ChatGPT and research: some interim information

ChatGPT is an artificial intelligence (AI) language model that uses complex algorithms to generate a relevant response to users' questions. It can be used to answer questions or generate responses to most questions or inputs as the model was trained using a huge array of text from the internet from inception to 2021.

Its use in academia and research is evolving with international consensus still developing on how to use these tools responsibly to facilitate research. This guide explains what ChatGPT is, how it uses the information inputted and how people are currently using it in health and medical research. We then provide some interim information on the known risks to using such AI models.

Please note that this is information rather than endorsed guidance as this is a rapidly evolving field and caution should be used when considering how you use ChatGPT in the context of your own work. This information is accurate as of 9 May 2023 and researchers should take care to refer to the most up-to-date guidance in this area.

What is ChatGPT?

ChatGPT is an artificial intelligence (AI) language model that uses an exceptionally large amount of data fed into the model up to 2021 and advanced algorithms to generate responses to users' questions or statements.

When you input text into ChatGPT, the model uses a complex algorithm to analyse and understand the meaning behind your input. Once the model has analysed your input, it uses this information to generate a response.

ChatGPT uses a type of machine learning called 'unsupervised learning' to continually improve its understanding of language and generate more accurate and relevant responses. This means that the model learns from the input it receives over time and adapts its responses based on patterns and trends it identifies in the data. This is based on a model to determine, word-by-word, what the next word in a response will/should be. *'Given the text so far, what should the next word be?'*

ChatGPT is just one of several online digital assistance tools using AI language models, and while some information provided here may be relevant to other tools, our guidance remains specific to ChatGPT and its capabilities at the time of writing. Advice regarding other digital assistance tools will be provided as more information comes to light. Each system will have its own idiosyncrasies and these need to be taken in to account. An example is Bard (Google) which continually draws on new information unlike ChatGPT that is limited to information available up to 2021.

How is ChatGPT being used in health and medical research?

The use and function of ChatGPT is rapidly evolving. The below list summarises how ChatGPT is likely being used for health and medical research. **This guidance is NOT**

endorsing the use of ChatGPT for these research tasks or methods but providing information on how it is being used. At the time of this guidance, ChatGPT is likely being used in the health and medical research sector to:

Research management

- Support administrative tasks e.g. drafting emails, sections of business cases, etc.
- Facilitate the drafting of research outputs, ethics applications and/or funding applications
- Facilitate development of research instruments, e.g. helping to simplify participant-facing information such as recruitment notices, feedback questionnaires, consent forms, surveys and/or instructions
- Conducting literature reviews
- Summarise one or more papers or theories in plain language
- Assist researchers to identify a gap in the existing scholarship and suggest a new research topic
- Find journals based on your research topic including ranking and providing short summaries of author requirements

Research analysis

There may be circumstances where ChatGPT may assist with research process. The concerns outlined below about scientific rigour, privacy and accuracy also apply to research analysis.

Ethical and scientifically rigorous research processes should be the first consideration.

Given that qualitative research frequently involves sensitive information from interviews and focus groups it is unlikely that this is an appropriate tool for qualitative research analysis. In addition, researchers understanding the data, developing codes on an iterative basis and interpreting the data in the context of the study is integral to high quality qualitative research and is unlikely to be achieved given the limitations of AI tools.

How does ChatGPT use the information I input?

ChatGPT is designed to operate without actively collecting or storing personal information, however it does store the information or 'conversations' you input into in order to continually learn from inputs. As such, **sensitive information must not be inputted into ChatGPT.**

Sensitive information includes personally identifiable information, financial information, health information or confidential business data (including finding and results arising from the conduct of research projects). For more information see here: https://intranet.mcri.edu.au/sites/policies/Pages/Cyber-Security-Standard---Information-Security-Classification.aspx

Intellectual Property, Confidentiality and ChatGPT

ChatGPT stores and uses the information or 'conversations' you input into it. As a consequence, putting information into ChatGPT can be considered to be a disclosure of that information to a third party and may be regarded as a public disclosure without appropriate confidentiality obligations. This means that **confidential information must not be inputted into ChatGPT**, as it may result in the researcher's organisation breaching its obligations to third parties with respect to keeping information confidential and it may impact the ability of the researcher's organisation to obtain patent protection for a new invention.

