

5 things to know about generative text AI tools... that might be outdated or upgraded by the time of publication

Natalia Tingle Dolan

Associate Professor and Business Collections & Reference Librarian, William M. White Business Library, University of Colorado Boulder, natalia.dolan@colorado.edu

Jennifer Freer

Liaison/Librarian, RIT Libraries, Rochester Institute of Technology, jlfwml@rit.edu

What are they and how to access them:

Generative text AI tools such as OpenAI's ChatGPT, Google's Bard, and Microsoft's Bing chatbot use machine learning algorithms to generate text based on input prompts. ChatGPT and Bing's new chatbot use the GPT language model, while Bard uses Google's LaMDA model. Some tools require a paid account, while others offer free access to beta versions. You can go pretty deep learning about LLMs (large language models), machine learning, NLP (natural language processing) and how the chat software can create an interactive, conversational search experience, but you can also start using these tools without any special knowledge. Signing up for an account on any of these services is the best way to get started. You can ask the tool to tell you some of the things that you can do with it, or you can try some of the following ideas. The OpenAI playground is a great place to experiment with different models and settings.

Cool stuff to try:

At the time of this writing, the generative text AI tools are remarkably strong at content creation. You can paste in your own text and ask it to punch up a presentation or make a wordy rough draft more cohesive. The tools can be useful for reference and teaching librarians in various ways, such as listing keywords for a topic, creating a lesson plan, drafting learning objectives, brainstorming classroom activities, summarizing content into bullet points or slide titles, or explaining a topic in simple terms. They can create a project schedule and include each component that must be completed by a given deadline. ChatGPT-4 can draft a decent business plan and provide a related elevator pitch without anything beyond a one-sentence prompt. Explore each tool's capabilities and limitations to determine how best to use them. If your prompt does not generate the result you expect, follow up with more instruction. Some tools, like the free version of ChatGPT, are not able to access the internet to search for current information, so they are not good at fact-checking or providing up-to-date sources for a reference question or a literature review. At the time of this writing, Google's Bard and Microsoft's Bing via the Edge browser do have live access to the internet, but, as generative text tools, they are still currently susceptible to hallucinations (Goode, 2023).

Sometimes the tools hallucinate:

One of the most significant issues with these tools is that they can hallucinate or invent realistic sounding text and citations for information. In some cases, they will even make up works that do

not exist. ChatGPT is gaining a reputation for these hallucinations (Smith, 2023) which have also referred to as “hallucinations” (Crawford, 2023), and this behavior may push potential users away. In one example, ChatGPT 3.5 suggested specific reports from Frost & Sullivan and IBISWorld in response to a prompt. Neither report exists, but both titles sounded like they could be real reports. Similarly, prompting the chatbot for a list of recent articles on your topic will likely result in at least one reference that does not exist. The journal in the citation is real, the authors may exist (and may even publish on that topic), but the article itself does not exist. The generated information is very convincing as well as unexpected. The hallucinated results will impact both librarians and patrons, making the mission of teaching critical thinking and information literacy skills even more important.

The caveats:

In addition to creative hallucinations, there are several caveats to keep in mind when using generative text AI tools. The output is impacted by the quality of the prompt, the corpus the model was trained on, and what the tool can access. There are misconceptions about what ChatGPT generates. It was trained to generate predictive text that sounds human, not to retrieve an accurate fact from a website like search engines do. This can be confusing because the text it was trained on is what we think of as information such as Wikipedia, Google Books; instructional manuals, journal articles, etc., (Halpern, 2023) but the text it generates is not always factual. The free version of ChatGPT “has limited knowledge of [the] world and events after 2021” (Natalie, 2023). This is because the dataset that this tool (and others using GPT) was trained on is a static dataset which ends in September 2021. In addition to this limitation, it may also be that the data or datasets on which these models are based have an English language or Western bias (Schaul, Chen, & Tiku, 2023). For a deep dive into the content many GPT tools were trained on, it is worth reading the in-depth analysis offered by *The Washington Post* on their website on April 19, 2023, through an interactive article called “Inside the secret list of websites that make AI like ChatGPT sound smart” (Schaul, Chen, & Tiku, 2023).

These tools are constantly changing. At the time of this writing, Bard is still in beta and there are rumors that it could be replaced before it is even officially released. Do not get too attached to any particular service or interface. There are dozens of startups releasing new products built on the OpenAI API every week. These are still early days, so it is uncertain which of these startups and tools will go the distance.

