



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Borger, JG;Ng, AP;Anderton, H;Ashdown, GW;Auld, M;Blewitt, ME;Brown, DV;Call, MJ;Collins, P;Freytag, S;Harrison, LC;Hesping, E;Hoysted, J;Johnston, A;McInnery, A;Tang, P;Whitehead, L;Jex, A;Naik, SH

Title:

Artificial intelligence takes center stage: exploring the capabilities and implications of ChatGPT and other AI-assisted technologies in scientific research and education

Date:

2023-11-01

Citation:

Borger, J. G., Ng, A. P., Anderton, H., Ashdown, G. W., Auld, M., Blewitt, M. E., Brown, D. V., Call, M. J., Collins, P., Freytag, S., Harrison, L. C., Hesping, E., Hoysted, J., Johnston, A., McInnery, A., Tang, P., Whitehead, L., Jex, A. & Naik, S. H. (2023). Artificial intelligence takes center stage: exploring the capabilities and implications of ChatGPT and other AI-assisted technologies in scientific research and education. *Immunology and Cell Biology*, 101 (10), pp.923-935. <https://doi.org/10.1111/imcb.12689>.

Persistent Link:

<https://hdl.handle.net/11343/340012>

Capabilities and implications of AI in science

## FUTURE CHALLENGES

[Free to View]

### Artificial Intelligence takes centre stage: exploring the capabilities and implications of ChatGPT and other AI-assisted technologies in scientific research and education

**Running title:** Capabilities and implications of AI in science

Jessica G Borger<sup>1\*</sup>, Ashley P Ng<sup>1</sup>, Holly Anderton<sup>1</sup>, George W Ashdown<sup>1</sup>, Megan Auld<sup>1</sup>, Marnie E Blewitt<sup>1</sup>, Daniel V Brown<sup>1</sup>, Melissa J Call<sup>1</sup>, Peter Collins<sup>1</sup>, Saskia Freytag<sup>1</sup>, Leonard C Harrison<sup>1</sup>, Eva Hesping<sup>1</sup>, Jaci Hoysted<sup>1</sup>, Anna Johnston<sup>1</sup>, Andrew McInnery<sup>1</sup>, Phil Tang<sup>1</sup>, Lachlan Whitehead<sup>1</sup>, Aaron Jex<sup>1</sup> and Shalin H Naik<sup>1\*</sup>

<sup>1</sup>WEHI (Walter and Eliza Hall Institute of Medical Research), Parkville, Australia; Department of Medical Biology, The University of Melbourne, Parkville, VIC, Australia.

\***Correspondence** should be addressed to:

Jessica G Borger<sup>1</sup>: [borger.j@wehi.edu.au](mailto:borger.j@wehi.edu.au)

Shalin H Naik<sup>1</sup>: [naik.s@wehi.edu.au](mailto:naik.s@wehi.edu.au)

**Contact Address:**

WEHI (Walter and Eliza Hall Institute of Medical Research), 1G Royal Parade, Parkville, VIC 3052, Australia.

**Keywords:** ChatGPT, AI, LLM, Science, Research, Education, Ethics

The emergence of large language models (LLMs) and assisted artificial intelligence (AI) technologies such as ChatGPT and Bard have revolutionised the way we interact with technology. These publicly available LLMs can generate cogent, human-like and human-level responses and have a diverse range of potential applications across diverse knowledge areas, including scientific research and education. However, with such advancements comes a new set of ethical, legal and social implications. Medical research and education are no exceptions, and our organisations must contend with new models of governance and responsibility. A recent Chat-GPT symposium at the Walter and Eliza Hall Medical Research Institute (WEHI)[1] explored the current practical applications of LLMs in medical research and canvassed the emerging ethical, legal and social implications for the use of AI-assisted technologies in the sciences. The symposium was led by early career researchers, lab heads, educators and policy-makers, representing the diverse academic landscape within medical research institutes who engage with ChatGPT in their work, who as experts in their fields are learning to navigate the appropriate, efficient and ethical application of AI and LLMs. Together, the speakers sought to provoke discussions within the 500+ in-person and online audience on the use of AI-assisted technologies in scientific research, including its use as an overtly friendly editor of scientific papers and grants, its ability to turn non-coders into bioinformaticians, and its ability to analyse big data at warp speed. In addition to the emergence of AI-driven tools such as

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1111/imcb.12689](https://doi.org/10.1111/imcb.12689)

Alphafold and protein hallucination, the symposium addressed broader societal implications of using AI-assisted technologies in science, including concerns around the ethics, privacy, confidentiality and security of research data and writing that is entered into the AI-assisted technologies ether. This paper provides an overview of the WEHI Chat-GPT symposium's key themes and discussions, highlighting the opportunities and challenges that lie ahead for scientific researchers and educators as we continue to explore the potential of this cutting-edge and emerging technology.

### ***An overview of generative AI models***

“Artificial intelligence” has certainly captured the popular imagination since the launch of ChatGPT, but models of AI have been widely used in scientific and clinical research for some time[2]. Models can drive cars, recognise images, and even create synthetic data – but ChatGPT is different in several ways. As a large language model (LLM), it is immediately accessible, and interacting with the model is as easy as having an online conversation. Designed to allow users to enter natural language “prompts” to “generate” a human-level response, depending on the nature of the prompt, the output can often surpass the knowledge, expertise and efficiency of the human entering it. This is relevant to many “human”-driven tasks from basic editing and distillation of a topic to complex analysis and collation of dispersed information.

Crossing the threshold from science-fiction into reality has required significant technological and financial investment into the development and training of neural network-based models by well-resourced AI technology focussed companies and collaborations. GPT4 for example, required significant investment from Microsoft to enable the OpenAI AI startup for the development and evolution to its current form. This highlights the scale of “training” required to produce their predictive text generating model and the important logistical considerations required for model implementation prior to public release, including how such models may interact with societal ethical, legal and social implications considerations.

OpenAI has made a splash in the AI Language Network space before. The earlier GPT2 LLM – a text completion neural network launched by OpenAI in 2019 – had “limited” release amid concerns a full version may be used to create fake news articles or be used for other nefarious purpose[3]. The full 1.5B parameter network was released soon after, when their claimed fears turned out to be an overestimation of the network’s performance and that its output did not traverse the “uncanny valley” to human-like responses (Figure 1). However, the leap between GPT2 to subsequent iterations of GPT3.5 and most recently GPT4 was stark. It is interesting that while OpenAI has given access to the full model, they have built in several safeguards to mitigate its malicious use, as anyone who has used it has almost certainly come across the phrase “As an AI language model, I can-not...”.