How accurate is ChatGPT?

It is very hard to determine the accuracy of ChatGPT as it depends on the task you ask of it and the quality and quantity of data used to train it on that particular topic. ChatGPT cannot access the internet to retrieve its own data but relies on data that it was 'fed' prior to 2021.

This means that the outputs generated by ChatGPT can be out of date, misleading and/or inaccurate. Because of this, users should be very cautious and check any outputs generated by ChatGPT for accuracy.

Using ChatGPT to write research outputs: what are the risks?

Scientific rigour

One of the emerging uses of ChatGPT is using it to assist with writing manuscripts or sections of funding applications. While ChatGPT can generate text that sounds like a funding pitch or scientific manuscript, it is not capable of performing the critical steps of data analysis, interpretation and contextualisation that is required for rigorous scientific research. It can also create misleading or incorrectly cited information.

Privacy

ChatGPT is designed to operate without actively collecting or storing personal information, however it does store the information or 'conversations' you input in order to facilitate its continual learning. As such, **sensitive information must not be inputted into ChatGPT.**

Sensitive information includes: personally identifiable information, financial information, health information or confidential business data.

Intellectual Property and Confidentiality

Please refer to the relevant section above.

Accuracy

It is very hard to determine the accuracy of ChatGPT as it depends on the task you ask of it and the quality and quantity of data used to train it on that particular topic. ChatGPT cannot access the internet to retrieve its own data but relies on data that it was 'fed' prior to 2021. This means that the outputs generated by ChatGPT can be out of date, misleading and/or

inaccurate. Because of this, users should be very cautious and check any outputs generated by ChatGPT for accuracy.

Whilst ChatGPT can be useful in managing the labour of writing (e.g., improving the readability or language of work already written), its potential inaccuracies mean it should be used with caution when generating large outputs of text. Citations can also be problematic if you are synthesising multiple papers. Although you can ask ChatGPT to do a literature search, references provided may not be up to date and in some cases may not exist. Caution is advised even when using this synthesis as a starting point for a literature review. Instead, asking ChatGPT for ideas on how to structure your writing or improve the readability of your already written work may be more useful.

Plagiarism

Research outputs require careful citation and attribution of sources. **Citation of original sources of information is something that ChatGPT does not do accurately**. As such, you may be unknowingly plagiarising sources if using text generated from ChatGPT. This is a breach of the Australian Code for the Responsible Conduct of Research, as it is a key responsibility of researchers to check their outputs for accurate referencing. of researchers to check their outputs for accurate referencing.

Authorship

Publishers are rapidly expanding their guidance on how ChatGPT can be used in the preparation of research outputs. You should check with the publisher of the journal you are submitting to for specific advice on how these tools can or cannot be used. Organisations such as Committee on Publication Ethics (COPE) have provided clear guidance that ChatGPT should not be listed as an author as it cannot take responsibility for the submitted work:

> COPE advise that *"Authors who use AI tools in the writing of a manuscript, production of images or graphical elements of the paper, or in the collection and analysis of data, must be transparent in disclosing in the Materials and Methods (or similar section) of the paper how the AI tool was used and which tool was used. Authors are fully responsible for the content of their manuscript, even those parts produced by an AI tool, and are thus liable for any breach of publication ethics."*

So can I use ChatGPT for research or not?

MCRI does not intend to regulate all areas where emerging tools such as ChatGPT are used by our researchers. Instead, we want to ensure our staff have a good understanding of how these tools work and the benefits and limitation of such technology including areas where it is inappropriate for use.

We understand that tools such as this are emerging all the time to support researchers, with tools such as referencing software, statistical programs and even spell/grammar checks actively encouraged. We know that these techniques can help our researchers with the administrative and manual tasks related to research. However, ChatGPT is an assistive tool

and not a replacement for rigorous scientific methods. Researchers should always conduct original research, analyse data, and interpret results themselves in order to produce scientifically sound and rigorous outputs.

Researchers are ultimately accountable for their output regardless of how it is generated.

Sensitive information (as outlined above) should never be fed into ChatGPT.