Paper ET05

What Role Does ChatGPT Have in your Clinical Programming?

Andy Ratcliffe, Instem plc, London, England

ABSTRACT

This paper explores the potential role of AI bots and chat bots in the creation of statistical analyses for regulatory review. With recent advancements in AI technology, bots are capable of generating essays, blog posts, and promotional copy, raising the question of whether they can also contribute to the creation of statistical analyses. The paper considers the advantages and limitations of using AI bots in this context and examines the feasibility of their application in regulatory review. While bots have the ability to automate tasks and reduce human error, they may not possess the necessary contextual understanding and expertise required for creating accurate statistical analyses. This paper is relevant for anyone interested in the potential applications of AI bots in regulatory review processes.

INTRODUCTION

Artificial intelligence (AI) is the field of computer science that aims to create machines or software that can perform tasks that normally require human intelligence, such as reasoning, learning, planning, or decision making. The concept of AI dates back to the early 20th century, when pioneers such as Alan Turing and John von Neumann explored the possibility of building machines that can think and communicate. Early AI systems relied on manual creation of rules and algorithms that specified how the machines should behave in different situations.

Machine learning (ML) is a subfield of AI that focuses on creating systems that can learn from data and improve their performance without being explicitly programmed. ML algorithms can analyze large amounts of data, identify patterns, and make predictions or decisions based on the data. ML has many applications in various domains, such as natural language processing, computer vision, recommender systems, and self-driving cars.

One example of ML is Tesla's full self-driving system, which uses generative models to control the vehicle's behavior. Generative models are a type of ML model that can generate new data or outputs based on the input data. For example, a generative model can create realistic images of faces or landscapes based on some input parameters.

Tesla's full self-driving system uses two types of generative models: one for training and one for deployment. The training model takes i) recorded video data from cameras mounted on the vehicle as input, and ii) the corresponding pedal and wheel movements that the driver made as expected output. The model is trained to behave like the human driver based on the video data. The deployment model takes live video data from cameras mounted on the vehicle as input, and outputs the pedal and wheel movements in real-time that the model chooses based on its learned behavior. The model is deployed to control the vehicle's actions in real time.

The advantage of using generative models for self-driving is that they can cope with a wide variety of situations and scenarios without relying on pre-planned rules or maps. The model can adapt to changing road conditions, traffic patterns, weather, and other factors. However, generative models also have some limitations and challenges. For example, Tesla's self-driving system sometimes fails to stop completely at stop lines or follow traffic signals correctly. This is because the model, having been trained on real-life data, may simply be aping the less-than-perfect behaviour shown in the input videos. Therefore, generative models need to be constantly monitored, evaluated, and improved to ensure safety and reliability.

WHAT IS CHATGPT?

ChatGPT (Chat Generative Pre-trained Transformer) is a chatbot system that uses a large-scale pre-trained language model called GPT-3 to generate natural and engaging responses to user messages. ChatGPT leverages the power of GPT-3 to produce coherent and diverse texts that can cover a wide range of topics and domains. ChatGPT also incorporates some techniques to improve the quality and safety of the generated responses, such as filtering, rewriting, and human evaluation. ChatGPT aims to provide a conversational experience that is both informative and entertaining for the users.

One of the most remarkable innovations in the field of natural language processing, ChatGPT was developed by OpenAI, a tech research company that aims to create artificial intelligence that can benefit humanity. The chatbot uses deep learning, a branch of machine learning that mimics the way the human brain works, to generate human-like text based on prompts from users. It does this by predicting the next word in a given text, based on the patterns it has learned from a massive amount of data during its training process. The data source for the chatbot is the whole internet, which means it can access a vast and diverse range of information and knowledge.

The chatbot was launched on November 30, 2022, and quickly became a sensation among users who wanted to try

out its capabilities and have fun with its responses. By January 2023, it was the fastest-growing consumer software application in history, with more than 100 million users worldwide. The chatbot also made worldwide news headlines for its ability to write essays, blog posts, and promotional copy on various topics and domains, with impressive quality and creativity. However, the chatbot also comes with a disclaimer that its answers are not guaranteed to be correct or accurate, and that users should exercise caution and critical thinking when using it.

CHATGPT IN ACTION

ChatGPT can be found at chat.openai.com. It offers a freeform text box at the bottom middle of its screen into which users can type their "conversations". Figure 1 shows a simple example in which I asked "what is phuse".

Notice how the free public version of ChatGPT has only been "fed with the internet" up to a not-so-recent date. The model retains the context of the previous question, so the next question can make use of the existing context, see Figure 2.

But ChatGPT (and other forms of the generic generative approach) can do more than reply in English. See a coding example in Figure 3. It may not be the fully formed, commented SAS program, meeting your company standards, that you might desire, but it's a good start. And Figure 4 shows how the quality can be upgraded, by request.