AI tools designed for academic researchers have seen significant growth and adoption in recent years. These tools are created to address various challenges faced by researchers, helping them streamline their work, improve efficiency and enhance the quality of their research. There now exists an expansive toolbox for academic researchers to support writing with inbuilt reference managers, image and video analysis, AI-assisted survey and experimental design platforms as well as plagiarism detectors within education (Table 1). These AI technologies and related tools have revolutionised the academic landscape, making it easier for researchers to tackle complex challenges and focus more on the creative aspects

of their work. While AI can offer tremendous benefits, it is essential for researchers to understand the limitations and potential biases of these tools to ensure reliability and validity of their research findings.

[Table 1 about here]

The WEHI Chat-GPT symposium identified two broad, overlapping themes to be considered when adopting ChatGPT and related LLM and AI-based tools in scientific and medical research: (1) the wide range of applications for AI in scientific research, communication and education and its potential to improve as well as confuse; (2) the implications, current challenges and potential future developments in the application of AI to broader academic domains, including ethics, law, security and intelligence, and analysed these in context of what it will mean to be a scientist in the future.

We will now discuss these themes, with consideration of how these LLM and AI-based tools will further impact science and academic domains as they become integrated into widely used word-processing, spreadsheet and multimedia software. The following discussions are to provoke thought and invoke discussions amongst the readers on these broad themes that as scientists, educators and ethicists we are navigating with limited, but growing, understanding and experience. Expert reviews on these themes can be found in Table 2.

[Table 2 about here]

### **Theme 1: The diverse potential and pitfalls for using ChatGPT in scientific research**

LLMs can be applied in various applications across a breadth of scientific and medical research, which is not limited to the work performed at the bench, including the potential for AI-assisted technologies to bridge language barriers in science. Indeed, there is an exciting potential for AI-assisted technologies to improve accessibility and facilitate collaboration between non-native English speaker scientists from different parts of the world. However, real concerns have been identified around bias and accuracy in text generation, particularly in the context of scientific research where accuracy and objectivity are paramount.

#### ***Use of ChatGPT to assist with writing and reading grants and papers***

Scientific literature is vital in advancing knowledge, but poor readability often poses a significant challenge. This issue goes beyond the use of technical jargon and incorrect English syntax. Common barriers to comprehension in scientific writing include excessive passive voice, long and convoluted sentences, and unnecessarily complex language. Poorly written articles can hinder effective communication and impede dissemination of scientific findings within the scientific community and beyond. Additionally, in the increasingly competitive landscape of scientific funding, it is essential to convey the quality of research, its significance and a sense of excitement, while remaining accessible. In this regard, ChatGPT has emerged as an effective writing assistant. ChatGPT can clarify convoluted text and ambiguous statements that may hinder a reader's understanding. It can also identify and simplify jargon and complex terminology, making scientific writing more accessible to non-experts. This is particularly beneficial in grant writing, where reviewers making career-defining decisions may lack subject matter expertise.

The pitfalls in generative AI can undermine ChatGPT's performance in simplifying and summarising scientific writing. The output from a generative model is statistically reconstructed language and does not guarantee scientific coherence. While an LLM can generate summaries encompassing research questions, main findings, methodology, results and implications, these summaries may be unreliable and inaccurate. Thus, an expert human author must critically assess AI-generated text through careful fact-checking and cross-referencing. Even with experts able to identify inaccuracies within their domain, the confidence with which generative AI can produce misinformation poses a real and present risk. Additionally, the training model's access to information is limited by its date restriction, excluding more up-to-date information unless newer applications that link the LLM to the internet are used.

Nonetheless, by appropriately leveraging ChatGPT's capabilities, scientists can optimise their time and expertise, allowing them to focus on improving the scientific quality of their work while using AI to help enhance the readability and accessibility of their writing.

### ***Using ChatGPT to break down language barriers for non-native English speakers in science***

As the world and scientific research increasingly move towards globalisation and international collaboration, the importance of proficient English language skills for biomedical researchers cannot be overstated. Non-native English speakers may face challenges with limited vocabulary, grammar rules and cultural nuances, creating a sense of separation among colleagues or collaborators.

In this context, ChatGPT has emerged as an empowering and inclusive tool. LLMs can accurately translate complex sentences in multiple languages, including those with different alphabets (Figure 2). This helps in breaking down barriers between English as the primary scientific language and the thousands of languages that are spoken by scientists around the world, thus advancing science, its collaborations and networks, globally. Excellent communication is an important attribute of being a successful scientist and the scientific community understands that individuals with English as a secondary language are disadvantaged. Communication between scientists, the foundation of a productive work environment and effective collaborations, can be compromised by the broad and potential unfamiliar written communication styles of diverse cultures, as experienced in different email practices. Here, LLMs can help bridge the communication gap. AI-aided translation of emails to colleagues or collaborators can improve context and vastly shorten the time that is needed to formulate complex matters that, in comparison, are more easily generated by native English speakers (Figure 3). No longer restricted only to basic translation, ChatGPT can serve as a personal assistant, teacher and translator, all in one platform. Specialised text-to-voice AI software can even help non-native English speakers improve their pronunciation, while applications such as Kick Resume AI aid in resume writing.

ChatGPT is a valuable resource for non-native English speakers for content-based questions, allowing them to retrieve information about specific biotechnological techniques, such as CRISPR, in their native language. Additional prompts can be used to request corresponding references to ensure accuracy. ChatGPT also excels as a proofreading tool. With Microsoft planning to integrate GPT4 as a co-pilot assistant, and Google intending to extend Bard LLM into the Google Suite of productivity tools, LLMs will inevitably help streamline many day-to-

day tasks. These integrations will facilitate AI assistance in activities such as drafting emails, summarising discussions, creating presentations and interpreting spreadsheet content, such that non-native English speakers stand to benefit greatly from the incorporation of these AI tools. Although LLMs have limitations, indeed translated text requires verification, effective communication is one of a professional's most powerful skills, and leveraging AI to improve individual communication skills will foster confidence and inclusion in the scientific community.

