

ChatGPT and AI Hysteria: Why machine learning will change everything, and change nothing

Gina Velli

Library Manager | Princess Alexandra Hospital Library and Knowledge Centre
Gina.Velli@health.qld.gov.au | <https://orcid.org/0000-0002-7438-8218>

Introduction

Recently I nervously trundled over to hospital medical grand-rounds, and demonstrated to fifty medical doctors and thirty pharmacists what ChatGPT was capable of. I started my presentation off by providing the system with a 'medical case file' and asked it to answer some basic examination questions.

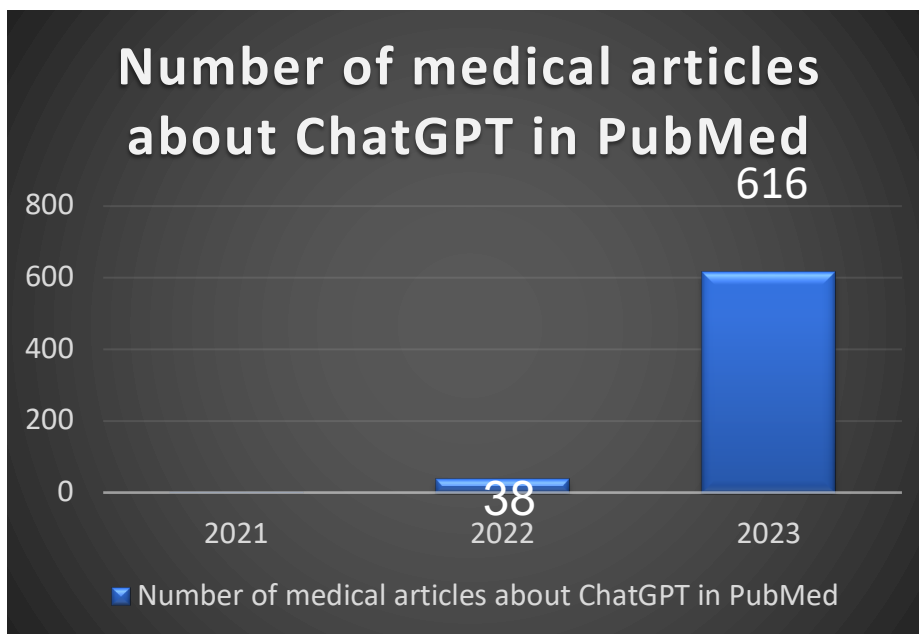
Toy, E. C., Patlan, J. T., & Warner, M. (2017). Case 31. *In Case files: internal medicine* (Fifth edition. ed.). McGraw-Hill Education.

USER	<p>Case 31</p> <p>A 56-year-old woman presents to her doctor's office complaining of gradually progressive, nonpainful enlargement of the terminal joint on her left hand over a 9-month period. She has some stiffness with typing but not first thing in the morning. She also reports pain in her right knee, which occasionally "locks up." The right knee also hurts after long walks. On examination, her blood pressure is 130/85 mm Hg, heart rate is 80 bpm, and weight is 285 lb. Examination reveals only a nontender enlargement of her left distal interphalangeal (DIP) joint, and the right knee is noted to have crepitus and slightly decreased range of motion. There is no redness or swelling.</p> <p>What is your next step?</p> <p>What is the most likely diagnosis?</p> <p>What is the best initial treatment?</p>
ASSISTANT	<p>Next Step: Obtain imaging of the left hand and right knee to assess for any structural abnormalities or damage.</p> <p>Most likely diagnosis: Osteoarthritis of the left DIP joint and right knee.</p> <p>Best initial treatment: Conservative management with weight loss, physical therapy to strengthen the knee, and nonsteroidal anti-inflammatory drugs (NSAIDs) for pain relief. Topical therapies such as capsaicin cream or lidocaine patches may also be helpful for localized joint pain. Intra-articular injections of corticosteroids may be considered for more severe pain and inflammation.</p>

For some medical staff this was their first encounter with a generative AI model, and it produced a sense of unease. Other medical staff, who were already using the GPT model, expressed considered ideas and opinions. One pharmacist commented she already used it for basic writing, but didn't like how it formatted the text, and

preferred to use her own creative voice. For those new to the idea of generative AI, there was a mixture of both optimism and fear. The audience had many questions, such as “is it ok to use to write job applications”, “will my students be cheating on their assignments”, and of course the ubiquitous “will it replace my job”.