The generic generative model technique that underpins ChatGPT has been adopted by many companies, producing a flurry of their own Chatbots with generative capabilities. Here are some:

OpenAI	ChatGPT
Microsoft	Bing AI
Github	Copilot
Google	Bard
Amazon	CodeWhisperer

OPPORTUNITIES AND BENEFITS

One of the opportunities and benefits is to use a natural language interface to query reference manuals and user guides so that they provide faster and quicker access to information for languages, editors and software development environments. For example, a user can ask a natural language question about the syntax or functionality of a programming language, and get a relevant answer from the documentation with uniquely formed syntax and semantics. This can help users learn new skills, troubleshoot problems, and improve their productivity.

Similarly, natural language interfaces can also be useful for accessing regulations, such as those related to clinical trials or data privacy. Users can ask questions about the requirements, procedures, or best practices for complying with the regulations, and get answers that are accurate and up-to-date. This can help users avoid errors, fines, or legal issues, and ensure that they follow the standards of their industry. That said, the caveats about accuracy still apply.

Another opportunity and benefit of using natural language interfaces is that they can generate tables, listings and figures (TLFs) on behalf of statistical programmers. This can be done by using large language models (LLMs) such as GPT-4, which can transform natural language queries into interactive, reproducible plots. For example, a user can ask a natural language question about the distribution or relationship of some variables in a clinical dataset, and get a corresponding TLF that visualizes the answer. This can help users derive insights from clinical data, without having to write complex code or use specialized software. Moreover, natural language interfaces can also generate mock shells, test data, and TLFs for new studies, by using the digital protocol, digital analysis plan, and metadata such as corporate data standards as inputs. This can help users speed up the development and validation of TLFs, and free up time for more creative and challenging tasks.

However, there are also some challenges and limitations of using LLMs for these purposes. One of them is the trustworthiness of the results. Users need to be confident that the answers or outputs they get from the models are correct and consistent with the source documents or data. This may require verification, quality control, or peer review processes to ensure that the results are reliable and reproducible. Another challenge is the training of the models. Users need to ensure that the models are trained on relevant and representative data and documents, and that they can adapt to different formats, standards, or terminologies. This may require customisation, fine-tuning, or feedback mechanisms to improve the performance and accuracy of the models.

But, is this just hot air, or is the opportunity real?

Research by Karma Dorje Tarap at BMS has shown the ability to LLMs to turn natural language queries into interactive, reproducible plots (in an internally hosted model, not the public ChatGPT base). This was shared in September 2023 and the results are impressive.

One of the challenges of the drug development process is the creation and use of digital documents that can facilitate study start-up and data flow. Several industry initiatives are working on this issue, such as TransCelerate and CDISC. TransCelerate's Digital Data Flow initiative aims to automate the generation of study start-up assets, such as Case Report Forms, Procedure Manuals, Statistical Analysis Plans, and Schedule of Activities, using an open-sourced, vendor agnostic technical solution. This would reduce cycle times and improve data quality for sponsors, third-party providers, sites and regulators. CDISC is collaborating with TransCelerate to develop a Study Definition Reference Architecture that will serve as a standard model for the development of conformant study definition technologies, including the TransCelerate-developed Study Definitions Repository Reference Implementation. These initiatives are examples of how the industry is moving towards a more efficient and effective digital data flow in drug development.

These initiatives are perhaps at an early stage in the process and not yet of benefit for our proposed purposes, but their time will come.

CONCLUSIONS AND NEXT STEPS

In this paper, we have explored the potential of using artificial intelligence to speed the delivery of statistical results from clinical trials. We have argued that a significant number of people believe that a beneficial outcome is possible, but that a combination of i) digital documents, along with ii) contextual metadata plus iii) copious historical trial data is required to facilitate a capable model. We have also discussed some of the challenges and limitations of this approach, such as data quality, ethical issues, and computational complexity. We conclude that progress will be stepwise, as the capability of the modelling software and associated techniques are advancing at pace. Some possible directions for future research include incorporating more diverse sources of data, developing more robust evaluation methods, and engaging with stakeholders and policymakers to ensure the relevance and usability of the models.