Be cognizant of what information you input with your prompt, especially personal or proprietary information. Your prompts may be used for training, and they cannot be deleted (Natalie, 2023). They may also be ingested into future iterations, as Samsung discovered (Dreibelbis, 2023). Copyright, ethical concerns, privacy, and misinformation are additional serious factors to consider when using these tools but are far beyond the scope of this article.

Ways to learn:

Google, OpenAI, and Microsoft provide FAQs and free online courses for their tools, making it easy for users to learn and gain foundational knowledge. Additionally, online learning platforms

such as LinkedIn Learning and Coursera offer courses. The University of South Florida offers a free online course with the option to purchase a digital badge upon completion. As the generative text AI tools rapidly evolve, YouTube and social media are likely to have the most cutting-edge use-case examples. Currently much of the social media content around these tools is from marketers, but there are also many students and educators sharing how they use these tools. For example, on the YouTube channel Science Grad School Coach, Dr. Alana Rister shares a variety of ways students might use many AI powered tools including ChatGPT. Her video titled “Using ChatGPT for Scientific Writing: The Dos and Don'ts” is a nice example of how one can use ChatGPT that also includes many cautions. (Science, 2023). Finally, the Congressional Research Service has a recent report on copyright and AI (Zirpoli, 2023).

During the writing of this article, we plugged an early outline into ChatGPT-4 and asked what else should be added. Figure 1 below shows the beginning of the conversation followed in our text by the full list.

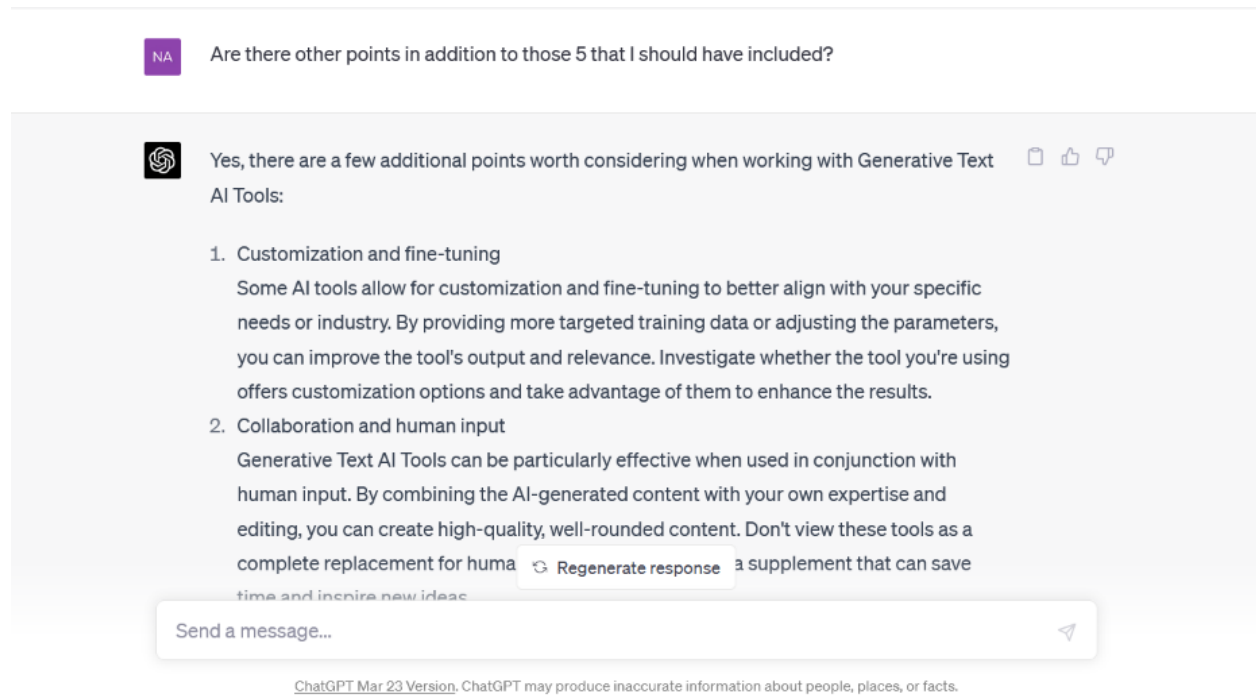


Figure 1. ChatGPT Conversation

Below is the prompt and full list ChatGPT generated (OpenAI, 2023):

Natalia: Are there other points in addition to those 5 that I should have included?