### ***Using ChatGPT to improve the experience of thesis writing for students and supervisors***

Supervisors of graduate students often provide multiple rounds of feedback as students construct their thesis. In addition to reviewing scientific content, correcting language structure and grammar can form a large part of thesis revision. While students may also have access to writing courses and professional copyediting services that proofread their thesis, such services are not universally accessible or cost-effective. AI typing assistants, like Grammarly, are already used with little controversy to provide real-time writing feedback on spelling, grammar, punctuation, and clarity, suggesting replacements for identified errors. With GPT4 soon to become part of the Windows365 suite, ChatGPT will become just another icon, located next to Grammarly, to provide writing assistance. In principle, this should reduce the editing load for students and supervisors. Much less time will be spent on removing commas and breaking up paragraphs, and more time talking about what is important – the science. Yet, it raises broader concerns within the education sector regarding the originality of content that is produced by the student.

Many supervisors and educators are worried that students will lose the critical skills of academic writing and editing that we are so passionate about developing in our students. However, we need to reflect on our own writing – with red underlined typos and blue underlined grammatical errors currently highlighted in Word and Powerpoint documents, we don't simply accept them; we review and correct them as appropriate. A thesis that has had any form of proofreading from a supervisor, a professional copyeditor or ChatGPT will still need to be revised by the student themselves.

Supervisors and educators should acknowledge that, if used appropriately, students will still learn the same writing and editing skills they traditionally have, just differently, and that is ok. As LLMs are currently unable to undertake logical or critical thinking in any true sense, their use is limited to writing assistance, not as a search engine that automates the critical skills required to produce exceptional researchers.

### ***Organising a Conference Program with ChatGPT***

A WEHI researcher asked ChatGPT to organise a conference for an Australian Medical Research Consortium whose members and interests are readily accessible from websites and publications. To enhance accuracy and richness, the prompt request initially included key details like title/theme, purpose, topics, target audience and number, location, duration and total budget. The aim was to see whether ChatGPT could produce more than just a generic program covering organisation and content.

Unfortunately, the result was rather disappointing. Firstly, of the 10 scientists named as members of the organising committee or as speakers, eight had no connection to the

consortium or the theme. Even for the two relevant individuals, their affiliations were incorrect. None of the individuals named on the organising committee were known consortium members. Attempts to improve the selection of relevant contributors was not improved by regenerating the request with additional scientific keywords or even suggested speakers; in fact, except for copying suggested speakers, the accuracy declined. Secondly, the scientific content did not align well with current or future themes, which would have been easily traceable. Presentation titles were too broad, e.g. 'Multiomics in research', 'Controversies', 'Future Directions'. Finally, although ChatGPT provided a generic program structure with sessions for keynote speakers, oral and poster presentations, workshops and a suggested budget breakdown, the closest it got to specific and accurate content was to suggest conference venues and registration websites. At this time, although ChatGPT can provide a useful checklist for organising a conference, it lacks the creative ability to go beyond this and craft meaningful scientific content.

### ***Using ChatGPT to help with data analysis, coding and interpretation using bioinformatics***

As the size and complexity of datasets in biomedical science have grown over time, the demand for programming skills has often outpaced the availability of such expertise, frequently leading to a bottleneck in experimental iteration. Analysing large datasets using programming languages can be daunting for bench scientists. However, relying on professional bioinformaticians for such analyses can introduce delays and communication challenges.

Tools such as ChatGPT, explicitly trained in part on programming code, can empower bench scientists to analyse their own datasets. A non-coding user can describe their inputs and desired outputs in natural language and use LLMs to generate the code required for their bioinformatic analysis. The improvement from GPT3.5 to GPT4 was evident in the reduced number of prompts required to refine the execution of a basic task, with GPT4 needing only a quarter of prompts as GPT3.5. Indeed, in a recent study 97.3% of bioinformatics tasks were solved in 7 or less prompts. Despite these excellent results, the study also points out that relying on AI generated programming code can lead to erroneous outputs in the absence of code comprehension[8].

The huge potential of ChatGPT for bioinformatics education was highlighted by prompting the model to explain the reasoning behind the selection of specific functions, which resulted in a detailed description of the underlying algorithm. Furthermore, ChatGPT can be employed to summarise and interpret code produced by a bioinformatician to provide a plain language interpretation and summary for the bench scientist, including an assessment of the script's limitations. Further prompts can enhance the script, making it more comprehensible and reusable for other applications. However, to examine the true impact AI has on learning bioinformatics, controlled experiments in a classroom setting should be conducted[9].

A final consideration is how adopting AI models might impact bioinformatics as a profession. We expect an immediate surge in productivity of computational laboratories as AI models automate code prototyping. In the medium term, tools like the ChatGPT code interpreter could democratise bioinformatics by enabling direct dataset analysis via natural language using an integrated Python interpreter. The required skill set will likely shift from writing syntactically correct code to better comprehension and testing. In the long term, we envision

professional bioinformaticians being liberated from routine analyses, allowing them to concentrate on more bespoke and creative tasks. This shift, in turn, may favour individuals with stronger logical, mathematical and creative skills over those focused on raw output.

### ***Using ChatGPT to assist with complex image analysis***

The rise in availability of large-scale imaging data through multi-channel, multi-dimensional, long-term and live-cell microscopy is providing information-rich data which requires bespoke pipelines to extract meaningful results. To cope with the challenge of quantifying this imaging data, non-coding users have a few options: (1) collaborate with researchers who do code (ideally an imaging specialist); (2) try and work with previously published pipelines; (3) rely on the proprietary software included with some imaging platforms which are limited in their application and cannot be modified.

Incorporation of previously published pipelines often requires workarounds and cannot always be applied directly to in-house datasets, even those using the same techniques from different laboratories. Challenges of achieving this when developing an opening source pipeline for detecting and counting cells, are clearly visible when using training data from different research groups around the world to increase the robustness of the model.

