Chatting about ChatGPT: How may AI and GPT impact academia and libraries?

Brady D. Lund and Ting Wang

January 2023

This is a pre-print of a manuscript that has been accepted for publication by

Library Hi Tech News.

Abstract

This paper provides an overview of key definitions related to ChatGPT, a public tool developed by OpenAI, and its underlying technology, GPT. The paper discusses the history and technology of GPT, including its generative pre-trained transformer model, its ability to perform a wide range of language-based tasks, and how ChatGPT utilizes this technology to function as a sophisticated chatbot. Additionally, the paper includes an interview with ChatGPT on its potential impact on academia and libraries. The interview discusses the benefits of ChatGPT such as improving search and discovery, reference and information services, cataloging and metadata generation, and content creation, as well as the ethical considerations that need to be taken into account, such as privacy and bias. The paper also explores the possibility of using ChatGPT for writing scholarly papers.

Keywords: ChatGPT, GPT-3, Generative Pre-Trained Transformer, AI, Academia, Libraries

This paper provides an in-depth examination of the cutting-edge technology behind ChatGPT, a highly sophisticated chatbot that has gained significant attention in recent months. The paper is divided into three parts. The first part provides definitions of some key concepts related to ChatGPT, such as natural language processing and artificial intelligence, and explains how they play a role in the technology. The second part delves into the history, technology, and capabilities of Generative Pre-Trained Transformer (GPT), the underlying technology of ChatGPT. It explains the concepts behind GPT, the process of its development, the scale of the program and the vast amount of data used to train it, and its ability to perform a wide range of language-based tasks such as translation, question answering, and text generation. The third part of the paper gives an example of ChatGPT's abilities by providing the output of an interview with ChatGPT on the topic of how AI and GPT will impact academia and libraries. This section will explore how ChatGPT can be used to improve various library services and the ethical considerations that need to be taken into account when using it.

Some Key Concepts Related to ChatGPT

Attention Mechanism: An attention mechanism is used in neural networks to allow a model to focus only on specific aspects of input data when making predictions (Niu et al., 2021).

Chatbot: A chatbot is a computer program designed to simulate conversation with human users, especially over the Internet (King, 2022).

Generative Model: A generative model is a type of model that generates new data, as opposed to only classifying or predicting based on input data (Pavlik, 2023).

Generative Pre-Trained Transformer: Generative Pre-Trained Transformer (GPT) is a machine learning model that uses unsupervised and supervised learning techniques to understand and generate human-like language (Radford et al., 2018).

Language Model: A language model is a type of artificial intelligence model is trained to generate text that is similar to human language (MacNeil et al., 2022).

Multimodal Neurons: Multimodal neurons are artificial neural network units that are able to understand and interpret the form of an object across different modes or representations, such as images, text, and speech (Goh et al., 2021).

Natural Language Processing: Natural language processing (NLP) is a field of artificial intelligence that involves using algorithms to analyze and interpret human language, such as text and speech, in order to extract meaning and extract useful information (Manning & Schutze, 1999).

Neural Network: A neural network is a machine learning model composed of interconnected processing nodes, that is trained on data to perform specific task by adjusting the strengths of connections between them (Bishop, 1994).

Supervised Fine-Tuning: Supervised fine-tuning is a machine learning technique in which a pre-trained model is further trained on a smaller, labeled dataset in order to improve its performance on a specific task (Lee et al., 2018).

Transfer Learning: Transfer learning is the capability of tools like ChatGPT to use knowledge gained from one task in order to improve its performance on another, related task (Pan & Yang, 2010).

Unsupervised Pre-Training: Unsupervised pre-training is a machine learning technique in which a model is trained on a large dataset without any labeled examples, allowing it to learn the underlying structure and patterns in the data (Lee et al., 2018).

GPT and ChatGPT

ChatGPT is a public tool developed by OpenAI that is based on the GPT language model technology (Kirmani, 2022). It is a highly sophisticated chatbot that is capable of fulfilling a wide range of text-based requests, including answering simple questions and completing more advanced tasks such as generating thank you letters and guiding individuals through tough discussions about productivity issues (Liu et al., 2021). ChatGPT is able to do this by leveraging its extensive data stores and efficient design to understand and interpret user requests, and then generating appropriate responses in nearly natural human language. In addition to its practical applications, ChatGPT's ability to generate human-like language and complete complex tasks makes it a significant innovation in the field of natural language processing and artificial intelligence. In this brief review paper, the details of how ChatGPT works and the potential impacts of this technology on various industries are discussed.