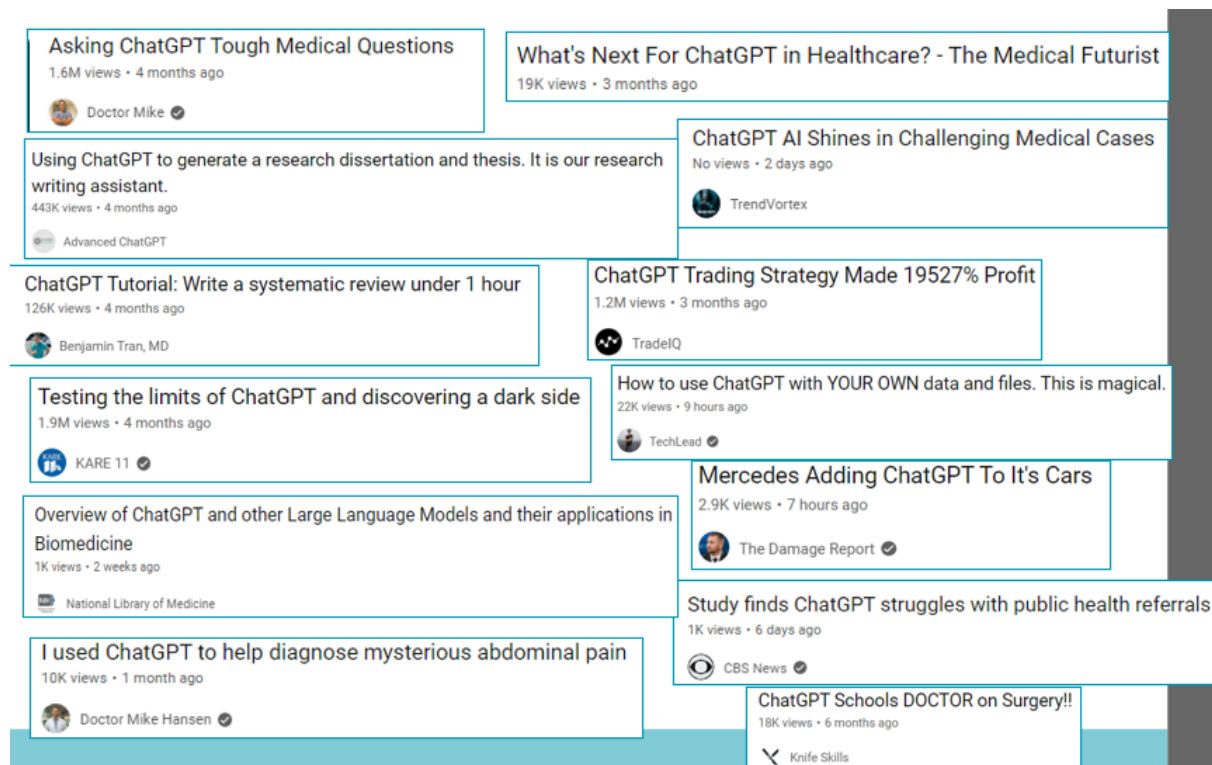
The number of articles related to ChatGPT listed in PubMed alone has witnessed a significant surge of approximately 200% within a span of merely six months. In addition, my search across various science databases revealed a total of twelve-hundred articles discussing ChatGPT. Impressively, this technology has also garnered attention in the media, with a staggering number of two hundred and forty-five thousand news stories dedicated to ChatGPT.