New imaging methods such as lattice light-sheet are now offering plug and play long-term (hours to days), video-rate, 3D, live-cell datasets. This has huge potential for answering complex biological questions for any researcher with access to a system. However, as with many imaging modalities, data handling and analysis is an ongoing challenge and even an afterthought. As such ChatGPT's coding capability offers an exciting opportunity to accelerate users in the rapid and efficient generation of bespoke coding packages. An example presented by imaging specialists at WEHI asked whether ChatGPT could be used to assist in generating a framework or at least in laying the coding foundations for an imaging data analysis workflow. The AI-generated code (in this case for Python) could import relevant libraries, load the imaging data, segment regions of interest using thresholding methods and attempt to quantify the data over time (Figure 4). The results could then be plotted accordingly. However, when running the image analysis code on real lattice light sheet data, it quickly became apparent that there are significant limitations that LLMs may struggle to overcome when developing novel coding approaches (Figure 4). For example, when prompted to track segmented cells over time, ChatGPT's coding attempt could only plot direct paths between all detected cells. While we can appreciate that cells do not move in perfectly straight lines, this may not be apparent to the LLM. Mathematical errors were also common, so any steps requiring quantification required manual checking. It was also common, upon asking to fix coding errors, for ChatGPT to offer 2 or 3 potentially non-functional alternatives, requiring an experienced coder's attention to resolve. While ChatGPT can provide a starting basis for bespoke analytic pipelines, its performance is still only as good as the researcher's own knowledge, the clarity of the question asked and, most importantly, the user's ability to fact-check and verify the code and the output.

Despite the limitations in GPT3.5, GPT4.0 offers major advances where researchers can generate bespoke code with little or no coding background. ChatGPT4.0 is already being incorporated into other pipelines; Omega (<https://github.com/royerlab/napari-chatgpt>) takes the image processing and analysis code from ChatGPT and, through napari, attempts to

fix bugs and errors in real-time. Thus, using LLMs as part of an analytic ecosystem could be a significant area of growth in the future. For researchers that do code, it is becoming an invaluable assistant when writing our own code, acting as an alternative to StackOverflow for finding fixes to errors.

**Remarks on Theme 1: The potential application of ChatGPT in scientific research lies in the expertise of its users.**

The emerging concepts within Theme 1 demonstrate that ChatGPT can have real-world applications, but its effectiveness should be tempered by the user's expertise. In assisting with writing, it is crucial for users to guide the AI with specific prompts and to understand and edit the output appropriately. In many instances, experts may find it easier to write the text themselves and use AI to handle mundane but time-consuming tasks, such as copyediting. The advantages of this are evident, particularly for users for whom English is a second language.

Opinions vary among academics as to how useful ChatGPT is for coding, which largely depends on the user's area of expertise. ChatGPT can output compelling looking code; however, if the user is unable to understand each step of the code to check for valid treatment of the data, use of this tool is risky. For a wet lab scientist with a good understanding of the source data and the ability to assess the output of an analysis, ChatGPT can be extremely useful for routine coding tasks. For expert coders or instances where large chunks of code are generated for multistep tasks, ChatGPT significantly slows down the process by introducing errors that require extensive debugging. Used judiciously and for discrete well-defined and well-understood tasks, ChatGPT can speed up data processing and be enabling for scientists with little coding experience.

**Theme 2: Implications, challenges and future developments in the use of AI in broader academia**

The second theme of the symposium focussed on the ethical, legal and social implications of using machine learning and LLMs in our day-to-day academic work. The discussion revolved around not only what we can do but what we should do. It raised questions about the impact AI-assistance may have on the role of scientists and educators and what intellectual contributions truly mean in our work.

The discussions explored what might be gained by using these AI platforms versus what could be lost or put at risk. These considerations have profound ethical, legal and social implications, particularly given that the scientists and students working with generative AI will have varying levels of understanding, experience and awareness of its shortcomings. As AI continues to shape the research landscape, it is essential that we continuously assess the broader impact of this transformative technology and actively strive to use it responsibly in research and education.

***The need of alignment research and the application of behavioural ethics***

ChatGPT and other emerging machine learning technologies provide a significant opportunity for researchers. However, to ensure that our use of these technologies has us working towards discoveries that promote human flourishing, AI needs to be aligned to these goals.

Due to the inherent difficulty in specifying the full range of desired and undesired behaviours, this has sparked a new field called alignment research. This research is at the core of ensuring that the risks AI poses for humanity are kept at a minimum. AI can harm humans in two ways; by influencing other humans to commit unethical behaviour or as an enabler of unethical behaviour. To navigate the corrupting effects of AI, researchers have suggested the application of behavioural ethics, a field of social scientific research based on empirical observations. Adopting such an approach will provide better foundations to design evidence-based effective policies. However, currently, this research is primarily undertaken by the companies that design and sell AI, as a research community we must insist on this changing. An excellent review by Kobis *et al.* [7], evaluating the available research from behavioural science, human-computer interaction and AI research, provides a novel evidence-based discussion (Table 2).

### ***Where do we start in trying to understand the legal implications of ChatGPT?***

Whilst there is a consensus that AI tools like ChatGPT will increase efficiency and allow scientists to focus their enquiring minds on discovery, there are specific concerns of which to be mindful. The content shared in queries or prompts can become part of the data pool that companies like OpenAI use to train their generative models. As a user, one permits that use when you sign up. There are currently no restrictions placed on the future use of content entered as text prompts into any LLM platform. This raises the potential risk that once shared on these platforms, the information is considered a public disclosure and is no longer confidential.

Companies such as Samsung have banned the use of ChatGPT after engineers used it to fix proprietary code inadvertently disclosing a valuable trade secret[10]. Following this incident, Walmart, Amazon and Microsoft were amongst many companies who implemented similar bans, at least until guidelines are developed[11]. Many Australian universities have also established guidelines or policies around the use of generative AI software for teaching, learning and assessment. They all remind their students of the value of academic integrity, but they also are strong advocates for (careful) adoption of AI tools.

Other concerns are around posting personal data onto the platform, including information that could constitute a data breach and contravene privacy or medical laws. According to the OpenAI ChatGPT privacy policy, a user intending to post personal data is thus bound by the General Data Protection Regulation, and must include a *Data Processing Addendum* with the AI provider.

