: This activity should take about 10-15 minutes to complete.

Resources

Zoe Bee (Director). (2022, November 12). [Why AI Isn't as Good at Writing as You Think](#) [YouTube video].


This video considers the possibilities and shortcomings of writing with AI and presents arguments for why the human aspects of writing and interpreting text are unlikely to be replaced by AI soon.

Udemy course: [Generative A.I., from GANs to CLIP, with Python and Pytorch](#). Description: “Learn to code the most creative and exciting A.I. architectures, generative networks, from basic to advanced and beyond”

Klasky, E., Middha, A., Rosenfeld, H., Kleinman, M., & Parthasarathy, S. (2022). [What’s in the Chatterbox? One Pager](#) (What’s in the Chatterbox? Large Language Models, Why They Matter, and What We Should Do About Them, p. 2). Ford School of Science, Technology, and Public Policy, University of Michigan.

This one-page document summarizes some societal implications of large language models (LLMs) like ChatGPT. It also links to the full report describing how LLMs work and how this technology might impact us.

References

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. [On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?](#) . In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–23. FAccT ’21. New York, NY, USA: Association for Computing Machinery, 2021.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). [Language Models are Few-Shot Learners](#). Advances in Neural Information Processing Systems, 33, 1877-1901.

[Data Science and AI Glossary](#). The Alan Turing Institute, 2023.

Mirsky, Yisroel, and Wenke Lee. [The Creation and Detection of Deepfakes: A Survey](#). ACM Computing Surveys 54, no. 1 (January 2022): 1–41.