FIGURES

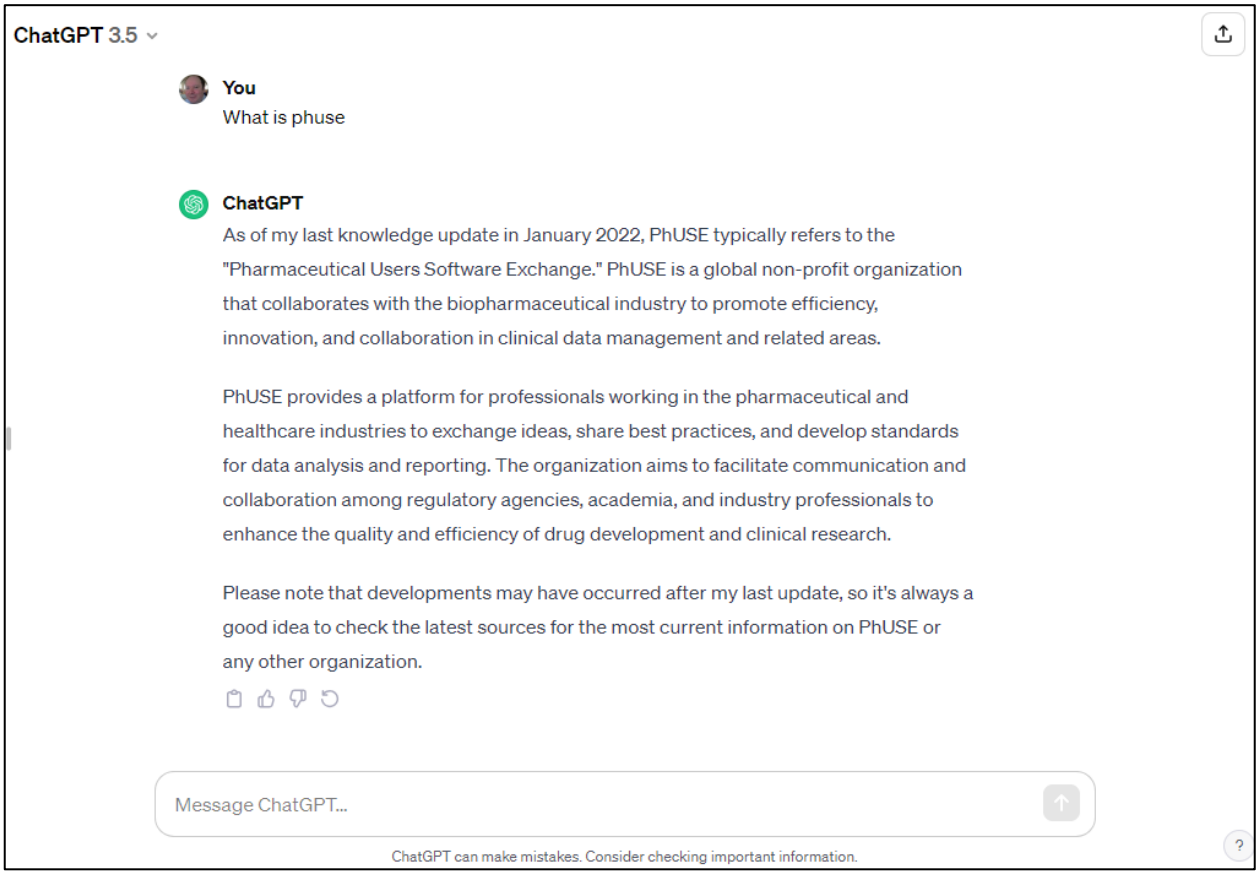

Figure 1. Simple ChatGPT conversation

Figure 2. ChapGPT retains context

You

show me a sas program for listing clinical demography data

ChatGPT

Certainly! Below is a simple example of a SAS program that can be used to list clinical demography data. In this example, I'm assuming you have a dataset named "Demography" with variables such as PatientID, Age, Gender, and other demographic information. Adjust the program based on your actual dataset structure.

sas Copy code

```
/* Example SAS program for listing clinical demography data */

/* Assuming 'Demography' is the name of your dataset */

/* Display the contents of the dataset */
proc print data=Demography;
  title 'Clinical Demography Data Listing';
run;

/* You can also use PROC CONTENTS to display information about the data
proc contents data=Demography;
  title 'Dataset Information';
run;
```

This program uses PROC PRINT to display the contents of the dataset and PROC CONTENTS to provide information about the dataset's structure. You may need to modify the program to suit your specific dataset and analysis requirements. Make sure to replace "Demography" with the actual name of your dataset.

If you have specific variables or criteria you want to include in your listing, you can use the VAR and WHERE statements in the PROC PRINT to customize the output. For example:

sas Copy code

```
proc print data=Demography;
  title 'Clinical Demography Data Listing';
  var PatientID Age Gender /* Add other variables as needed */;
  where Age >= 18; /* Example: Display only patients 18 years and older
run;
```

This is a basic example, and the actual program will depend on the structure of your dataset and the specific information you want to include in the listing. Adjust the code based on your data and analysis requirements.