Questions have been raised about copyright ownership and moral rights' infringement in the use of LLMs and other AI-based tools. Copyright subsists in a work that has been generated by a human and must have a human author. Since ChatGPT is an AI language model, it is not capable of creating content on its own but rather responds to user prompts. Copyright subsists in PhD or Master's theses or any scientific articles that may become part of the ChatGPT data pool. Copyright owners are concerned that works may have been used to train or inform AI tools without their knowledge or permission. In the case of scientific work, this could mean fewer citations. Since citations play a crucial role in assessing research value and validity, this type of adoption by ChatGPT may be of concern.

Moral rights, such as the right to attribution and reputation, have also been a subject of concern. Well-known singer-songwriter, philosopher and Australian music icon, Nick Cave voiced concern over a fan's use of ChatGPT to produce lyrics "in the style of Nick Cave". Cave labelled the results a "grotesque mockery" of his lyrics and a "travesty"[12]. Couple these legal concerns with those raised earlier about inaccuracies and reliability and the message to the user is to be prudent and cautious. Be vigilant in what one posts, scrutinise outputs to verify accuracy and currency, and do not reproduce any third-party material in any way that may infringe copyright or raise moral rights issues.

Finally, we turn to resource concerns. Will we still need lawyers and IP professionals in the future? Many lawyers have test-driven ChatGPT with mixed results. When asked by the Dean at Suffolk Law School to find answers to several legal queries, the answers were "*imperfect and at times problematic. For example, the legal documents are incomplete, and its discussion of certain topics lacked nuance*"[11]. Law is, after all, complex, and each case unique. Users must be wary that the use of an AI tool for law-related services raises a host of regulatory and ethical issues, such as those relating to the unauthorised practice of law.

One also may be surprised to learn that ChatGPT can draft a patent specification. But whether it should be used to draft novel patent claims is contentious, as it is likely to miss the intricacies of the crux of the invention that require specialty content knowledge, as much as being potentially unaware of specific legal frameworks of the jurisdictions to which an application may apply. As data used by the GPT4 foundation model has a cut-off for September 2021, more recent "prior art" will be absent from its knowledge bank, unless links to other appropriate references are created. While this lack of currency and relevance can be addressed with development of bespoke future applications, the verdict for now is that intellectual property professionals will not be replaced by LLMs anytime soon.

### ***Should we be concerned about the safety and security of AI, and how do we use it responsibly?***

ChatGPT, along with other generative AI platforms, can incredibly powerful productivity tools. However, as with any new technology, we must ask questions about how safe our information is in these systems and how they might be used against us. We need to consider who is delivering our choice of AI tool and their privacy policies, as all platforms have the potential to use our inputs as prompts or feedback, to train their models. We must also be mindful that the websites, plugins or other tools we interact with may collect our inputs or have technical vulnerabilities that could lead to data breaches.

Another, more concerning area of security concern regarding generative AI is its potential misuse by malicious parties. Tools like ChatGPT are designed to finely craft passages of text. When this is combined with AI-generated images, video, or audio, new types of scams known as "deepfakes" are emerging that can be incredibly difficult to spot. The spread of misinformation through the misuse of generative AI tools is a real and present danger. In the future, public and private organisations also face the possibility of cyberattacks that may target AI models and assistants, that could use AI jailbreaking techniques to try and coax AI assistants into revealing sensitive information or acting on the attacker's behalf by biasing the behaviour of the model.

We will all need to stay vigilant and up to date on AI-based cyber threats – this is a whole new arena in cybersecurity and, like all other areas of AI, we still have a lot to learn about the technology's capabilities and vulnerabilities.

***What does intelligence mean in the context of ChatGPT, and can it be defined and measured?***

The rapid adoption of ChatGPT, the fastest of any technology in human history, is a testament to the reach and power of usable AI. However, many commentators (and the majority in this forum) focussed on its current limitations and problems. Recognising AI's flaws is important to establish clear moral frameworks for its use and to steer technological improvements. Nonetheless, these limitations should not dim our appreciation of the revolutionary advance it represents. While in-built safety features may prevent GPT4 from passing the Turing test, OpenAI's experiments strongly suggest that it can convincingly appear human-like to casual observers. Hence, it will not be long until we share our planet, for the first time, with true intelligence of non-organic matter.

The "intelligence" of AI, now and in the future, is likely to be profoundly different to our own. Geoffrey Hinton, winner of the 2018 Turing Award, points to the capacity and speed advantage AI has compared to humans, famously warning that "it's quite conceivable that humanity is just a passing phase in the evolution of intelligence"<sup>[13]</sup>. AI can learn instantaneously and at large scale. As it develops to also encompass the development of creativity, this different type of "task directed intelligence" may present new challenges and dangers to a society that will ultimately be forced to adopt to these changes. The workplace, in particular white-collar jobs, will radically alter with the adoption of this technology. Already IBM has put a stop to hiring in HR-related roles as the company expects that many of the tasks will be undertaken by AI in the very near future.

As with all truly revolutionary technologies, AI will become part of our daily lives and will therefore augment our own intelligence. It will also embed itself within our own knowledge network and, as such, may gain sentience, the ability to experience and feel on the same level as humans. This is not an outrageous claim; consider that sentience has not been linked to biological circuits and as such it is hard to construct an argument why it could not emerge in silicone circuits. Even though, according to most experts, the scenario of AI gaining sentience, if it ever does, is still in our distant future, we should get prepared now. What rights would sentient AI systems have? How should we behave towards them? Would we have a duty to promote and foster their existence? While the emergence of sentience in AI circuits is considered a distant possibility by most experts, it is a scenario that is worthy of contemplation using more comprehensive frameworks than those that have already been explored in science-fiction narratives.

These broad topics in Theme 2 focussed on the current limitations of AI, informed and uninformed preconceptions, and concerns about the potential loss of academic skills, IP and plagiarism. These are important points. But if we assume, for the moment, that they will be resolved, we must consider how scientists can best leverage AI to produce the highest-quality science. The challenge will be to explore what AI does well and how we can best use it.

***What will it mean to be a scientist in a world shaped by AI?***

## Capabilities and implications of AI in science

First, what are the primary ingredients of high-quality science and scientists? Every field has its own attributes. However, broadly speaking, these are our skills, knowledge, understanding and creativity (Figure 5). Knowledge and skills can include our fundamental expertise in the scientific method and specialist expertise within our chosen discipline. Our understanding includes a higher-order mental map of the principles, theories and knowledge of our field, the key gaps requiring further study, and how our and others' research fits within this. Creativity, in a broad sense, comprises those intuitive leaps and generation of novelty that add value.