ChatGPT-4: Yes, there are a few additional points worth considering when working with Generative Text AI Tools:

6. Customization and fine-tuning

Some AI tools allow for customization and fine-tuning to better align with your specific needs or industry. By providing more targeted training data or adjusting the parameters, you can improve the tool's output and relevance. Investigate whether the tool you're using offers customization options and take advantage of them to enhance the results.

7. Collaboration and human input

Generative Text AI Tools can be particularly effective when used in conjunction with human input. By combining the AI-generated content with your own expertise and editing, you can create high-quality, well-rounded content. Don't view these tools as a complete replacement for human creativity, but rather as a supplement that can save time and inspire new ideas.

8. Ethical use and transparency

When using AI-generated content, it's crucial to be transparent about its origins. If you're using AI-generated text in articles, educational materials, or marketing materials, consider disclosing that the content was produced with the help of AI. This maintains trust and ensures that readers are aware of the nature of the content they're consuming.

9. Regular updates and maintenance

As AI technologies advance rapidly, the tools you're using might become outdated or less effective over time. Stay informed about new releases, updates, and improvements to ensure that you're using the most up-to-date and efficient version of your chosen AI tool.

10. Evaluating cost and ROI

While some AI tools are available for free, others require a subscription or per-use fee. When deciding on which tool to use, consider the cost of the tool and the potential return on investment (ROI). Evaluate whether the time saved and the quality of the generated content justify the expense. If the tool helps streamline your workflow and produces valuable results, it may be a worthwhile investment.

By considering these additional points, you can make a more informed decision when working with Generative Text AI Tools and maximize their potential benefits.

Conclusion

Many of us already use text completion/predictive text through tools like Gmail, Outlook, text messaging, Google search, etc., but these new GPT products feel a little bigger in scale and scope. Artificial Intelligence has been under development for almost 80 years (Asaro, 2005) and

yet with the very sudden emergence of ChatGPT, Bing, and Bard there is a sensation that a leap has been made which will impact librarians and information professionals in many ways. Having an awareness or working knowledge of these tools as they evolve, or possibly fade, can help inform decisions whether one is working with a patron or shaping one's own workflow.

References

- Asaro, P. M. (2005). Artificial intelligence. In C. Hempstead, & W. E. Worthington (Eds.), *Encyclopedia of 20th century technology*. Routledge. Credo Reference: https://ezproxy.rit.edu/login?url=https://search.credoreference.com/content/entry/routt/artificial_intelligence/0?institutionId=3255
- Crawford, K. [@katecrawford]. (2023, April 4). Guess we could start calling this a 'hallucination'? [Tweet]. Twitter. <https://twitter.com/katecrawford/status/1643323088644235266>
- Dreibelbis, E. (2023, April 7). Samsung software engineers busted for pasting proprietary code into ChatGPT. PCMag. <https://www.pcmag.com/news/samsung-software-engineers-busted-for-pasting-proprietary-code-into-chatgpt>
- Goode, L. (2023, March 30). Review: We put ChatGPT, Bing Chat, and Bard to the test. *Wired*. <https://www.wired.com/story/review-ai-chatbots-bing-bard-chat-gpt/>
- Halpern, S. (2023, March 28). What we still don't know about how A.I. is trained. *The New Yorker*. <https://www.newyorker.com/news/daily-comment/what-we-still-dont-know-about-how-ai-is-trained>
- Natalie. (2023). *What is ChatGPT? Commonly asked questions about ChatGPT*. OpenAI. <https://help.openai.com/en/articles/6783457-what-is-chatgpt>
- OpenAI. (2023). ChatGPT (Mar 23 version) [Large language model]. <https://chat.openai.com/chat>
- Science Grad School Coach. (2023, March 8). *Watch this before using ChatGPT for scientific research: Dos and don't of ChatGPT* [Video]. YouTube. <https://youtu.be/TwnQGNdiAnk>
- Schaul, K., Chen, S.Y., and Tiku, N. (2023, April 19). Inside the secret list of websites that make AI like ChatGPT sound smart. *The Washington Post*. <https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning>
- Smith, C. S. (2023, March 13). Hallucinations could blunt ChatGPT's success. *IEEE Spectrum*. <https://spectrum.ieee.org/ai-hallucination#toggle-gdpr>
- Zirpoli, C.T. (2023, February 24). Generative artificial intelligence and copyright law. (Legal Sidebar No. LSB10922). *Congressional Research Service*. <https://crsreports.congress.gov/product/pdf/LSB/LSB10922>