

**PSI Careers Lunch & Learn: ChatGPT & Beyond**  
**Tuesday 12<sup>th</sup> September 2023**

As mentioned during the webinar, we received some great questions from our audience, though unfortunately there were more than time allowed! So we asked our Panel to provide further responses to questions that we didn't have time to address, and they have very kindly obliged. Please therefore see below a list of the unanswered attendee questions, with responses from panel.

- 1. 2-parter question. What safeguards do you think should be implemented to protect sensitive data, especially in pharmaceutical industry, when utilising these tools? Also, do you see ChatGPT and other similar tools will be implemented into existing programming languages such as SAS and R to further develop these languages (via SAS macros or PROC FCMP) and aid automating programmer tasks or do you envision these be separate tools used independently such as the code generator for figures you have shown?**

*[Tom Fisher] Internal versions with self-contained pipelines of these tools are important! Definitely, there are coding assistants built using LLMs that are becoming very popular.*

- 2. If these tools are free (or become free), then how do the businesses behind them make money?**

*[Tom Fisher] I doubt the top-end tools will ever be completely free given the compute time required for training such large models. I expect companies will continue to charge for large-scale API calls to the more advanced models and for custom internal solutions.*

- 3. If i ask the same question twice it gives me a different answer, how to we ensure reproducibility when using these tools?**

*[Tom Fisher] One option could be to look at the probabilities from the models in the back end - these are what are sampled from when responses are generated.*

*[Kleio Kipourou] To add on Tom's answer; ChatGPT has a parameter called "temperature" that controls the randomness. If temperature is set to 0 then the model chooses the most probable word (every time) and becomes more deterministic. Temperature can be used in a different context to help you with your task in hand. For more info please see here <https://community.openai.com/t/cheat-sheet-mastering-temperature-and-top-p-in-chatgpt-api-a-few-tips-and-tricks-on-controlling-the-creativity-deterministic-output-of-prompt-responses/172683>*

- 4. If you ask ChatGPT to critique its responses (e.g. which of these references did you make up?) Does it admit to its hallucinations?**

*[Tom Fisher] Great question! For the new versions with web search plugins, it will provide links to the websites that it claims are references so it should be able to check where it's gone wrong. However with base ChatGPT, due to the probabilistic nature of the output, it might not always admit to hallucinations, despite sometimes (rather scarily), ChatGPT claiming it is "incapable of making mistakes!"*

- 5. A friend of mine(engineer) who uses ChatGPT a bit more than me suggested it may be effectively Google 2.0. As in you get the results back quicker from searches but you still need to know a bit what you're looking at in order to really know whether it's correct. Is that a fair assessment?**

*[Tom Fisher] A sanity check is definitely always wise, yes. [Andy Nicholls] 100%.*

- 6. "You mentioned that ChatGPT can be used to do simple programming tasks or to use it as a slightly better search in stack overflow when searching for code snippets. However, this feels a bit ""weak"" as you already know what you would like to do. Letting it summarizing papers sounds great, but if Chat GPT likes to make up some facts you anyhow have to read the complete paper completely - so I see limited use for using Chat GPT as a statistician right now. Do you have more concrete examples where Chat GPT really helped you / solved some problems?"**

*[Tom Fisher] Here's just one example from me: I've used ChatGPT in personal projects to create a first draft of code in a language I'm not very familiar with and I'm just starting to learn. Sometimes it needs*

*tweaking but if you build up to the final answer via multiple prompts then it tends to perform better. At each stage, asking ChatGPT to explain what's going on and checking the documentation can fast-track you through learning and producing good code.*

*[Kleio Kipourou] I also used it to add comments to my code, to suggest me alternative functions in order to improve my code's speed (with justification) and even ask it to write an abstract based on a paper.*

*[Andy Nicholls] It's certainly possible to do more. A colleague of mine used only ChatGPT queries to build an entire JavaScript app, without any prior Javascript knowledge. However debugging then proved ... tricky! The message here is that you need a bit of subject matter knowledge to know if what it has returned is useful. Most of the time I do know roughly what I'm looking for in Stack Overflow but I need to search because I'm rusty and have forgotten the function name or the programming trick required. ChatGPT gets straight to the point. As a concrete example I recently asked, "what's the stringr function to trim white space from text?" ChatGPT returned the code with a neat (usually) reproducible example.*

**7. Do any of you have experience using Google's Bard for programming &/or analysis? If so, any thoughts on how it compares with ChatGPT?**

*[Tom Fisher] Bard's Pathways Language Model (PaLM) makes it very good for coding and maths, comparable with ChatGPT.*

**8. Does code generated by ChatGPT need less or more validation than self-written? Regulatory view?**

*[Tom Fisher] Personally I would recommend it undergoes the same checks as any other code. Given the points on inaccuracy mentioned by Andy during the call, it is crucial for code and outputs to be comprehensively reviewed. Regarding the regulatory side, I cannot see purely AI-written code being adopted any time soon, but it could perhaps be used as an additional further quality control check to make this possible with additional speed AI provides.*

*[Kleio Kipourou] Totally agree with Tom. Regardless of the author of the code, any code should be subject to scrutiny and depending on the purpose, it must follow specific standards. That at the moment translates to human supervision and further quality checks by appropriate committees.*

**9. Is it significant that most of the speakers seem to be from the "exploratory" world rather than "confirmatory"?**

*[Tom Fisher] Perhaps! I hope we haven't given too biased a view given how we all mentioned that bias is of such a concern!*

*[Andy Nicholls] Actually, ~60% of my team's work is on the submission end of innovation and we're exploring some very real use cases. It's fair to say that most of the work around LLM's is still very exploratory in nature.*