If we accept this generalisation, then, for the foreseeable future, where is AI better than humans and where are humans better than AI? AI systems excel at storing, condensing and summarising knowledge. With self-directed learning, AI can increasingly complement or replicate human skills. However, AI lags behind humans in truly understanding this information. Similarly, although AI can create novel concepts, it falls far short of humans in judging or prioritising their quality. However, we may predict that, as AI improves, it will further improve in the automation of storing, summarising and recalling human knowledge and, as models are further "tuned" by human feedback, such systems will develop embedded features that will allow prioritisation of that knowledge. In addition, further development of AI capabilities may provide an excellent platform for consolidating and developing specialist skills and automating many aspects of their application when foundation systems can be tuned and personalised. The more this occurs, the harder it may be for scientists to differentiate themselves or lead their field primarily on their specialist knowledge or skills as the barrier to accessing these, can be lowered by the appropriate deployment of an AI model.

As AI's influence will undoubtedly grow, we scientists will likely place a greater value on understanding and creativity of "new" and cutting-edge knowledge, areas that AI models will find harder to synthesise and effectively validate. This perspective can help guide principles on how to leverage AI to automate our work and boost our productivity. However, even if we accept that AI can surpass some human capability for collating and synthesising knowledge at scale, these skills remain essential building blocks of human understanding and creativity. We do not have the luxury of giving over to AI the task of "base learning", leaving us to have all of the "fun" without having first put in the hard yards to further our understanding of a field. The best approaches will use AI to boost productivity and accelerate development, particularly where it supports and streamlines workflows to further our understanding of a given area and facilitating our creativity.

Moving forward, we will need to be wary of allowing AI to boost human productivity without also attending to our own growth or, worse, sacrificing human potential to AI models. Arguably, AI-based automation should be left to tasks for which we already have significant expertise, freeing us to engage more valuable pursuits. Leveraging AI well will allow us to produce astounding science. However, leveraging it poorly will risk impeding our development, leaving us less capable than peers who effectively use AI and potentially making us redundant to the AI itself.

### Concluding remarks

The WEHI ChatGPT symposium was the most well-attended event at Australia's foremost medical research institute in years. This was a testament to our collective, intuitive

understanding of the transformative potential of LLMs. As individuals, institutions and governments are racing to understand the capabilities of these models and how they will change work, play and society, two fundamental questions persist: What changes can we anticipate? And how can we effectively prepare for them?

It is likely that new tools will emerge between the time this article is written and the time you read it. Indeed, in the 20 days since the symposium, Google introduced Bard. Unlike the foundation model of GPT4, that can only access internet data from before September 2021, Bard can access data from Google's current results. OpenAI has subsequently released new GPT4 capabilities with their application programming interface to allow linkages to the Internet and other applications, while Microsoft has integrated GPT4 into its Bing search engine.

This prompts us to reflect on what might be released in another 20, 200 and 2000 days. Author Shalin Naik has already put Bard to the test, using it to draft text for his lab website. It did a pretty good job. However, like ChatGPT, Bard can also 'lie'. Apparently, Shalin has published three papers in Nature, Cell and Science in the last three years and was noted to have been born in Mumbai, India: significant achievements our Director would surely desire, but his mother would fervently dispute.

A common sentiment is that LLMs currently have more limitations than use and, despite significant advancements, are viewed as more of a party trick. Perhaps we need to add a 'for the moment' coda to the end of such pessimistic statements. We must remember that the current iterations have been developed to stochastically generate language, such that LLMs may indeed 'lie', for the moment. LLMs may not be able to effectively summarise papers or comprehensively find and prioritise accurate references, for the moment. LLMs cannot autonomously read the literature, generate a testable hypothesis, design and run a sensible experiment, make figures and draft a paper... for the moment. Every day, Twitter/X is flooded with new AI tools and applications which claim to address this functionality. As scientists, we are struggling to evaluate these tools before the next, more advanced iteration appears.

How then, should we navigate this constantly changing landscape? Do we attempt to halt the tidal wave or try to channel it with safety, ethics and equity at our core? Do we ban the use of LLMs in thesis writing and essay composition, or redefine the nature of what it means to be educated? If AI can someday generate testable hypotheses and conduct experiments, who is the scientist? Does it really matter? If we agree that science is the pursuit of knowledge and that human scientists are just one means of accomplishing that goal, perhaps we are just temporary custodians of knowledge gathering. Instead, there may be different ways for humans to engage their minds in the future. In the meantime, let's use AI to accelerate scientific discovery and bring everyone along for the ride, ethically and responsibly!

### **Author contributions**

Jessica G Borger\* - conceptualised, drafted, wrote and edited the manuscript

Ashley P Ng - wrote and edited the manuscript

Holly Anderton - wrote and edited the manuscript

George W Ashdown – wrote the manuscript

Megan Auld – edited the manuscript  
Marnie E Blewitt - wrote the manuscript  
Daniel V Brown - wrote the manuscript  
Melissa J Call - wrote the manuscript  
Peter Collins - wrote the manuscript  
Saskia Freytag - wrote the manuscript  
Leonard C Harrison - wrote the manuscript  
Eva Hesping - wrote the manuscript  
Jaci Hoysted – edited the manuscript  
Anna Johnston - wrote the manuscript  
Andrew McInnery - wrote the manuscript  
Phil Tang - wrote the manuscript  
Lachlan Whitehead - wrote the manuscript  
Aaron Jex - wrote and edited the manuscript  
Shalin H Naik\* - wrote and edited the manuscript

## CONFLICTS OF INTEREST

The authors declare not conflicts of interest.