OpenAI is a research laboratory founded in 2015 (Brockman et al., 2016). This laboratory has made rapid progress in the development of AI technologies and has released a number of machine learning products for the general public, including DALL-E and ChatGPT (Devlin et al., 2018). DALL-E, which uses a combination of machine learning technologies to generate novel images based on user inputs, gained extensive public attention in early 2022 (Marcus et al., 2022). Its ability to understand user requests through NLP principles, similar to those used in ChatGPT, and to create images using artificial neural networks with multimodal neurons allows it to produce a wide range of novel images (Cherian et al., 2022; Goh et al., 2021). DALL-E's availability to the public has also contributed to the rapid popularity of ChatGPT, which achieved over one million unique users within one week of its launch (Mollman, 2022).

Generative Pre-Trained Transformer (GPT) is a language model developed by OpenAI that is capable of producing response text that is nearly indistinguishable from natural human language (Dale, 2021). The concepts behind GPT are refined through a two-step process: generative, unsupervised pretraining using unlabeled data and discriminative, supervised fine-tuning to improve performance on specific tasks (Erhan et al., 2010; Budzianowski & Vulić, 2019). During the pretraining phase, the model learns naturally, similar to how a person might learn in a new environment, while the fine-tuning phase involves more guided and structured refinement by the creators (Radford et al., 2018).

GPT-3 and ChatGPT, along with other models like BERT, RoBERTa, and XLNet are all state-of-the-art language models developed by OpenAI (GPT), Google (BERT), and Microsoft (XLNet). GPT-3 and ChatGPT are both based on the GPT-3 architecture and have the ability to generate human-like text, making them useful for a variety of natural language processing tasks

such as language translation, summarization, and question answering. BERT, RoBERTa, and XLNet, on the other hand, are primarily focused on understanding the underlying meaning of text and are particularly useful for tasks such as sentiment analysis and named entity recognition. One of the key benefits of GPT-3 and ChatGPT is their ability to generate high-quality text, while BERT, RoBERTa, and XLNet excel at understanding and analyzing text.

Developed by OpenAI, ChatGPT is a public tool that utilizes GPT technology. As a sophisticated chatbot, it is able to fulfill a wide range of text-based requests, including answering simple questions and completing more advanced tasks such as generating thank you letters and addressing productivity issues. It is even capable of writing entire scholarly essays by breaking a main topic into subtopics and having GPT write each section, it is possible to create an entire article using the tool. With a full version that allows for longer responses, it is even possible to write an entire paper in a matter of seconds with minimal input from a researcher. In addition to its potential impact on the writing profession, ChatGPT could also have significant consequences for a range of other industries. Its natural language processing capabilities make it an ideal tool for handling basic customer service inquiries, such as the "ask me" feature on websites. Its ability to analyze and interpret large amounts of text could also make it valuable in the legal profession, potentially assisting with research and document preparation tasks. Additionally, ChatGPT's ability to provide oversight on the quality of written work could be useful in the field of education, potentially helping to grade and provide feedback on student assignments.

GPT technology is a powerful tool for natural language processing tasks, but it does have its limitations. One of the main limitations is that GPT models are based on a statistical approach that learns patterns from a large dataset of text, which can perpetuate biases and stereotypes present in the data (Dale, 2017; Lucy & Bamman, 2021). This means that the model may generate offensive or harmful output. Additionally, GPT models are not able to fully understand the context and meaning of the text they generate and they are not able to perform well in tasks that require common sense reasoning or logical reasoning which is not covered in the training data (Strubell et al., 2019). Furthermore, GPT models are computationally expensive to train and require large amounts of data and computational resources, making them difficult to implement for some organizations and individuals. Additionally, operating these algorithms and data stores at the scale that OpenAI does requires a significant amount of energy (Zhou et al., 2021). Therefore, it is important to be aware of these limitations and to use GPT responsibly.

Four questions with ChatGPT About Its Impact on Academia and Libraries

In this section, each question, labelled with “Q:” is written by a human interviewer, and each response, labelled with “A:” is written entirely by ChatGPT. Credit for these responses goes to the ChatGPT tool, which can be accessed through https://chat.openai.com/chat.

Q: How can ChatGPT be used to improve research and scholarship in academia?

A: ChatGPT can be used to improve research and scholarship in academia in a number of ways:

1. Literature review assistance: ChatGPT can assist researchers in identifying relevant literature by generating summaries of articles or providing a list of relevant papers based on a given topic or keyword.
2. Text generation: ChatGPT can generate text in a specific style or tone, allowing researchers to easily generate draft versions of research papers, grant proposals, and other written materials.
3. Data analysis: ChatGPT can assist researchers in analyzing large amounts of text data, such as social media posts or news articles, by providing insights and identifying patterns in the data.
4. Language Translation: ChatGPT can be used for machine translation, allowing researchers to access and understand research materials in multiple languages.
5. Automated summarization: ChatGPT can be used to automatically summarize scientific papers, reports, or other documents, making it easier for researchers to stay up-to-date with the latest developments in their field.
6. Question answering: ChatGPT can be fine-tuned to provide answer to domain specific questions, making it a powerful tool for scholars to find answers quickly and efficiently.

These capabilities can help researchers to save time and effort, allowing them to focus on the more creative and analytical aspects of their work.

Q: What are the ethical and privacy implications of using ChatGPT in academic settings?

A: There are several ethical and privacy implications to consider when using ChatGPT or other large language models in academic settings. One concern is the potential for bias in the model's responses, as it may reflect the biases present in the training data. Additionally, there are concerns about the privacy of individuals whose data was used to train the model, as well as the potential for the model to be used for nefarious purposes. It's also important to consider the ethical implications of creating highly realistic synthetic text or speech, as it could be used to impersonate or deceive others.

Additionally, it is important to consider the data privacy and security implications of using ChatGPT. The model has the ability to generate highly sensitive information, such as personal data, financial data, and even medical data.

In light of these concerns, it is important to use these models responsibly and with caution, and to consider appropriate measures to mitigate any potential risks.

Q: What are the implications of using ChatGPT for natural language processing activities in libraries?

A: Using ChatGPT for natural language processing activities in libraries can have several implications, both positive and negative. On the positive side, the model's ability to understand and generate human language can be used to improve a variety of library services, such as:

- Search and discovery: ChatGPT can be used to improve the accuracy and effectiveness of library search systems, by understanding the nuances of natural language queries and providing more relevant results.

- Reference and information services: ChatGPT can be used to provide automated reference and information services, such as answering common questions or providing information on library resources.
- Cataloging and metadata generation: ChatGPT can be used to automatically generate cataloging and metadata for library resources, such as book titles and descriptions.
- Content creation: The model can be used to generate summaries, abstracts, and other forms of content that can be used to improve access to library resources.

On the negative side, there are several ethical and privacy concerns to consider when using ChatGPT in libraries:

- Bias: The model's responses may reflect the biases present in the training data, which could result in inaccurate or unfair results.
- Privacy: The model's ability to generate highly realistic synthetic text or speech could be used to impersonate or deceive others, which would be a violation of user's privacy.
- Security: The model has the ability to generate highly sensitive information, such as personal data, financial data, and even medical data, which could be a security risk.

In light of these concerns, it is important to use ChatGPT responsibly and with caution in libraries, and to consider appropriate measures to mitigate any potential risks.

Q: What ethical issues should be considered when using ChatGPT within academic libraries?

A: When using ChatGPT within academic libraries, there are several ethical issues that should be considered, including:

- Bias: The model's responses may reflect the biases present in the training data, which could result in inaccurate or unfair results. This could disproportionately affect marginalized groups, and undermine the library's mission to provide equitable access to information.
- Privacy: The model's ability to generate highly realistic synthetic text or speech could be used to impersonate or deceive others, which would be a violation of user's privacy. Additionally, the model may be able to generate sensitive information, such as personal data, financial data, and even medical data, which should be protected and not shared without explicit consent.
- Autonomy and informed consent: Users should be informed of the use of ChatGPT and should have the ability to opt-out or give their consent to use their information.
- Transparency and accountability: There should be transparency in the design, training, and deployment of ChatGPT, including the data used to train the model, so that any potential biases or errors can be identified and addressed. Additionally, there should be clear policies and procedures in place to handle any issues that may arise.

- Intellectual property: ChatGPT can generate text which may be copyright protected. Therefore, it is important to ensure that the use of the model complies with intellectual property laws and regulations, such as obtaining permission from copyright holders when necessary.

These ethical issues are of paramount importance for academic libraries to consider in order to maintain their role as trusted providers of information and to ensure that the use of ChatGPT aligns with their mission and values.

Conclusion

In conclusion, this paper has provided an in-depth examination of the technology behind ChatGPT, a highly sophisticated chatbot that has gained significant attention in recent months. It began by defining key concepts related to ChatGPT and explaining how they play a role in the technology. The paper then delved into the history, technology, and capabilities of Generative Pre-Trained Transformer (GPT), the underlying technology of ChatGPT, and how it can perform a wide range of language-based tasks. The paper also provided an example of ChatGPT's abilities by giving the output of an interview with ChatGPT on the topic of how AI and GPT will impact academia and libraries. This section explored the benefits of ChatGPT such as improving search and discovery, reference and information services, cataloging and metadata generation, and content creation, as well as the ethical considerations that need to be taken into account, such as privacy and bias. As the paper has shown, ChatGPT has considerable power to advance academia and librariasnhip in both anxiety-provoking and exciting new ways. However, it is important to consider how to use this technology responsibly and ethically, and to uncover how we, as professionals, can work alongside this technology to improve our work, rather than to abuse it or allow it to abuse us in the race to create new scholarly knowledge and educate future professionals.

Reference

Bishop, C. M. (1994). Neural networks and their applications. *Review of Scientific Instruments, 65*, article 1803. https://doi.org/10.1063/1.1144830

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). Openai gym. arXiv. https://doi.org/10.48550/arXiv.1606.01540

Budzianowski, P., & Vulić, I. (2019). Hello, it's GPT-2--how can I help you? towards the use of pretrained language models for task-oriented dialogue systems. arXiv. https://doi.org/10.48550/arXiv.1907.05774

Cherian, A., Peng, K. C., Lohit, S., Smith, K., & Tenenbaum, J. B. (2022). Are Deep Neural Networks SMARTer than Second Graders?. arXiv. https://doi.org/10.48550/arXiv.2212.09993

Dale, R. (2017). NLP in a post-truth world. *Natural Language Engineering, 23*(2), 319-324.

Dale, R. (2021). GPT-3 What's it good for? *Natural Language Engineering, 27*(1), 113-118.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv. https://doi.org/10.48550/arXiv.1810.04805

Erhan, D., Bengio, Y., Courville, A., Manzagol, P., & Vincent, P. (2010). Why does unsupervised pre-training help deep learning. *Journal of Machine Learning Research, 11*, 625-660.

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines, 30*(4), 681-694.

Goh, G., Cammarata, N., Voss, C., Carter, S., Petrov, M., Schubert, L., Radford, A., & Olah, C. (2021). *Multimodal neurons in artificial neural networks*. Retrieved from https://doi.org/10.23915/distill.00030

King, M. R. (2022). The future of AI in medicine: A perspective from a chatbot. *Annals of Biomedical Engineering*. https://doi.org/10.1007/s10439-022-03121-w

Kirmani, A. R. (2022). Artificial intelligence-enabled science poetry. *ACS Energy Letters, 8*, 574-576.

Lee, C., Panda, P., Srinivasan, G., & Roy, K. (2018). Training deep spiking convolutional neural networks with STDP-based unsupervised pre-training followed by supervised fine-tuning. *Frontiers in Neuroscience, 12*, article 435.

Liu, X., Zheng, Y., Du, Z., Ding, M., Qian, Y., Yang, Z., & Tang, J. (2021). GPT understands, too. arXiv. https://doi.org/10.48550/arXiv.2103.10385

Lucy, L., & Bamman, D. (2021). Gender and representation bias in GPT-3 generated stories. *Proceedings of the Workshop on Narrative Understanding, 3*, 48-55.

MacNeil, S., Tran, A., Mogil, D., Bernstein, S., Ross, E., & Huang, Z. (2022). Generating diverse code explanations using the GPT-3 large language model. *Proceedings of the ACM Conference on International Computing Education Research, 2*, 37-39.

Manning, C., & Schutze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.

Marcus, G., Davis, E., & Aaronson, S. (2022). *A very preliminary analysis of DALL-E 2*. ArXiv pre-print. Retrieved from https://doi.org/10.48550/arXiv.2204.13807

Mollman, S. (2022). *ChatGPT gained 1 million users in under a week*. Retrieved from https://www.yahoo.com/lifestyle/chatgpt-gained-1-million-followers

Niu, Z., Zhong, G., & Yu, H. (2021). A review on the attention mechanism of deep learning. *Neurocomputing, 452*, 48-62.

OpenAI. (2022). OpenAI about page. Retrieved from https://openai.com/about/

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism and Mass Communication Educator*. https://doi.org/10.1177/10776958221149577

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training*. Retrieved from https://www.cs.ubc.ca/~amuham01/LING530/papers/radford2018improving.pdf

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *Proceedings of the Annual Meeting of the Association for Computational Linguistics, 57*, 3645-3650.

Zhou, X., Chen, Z., Jin, X., & Wang, W. Y. (2021). HULK: An energy efficiency benchmark platform for responsible natural language processing. *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations, 16*, 329-336.