Figure 3. Code generation

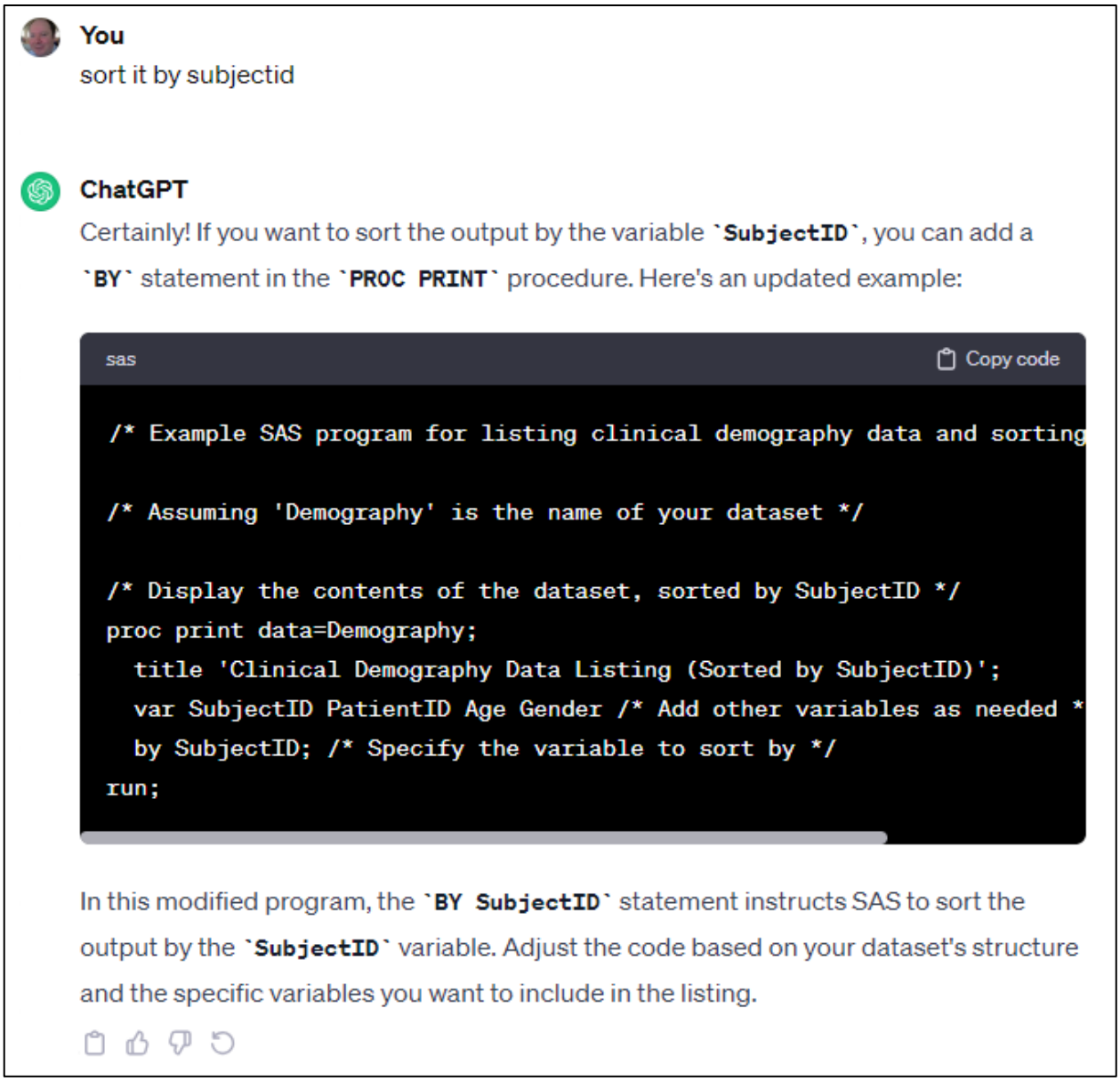

Figure 4. Augmented coding

RECOMMENDED READING

What Kind of Mind Does ChatGPT Have? Cal Newport
https://www.newyorker.com/science/annals-of-artificial-intelligence/what-kind-of-mind-does-chatgpt-have

ChatGPT
https://chat.openai.com

LLMs Unleashed: A New Chapter in Clinical Programming? Karma Dorje Tarap
https://www.linkedin.com/pulse/llms-unleashed-new-chapter-clinical-programming-karma-dorje-tarap

Digital Data Flow Solutions, Transcelerate
https://www.transceleratebiopharmainc.com/assets/digital-data-flow-solutions/

Digital Data Flow, CDISC
https://www.cdisc.org/ddf

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Andy Ratcliffe:
Company: Instem plc:
Email: Andrew.ratcliffe@instem.com
LinkedIn: https://www.linkedin.com/in/andrewratcliffeuk/