## References

1. (WEHI), W.E.H.I. [cited 2023 June 3rd]; Available from: <https://www.wehi.edu.au/>.
2. Xu YJ, Xin Liu, Cao X, Huang CP, *et al.* Artificial intelligence: a powerful paradigm for scientific research. *Innovation* 2021; **2**: e10017
3. Vincent, J. OpenAI has published the text-generating AI it said was too dangerous to share. 2019 [cited 2023 June 3rd]; Available from: <https://www.theverge.com/2019/11/7/20953040/openai-text-generation-ai-gpt-2-full-model-release-1-5b-parameters>.
4. Federspiel F, Mitchell R, Asokan A, Umana C, McCoy D. Threats by artificial intelligence to human health and human existence. *BMJ Glob Health* 2023; **8**: e010435.
5. Zawacki-Richter O, Marin VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education – where are the educators? *Int J Educ Technol High Educ* 2019; **16**: e39 .
6. Birhane A, Kasirzadeh A, Leslie D, Wachter S. Science in the age of large language models. *Nat Rev Phys* 2023; **5**: 277–280.
7. Kobis N, Bonnefon JF, Rahwan I. Bad machines corrupt good morals. *Nat Hum Behav* 2021; **5**: 679–685.
8. Piccolo SR, Denny P, Luxton-Reilly A, Payne S, Ridge PG, Many bioinformatics programming tasks can be automated with ChatGPT. *arXiv* 2023. doi.org/10.48550/arXiv.2303.13528
9. Shue E, Liu L , Li B, Feng Z, Li X, Hu G. Empowering Beginners in Bioinformatics with ChatGPT. *bioRxiv* 2023. doi.org/10.1101/2023.03.07.531414
10. Ray S. Samsung Bans ChatGPT Among Employees After Sensitive Code Leak. *Forbes* 2023 [accessed 3 June 2023]

- <https://www.forbes.com/sites/siladityaray/2023/05/02/samsung-bans-chatgpt-and-other-chatbots-for-employees-after-sensitive-code-leak/?sh=559a51a60781>.
11. Sabin S. Companies are struggling to keep corporate secrets out of ChatGPT. Axios 2023 [accessed 3 June 2023] <https://www.axios.com/2023/03/10/chatgpt-ai-cybersecurity-secrets>
  12. Savage M. Nick Cave says ChatGPT's AI attempt to write Nick Cave lyrics 'sucks'. BBC News 2023 [accessed 3 June 2023] <https://www.bbc.com/news/entertainment-arts-64302944>
  13. Brown S. Why neural net pioneer Geoffrey Hinton is sounding the alarm on AI. MIT Management Sloan School 2023 [accessed 3 June 2023] <https://mitsloan.mit.edu/ideas-made-to-matter/why-neural-net-pioneer-geoffrey-hinton-sounding-alarm-ai>

## FIGURE CAPTIONS

**Figure 1.** Network architecture diagram of a "Deep Learning" network that takes a sequence of eight inputs and predicts one output (for example, the first part of a sentence), transforms through four "hidden layers" to predict the next word. This network is relatively simple and contains only 43 nodes (open circles), but even so, contains over 300 parameters. This gives some indication of the size and shape of the GPT models. (Figure generated using NN-SVG <https://alexlenail.me/NN-SVG>)

**Figure 2:** ChatGPT is a multilingual chatbot supporting over 50 languages.

**Figure 3.** ChatGPT can aid the email writing process through simple instructions by providing clear and human-like text suggestions.

**Figure 4.** Using code written by ChatGPT 3.5 for Python, users can import all relevant libraries, and their imaging data. Through some optimisation this data (Original Image) can be thresholded to detect individual cells (Threshold image) and segmented (Segmented image) successfully. However, more complex tasks requiring some degree of problem-solving become challenging, for example, tracking segmented cells through time (Tracking image).

**Figure 5.** Using AI strategically and ethically to produce high-quality science and outstanding scientists. **(a)** Like any disruptive technology, AI presents a classic dilemma around whether and when to allow or prohibit its use. Our view is that the best incorporation merges the strengths of AI with the capabilities of high-quality scientists to produce the most impactful science. **(b)** AI clearly outperforms humans in its ability to retrieve and synthesise skills and knowledge but lags well behind in demonstrating a qualitative understanding or engaging in high-value creative thought. As conventional skills and knowledge are increasingly automated, the capability of future scientists to contribute outstanding understanding and creativity is likely to be of increasing value. **(c)** Considering this, in incorporating AI into routine science, the primary question for each scientist is to subjectively consider where they can leverage the increased productivity that AI provides without compromising or, best of all, while enhancing their abilities and accelerating their development.

**Table 1.** AI technologies and related tools to support academic researchers and educators

Research task	AI tool/platform	Use
Reference Management	Zotero Mendeley EndNote	Organising, storing and citing academic papers to make it easier to manage references and bibliographies
Reading articles	ChatPDF Scholarcy Explainpaper	Summarise and answers questions about research PDF documents, breaking them down into bite-sized sections and highlighting key information
Data Analysis	IBM SPSS R Pandas NumPy	Support researchers in processing large datasets, perform statistical analyses and visualise results
Language Translation and NLP	Google translate NLTK spaCy	Globalisation of research has been supported through translation tools to breakdown language barriers, and NLP framework enable researchers to analyse and extract insights from textual data
Image and Video Analysis	OpenCV TensorFlow	Allow for automatic recognition of objects, patterns and anomalies in visual data
Creating images	DALL-E2 Cariyon	Create realistic images from a description in a natural language and text prompts
Collaboration	Microsoft teams Slack Google Workspace	Facilitation of communication, file sharing and real-time collaboration
Survey and Experiment Design	Elicit Qualtrics SurveyMonkey	AI-incorporation optimises survey questions, analyse responses and derive meaningful insights from data
Literature Reviews	Semantic Scholar Iris.ai Research rabbit R Discovery	Assist researchers in quickly finding relevant academic papers, identify key insights and explore connections between different studies
Data visualisation	Tableau Power BI	Create visually compelling representations of data and make complex findings more accessible to a broader audience
Proofreading and editing	Grammarly	Enhance quality of academic writing and ensure proper grammar, punctuation and style
Plagiarism detection	Turnitin iThenticate Copyscape	Maintain academic integrity by identifying instances of potential plagiarism

**Table 2.** Expert reviews on the application of AI in higher education, scientific and medical research and ethical considerations of AI in science

<b>Topic</b>	<b>Article Summary</b>	<b>Reference</b>
Scientific research	Provides a comprehensive survey and guidelines on the development, integration and application of AI in different aspects of fundamental sciences, including medical science discussing challenges and the potential of AI techniques to handle these challenges, to help motivate researchers to understand the applications of AI platforms and tools and thereby promote the continuous development of these fundamental sciences.	[2]
Medical research and human health	A critical discussion on how AI poses a number of threats to human health and well-being and the need for effective regulation in access and application of certain types of AI	[4]
Higher Education	A systematic review on research publications on AI applications in higher education discussing the use of AI in (1) profiling and prediction, (2) assessment and evaluation, (3) adaptive systems and personalisation, (4) intelligent tutoring systems, concluding that all areas have limited understanding and reflection of the challenges and risks posed by AI, that also lack pedagogical considerations	[5]
Ethics	Interview with 4 experts in AI ethics and policy discussing the potential risks in science with the rapid emergence of LLMs in recent years and the considerations to apply the ensure good scientific practice and trust in science	[6]
Alignment research	Evaluation of current available evidence aligning behavioural science, human-computer interaction and AI research, with a proposed research agenda outlining perspectives on how to gain behavioural insights for better AI oversight.	[7]

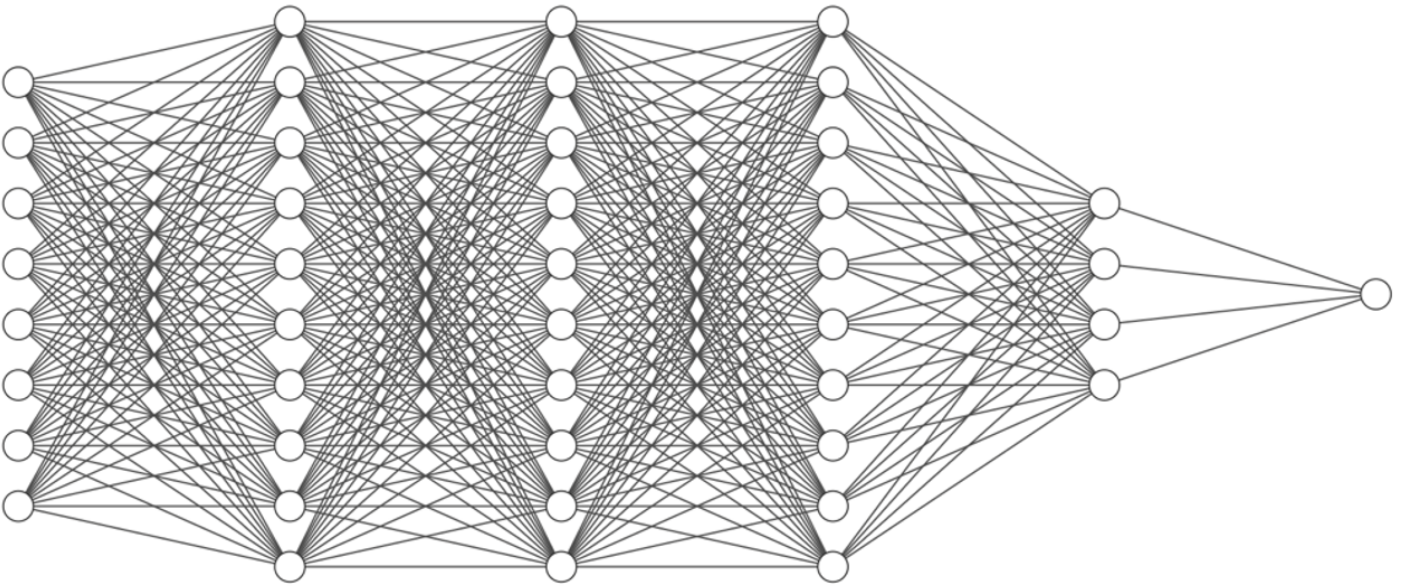


Figure 1.png

**"Together, we are striving to eradicate malaria"**

"معاً، نحن نسعى للقضاء على الملاريا" - Arabic

"我们一起努力消除疟疾" - Mandarin

"Gemeinsam streben wir danach, Malaria auszurotten" - German

"Insieme stiamo lavorando per eradicare la malaria" - Italian

Figure 2.png

E

I want to criticise my colleague for doing something wrong without being rude. I also want to give some suggestions how to perform better next time. How do I say that in an email?  
Australian cultural context



When giving constructive criticism to a colleague in an Australian cultural context, it's important to be direct and respectful in your communication. Here is an example of an email you could write:

Subject: Feedback on Project X

Hi [colleague's name],

I wanted to give you some feedback on Project X. First of all, I want to say that I appreciate the effort you put into it. However, I noticed that [describe the issue/problem]. While I understand that mistakes can happen, I wanted to bring it to your attention so that we can avoid similar issues in the future.

Moving forward, I think it would be helpful if [provide suggestions for improvement]. By doing so, we can [state the benefits of implementing your suggestions].

Please don't hesitate to reach out if you have any questions or if you need any support. I'm looking forward to working together to make our projects even better.

Best regards,

[Your name]

Figure 3.png

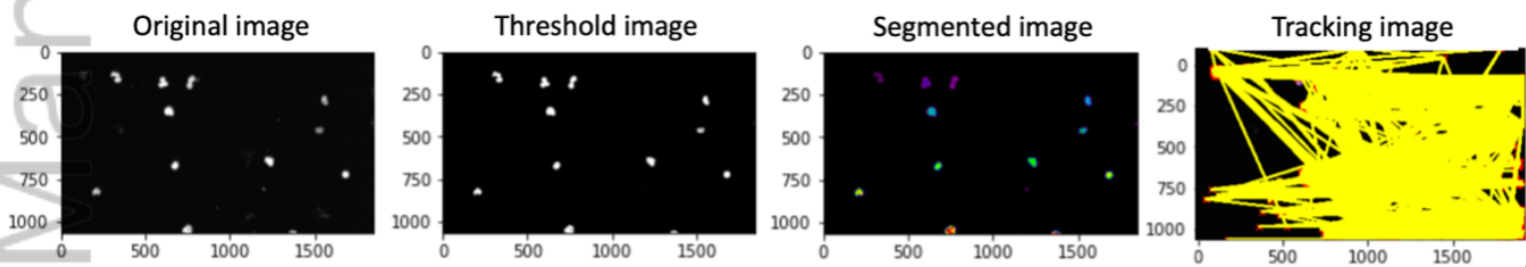


Figure 4.png