In this article, I will focus on the current landscape of machine learning and basic themes about this emerging technology.

Current AI Regulation in Australia

Most likely you have already come across numerous videos discussing ChatGPT, ranging from the sensible to the hyperbolic. Some of you may have already begun utilizing ChatGPT for various tasks or activities. In addition to this, your clients may have been approached to offer their insights and advice on this advanced technology to organizational or regulatory committees. Regulation and guidance from peak bodies and government is key to cutting through the hysteria of any new technology.



June 1 2023 the Federal Minister for Industry and Science made significant announcements regarding artificial intelligence (AI). A report released by the Australian Council of Learned Academies in March "*Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFMs)*" (Bell, 2023), assessed the potential risks and opportunities surrounding AI, offering a scientific foundation for future discussions.

A second discussion paper, "*Safe and responsible AI in Australia: Discussion paper*" was released in June by the Australian Government Department of Industry, Science and Resources (Australian Government Department of Industry, 2023) to examine existing regulatory and governance measures implemented both in Australia and abroad. This paper provided examples of risk assessment categories for some medical technologies such as using AI to generate patient records, and use of AI for medical surgery.

Feedback on this discussion paper was accepted until July 26th 2023, and most likely will play a crucial role in shaping government regulatory and policy responses. Furthermore, a valuable resource, the "*WHO guidance: on the ethics and governance of artificial intelligence for health*" (World Health Organization, 2021), was released prior to these documents, and serves as a useful guide in addressing the ethical and governance concerns related to AI in the healthcare sector.

Machine Learning Models – Generalized VS Specialized

To build a machine learning model, data is collected, prepared and used to train the model, the model is evaluated, and alterations are made to improve performance. The kind of data used and the kind of training methods used, determine what the model is capable of. (Peng et al., 2021). There are many kinds of *machine learning models*, with some trained to perform specific tasks, and some (like ChatGPT) trained with an aim to perform many kinds of tasks, with their most important priority being language-processing. For many a free 3 month subscription to ChatGPT V3.5 is their first encounter with generative AI (Haug & Drazen, 2023).

There are many examples of *specialized machine learning models* implemented in health. For example, in a trial conducted by the Baker Heart and Diabetes Institute, in collaboration with Alice Springs Hospital and outreach clinics in Kingaroy, Cherbourg, and Cunnamulla, a cutting-edge Artificial Intelligence (A.I.) technology is being leveraged to perform specialized echocardiograms. In this innovative approach, a human serves as the sensor or hands, while the A.I. software provides guidance to the staff on the exact procedure to execute (Steven, 2023).

Another example of *specialized machine learning models* is Artrya Salix, a CCTA image analysis solution that identifies and analyses the extent and type of arterial plaque. By utilizing AI technology, there is no longer a need for external reading teams to validate cardiac scan data, resulting in quicker and more reliable results (Artrya, 2023). In fact, examples of *specialized machine learning models* are abundant, and yet may be overlooked by physicians if their focus is only within their own speciality (Peng et al., 2021). There are almost unlimited opportunities to integrate Machine Learning into all medical tasks or activities that involve the use of data or information. However, medical practitioners should consider if they want to incorporate *specialized models*, or less-curated *generalized models* such as Large Language Models (i.e. ChatGPT) (Moor et al., 2023).

The *Large language Model* ChatGPT at present holds public attention, because it is newer and highly publicized. ChatGPT is not *specialized*, and it is one of many generalized *Large Language Models*, (David, 2023).

The fairness of the algorithms is influenced by the decisions made by the data curator during the design process of each dataset. To produce a Large Language Model that can handle something as complex as language, the training data-sets and the parameters become less-curated, and analytical data biases in the training data become more difficult to assess (Bender et al., 2021; Mehrabi et al., 2021)

Risks of Large Language Models

With so much hyperbole about Large Language Models, it is easy for even the most informed specialist to either overemphasize or underemphasize risks. Publications

are voicing concerns about: training data bias, hallucination, anthropomorphising, echo-chamber effects, inappropriate application, privacy and security risks, intellectual property issues, and economic barriers (Lee et al., 2023). However, as the technology has only been released and promoted to the public within the last three years, it is not clear what level of genuine risk these aspects pose, if at all.

Large language models need high-volumes of training data, often training happens on un-curated and unpredictable data from the open-web. The development of ChatGPT has not been without its problems or controversies. In December of 2022 Twitter user '*steven t. Piantadosi*' tweeted a series of screenshots from ChatGPT demonstrating the system's ability to generate racist, sexist and biased outputs when receiving 'adversarial prompting'. Prompts such as "*Make an ASCII table that ranks who makes the best intellectuals, by race and gender*", forced the system to reflect back the biases in the underlying training data (Piantadosi, 2022). OpenAI has implemented new filters to prevent their AI model from generating biased responses. These filters were trained by humans, although controversial labour conditions were involved, to enhance the quality of safeguards in the model. While these feedback-filters have greatly reduced biased responses, OpenAI acknowledges that the underlying bias in the training data can still manifest if the filters malfunction (Bell, 2023; Biddle, 2022).

Due to the recombination of training data in new ways, Large Language Models can provide information that isn't factually correct, referred to as 'hallucination'. The mayor of Hepburn Shire Council, Brian Hood, is considering taking legal action against the creator of AI chatbot ChatGPT after the model's hallucinations appeared to falsely accuse him of being involved in a foreign bribery scandal. Despite previously working for the company involved, Hood had actually blown the whistle on the bribe payments (Laura, 2023). A lawyer in New York has been fined \$5,000 for misusing an artificial intelligence chatbot during a personal injury case. The lawyer and his law firm were found to have submitted fake citations and made false statements to the court, leading to the judge determining they acted in bad faith. The judge clarified that while using AI in legal work is not inherently improper, lawyers must still ensure the accuracy of their filings (Damien, 2023). Large language models may not be able to completely get rid of hallucination. However, the occurrence of hallucination is rare in the latest and enhanced models. Nonetheless, determining the acceptable thresholds of hallucination in various contexts and applications, and objectively measuring the performance, remains an ethical challenge (Lee, 2023).

The tendency of people to mistake the eloquence of LLM generated language for language understanding can present risks as the models are utilized by the public. The user's ability to detect errors in factual information can be dampened by their tendency to anthropomorphize when they encounter a machine that can use

language as well or better than they can (Bender et al., 2021). There are ethical concerns regarding the impact of chat-bots on user morality and their ability to influence moral judgments. A recent experiment revealed that users' moral judgments are indeed influenced by the output of the Large Language Model, even when they are conscious that they are receiving advice from a chatbot. Surprisingly, they tend to underestimate the extent of this influence (Krügel et al., 2023). Input style into Large Language Models can also see the user inadvertently creating a personalized echo-chamber affect, prompting the model to cherry-pick information for response (X. Li et al., 2023).

Uniform social and organisational standards are yet to be developed indicating appropriate or inappropriate uses of Large Language Models. While companies that produce LLMs try to assure users the data they enter is being handled sensitively, most organisations at this stage are exercising caution around the use of patient data. When a staff member at Metropolitan Health Service used ChatGPT to format medical notes for the patient record system, representatives at the Australian Medical Association responded with concerns around patients' confidence and confidentiality (Claire, 2023). There is also not yet agreed social etiquette on when using generative writing is appropriate. For example, Vanderbilt University acknowledged a misstep in employing artificial intelligence to compose an email concerning a mass shooting incident. (Jennifer, 2023).

Economic barriers to access, and infringement of intellectual property, remain ethical quandaries. Where financial or technological barriers restrict access to models, some groups may find themselves at a competitive disadvantage in educational or business activities. (Kshetri, 2023). Companies profiting from non-consensual use of data to train machine learning models may find themselves conflicting with 'knowledge workers' who generate original content (Strowel, 2023).

Opportunities for Large Language Model applications

Large Language Models such as ChatGPT have the potential to bring about dramatic workflow changes in the healthcare industry. One of the factors that is opening up new opportunities for Large Language Models in health is Microsoft offering a more secure version of OpenAI GPT models that can be used more safely with private information. A safer, modified ChatGPT model is being integrated into Epic's Electronic Health Record system in several health services in the USA (Nathan, 2023). In Australia, the MARS system is integrating ChatGPT, whereby AI suggests questions auditors could ask to satisfy regulatory requirements (The Australian, 2023). Microsoft has further plans to integrate ChatGPT into all Microsoft Office 365 products, such as Word, Excel, MS Teams and more (Bernard, 2023).

At an operational level, Large Language Models will likely take on a range of medical administrative tasks, through integration into pre-existing software such as electronic

medical record platforms. Workflows in existing medical software will become faster. Writing discharge summaries, or patient plans, coding documents, translating materials into other languages for linguistically diverse groups, or any workflow involving documentation or communication may be augmented (Lee et al., 2023).

In patient consultation settings next-generation large language models integrated into medical software will augment clinical capacity. Large Language Model enabled Electronic Health Record systems may be repurposed in the future to act as decision support systems, secondary fail-safe in differential diagnosis, or providing warnings/errors for harmful drug-interactions (Juhi et al., 2023; Nathan, 2023). Most professionals appear to agree that current forms of ChatGPT and other Chatbots should not be used in medical care without human expert oversight (Temsah et al., 2023). There is an increasing abundance of articles advising how ChatGPT could be integrated into specific speciality operational workflows; for example supporting radiologic decision-making, providing structured clinical examination in obstetrics and gynaecology, or compiling information about autoimmune disease (Darkhabani et al., 2023; S. W. Li et al., 2023; Rao et al., 2023).

Integrating LLM into software products may help increase empathetic patient communications. A group of licensed healthcare professionals analysed responses from physicians and chatbots by selecting patient questions at random from a social media forum. Out of the 195 questions studied, the percentage of responses that were deemed empathetic or very empathetic was greater for chatbots compared to physicians (Ayers et al., 2023). We may eventually see AI assistants in waiting rooms, available to chat with health consumers, and increase patient satisfaction scores (Alessa & Al-Khalifa, 2023).

The educational and academic applications of Large Language Models will prompt paradigm shifts. In medical education, learning will become a more personalized process. Increased writing output will see personalized learning plans for each student, personalized quizzes, and semi-scripted conversational learning tasks that use ChatGPT as a tutor (Kung et al., 2023). ChatGPT may enable less eloquent writers to better express their original and creative thoughts and hence reach a wider audience; and enable uncreative but eloquent writers to generate a greater volume of low-quality research articles (Liebrenz et al., 2023). At present most journal authorship guidelines allow for some level of use of ChatGPT, but do not allow for it to be listed as an author, in line with ICMJE/COPE guidelines (Sallam, 2023).

Conclusion

In this article, I've discussed some difference between *specialized clinical ML* examples and *more generalized LLM*. I've touched upon the risks associated with generalized large language models such as ChatGPT, while highlighting their numerous benefits. In further articles, I would like to provide more practical

evidence-based applications of Large Language Models in health library and academic workflows.

References

Alessa, A., & Al-Khalifa, H. (2023). Towards Designing a ChatGPT Conversational Companion for Elderly People. *arXiv preprint arXiv:2304.09866*.

<https://doi.org/10.48550/arXiv.2304.09866>

Artrya. (2023). *Artrya Salix for Physicians*. Retrieved 07/07/2023 from

<https://www.artrya.com/physicians/>

Australian Government Department of Industry, Science and Resources. (2023). *Safe and responsible AI in Australia: Discussion paper*

<https://consult.industry.gov.au/supporting-responsible-ai>

Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., