V. P. Maidaniuk
O. O. Kavka

CHATGPT AS AN INSTRUMENT FOR A PRELIMINARY DATA GATHERING IN SCIENTIFIC RESEARCHES

Vinnytsia National Technical University

Анотація
Тези описують способи використання ChatGPT у якості інструментарію для попереднього збору даних для наукових досліджень.
Ключові слова: штучний інтелект, ChatGPT, наукове дослідження.
Abstract
The thesis elaborates on using ChatGPT as a toolset for preliminary data gathering in scientific researches.
Keywords: artificial intelligence, ChatGPT, scientific research.

Introduction

After ChatGPT [1] entered the public beta stage in 2022, the world was shaken by the enormous potential in the field of AI. Hundreds of new AI services rapidly reshape the whole world's economics, making it obligatory for almost every employee to embrace new technologies and tools. Scientists are no exception from this.

The Potential of ChatGPT

ChatGPT can be a valuable tool for researchers to gather data for their studies. Some possible use cases are:

- **Information retrieval**: Researchers can ask ChatGPT questions or request specific information on their research topics, and it will provide relevant answers based on its training data.
- **Exploratory discussions**: Researchers can have open-ended conversations with ChatGPT to explore ideas, gain fresh perspectives, or discover new research directions. ChatGPT's contextual responses can help generate innovative insights.
- **Preliminary data collection**: ChatGPT can assist in collecting initial data by answering questions or offering opinions on a subject. This can provide a starting point for further investigation or hypothesis development.
- **Literature review assistance**: ChatGPT can help researchers in their literature reviews by suggesting relevant papers, summarizing key findings, or providing insights into existing research. It can save time by offering relevant information and aiding in source identification.
- **Expertise in niche areas**: If ChatGPT has been trained on specialized data, it can provide expertise in specific domains. Researchers can tap into this knowledge to gain insights and gather data from a specialized perspective.
- **Data synthesis and analysis**: Researchers can input data or summaries into ChatGPT for analysis and synthesis. By receiving responses, researchers can identify patterns, generate hypotheses, or interpret the collected data.

Precautions

When utilizing ChatGPT for preliminary data gathering, it is crucial to consider several precautions. Here are some important points to bear in mind:

- **Data reliability and bias**: The responses provided by the language model are generated based on patterns and examples from its training data. Consequently, there is a possibility of inadvertent biases or inaccuracies. Exercise caution regarding potential prejudices, inaccuracies, or outdated data that might be present in the training material.
- **Fact-checking and validation**: While the language model can offer valuable insights, it is essential to fact-check and validate the information it provides [2]. Relying solely on the responses without independent verification can lead to misleading or incorrect conclusions. Verify the information obtained by cross-referencing with reliable sources or conducting further research to ensure accuracy.
- **Contextual understanding**: The language model might not fully comprehend the nuances or context of a particular research topic. It lacks real-world experiences and common-sense reasoning. Therefore, carefully consider the limitations of the model's understanding and critically evaluate its responses in the context of your research.
- **Subject expertise**: The language model's knowledge is derived from its training data, which covers a broad range of topics but may not encompass the most up-to-date or specialized knowledge in certain fields. Exercise caution when relying on the model for highly specialized or domain-specific information.
- **Limited source evaluation**: The language model does not possess the ability to evaluate the credibility or quality of its training sources. It is important to independently assess the reliability and trustworthiness of the information provided by the model.
- **Ethical considerations**: When interacting with the language model, ensure that the data collection process adheres to ethical guidelines. For instance, if engaging with human participants, obtain informed consent, and respect privacy and confidentiality.
- **Preliminary nature of data**: Recognize that the data collected through the language model should be treated as preliminary or exploratory. It should serve as an initial foundation for further research rather than being considered conclusive evidence. Validate and verify the data using rigorous scientific methods before drawing final conclusions.
- **Transparent documentation**: If you intend to employ the data collected through the language model in your research, transparently acknowledge the source and limitations of the data. Clearly document and attribute the use of the language model to ensure proper citation and provide context for readers.

By considering these precautions, researchers can make informed decisions about employing the language model for preliminary data gathering while mitigating potential risks or limitations associated with its use.

Conclusions

ChatGPT, as well as some other AI startups, is an extremely useful instrument, which should be used with precautions. It is recommended for individual researchers to learn how to use this tool, and how to use it properly. Scientific organizations, such as universities, should establish dedicated programs to assist their employees with embracing the new available tooling.

REFERENCES

1. *Introducing ChatGPT*. (n.d.). https://openai.com/blog/chatgpt
2. *What is ChatGPT?* (n.d.). https://help.openai.com/en/articles/6783457-what-is-chatgpt

Майданюк Володимир Павлович – д. т. н., доцент кафедри програмного забезпечення, Вінницький національний технічний університет, м. Вінниця, e-mail: maidaniuk@vntu.edu.ua

Кавка Олексій Олександрович – аспірант кафедри програмного забезпечення, Вінницький національний технічний університет, м. Вінниця, e-mail: oleksii.kavka@gmail.com

Maidaniuk Volodymyr P. – Doctor of Science (Tech.), Associate Professor of the Department of Software Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: maidaniuk@vntu.edu.ua;

Kavka Oleksii O. – Graduate Student of the Department of Software Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: oleksii.kavka@gmail.com.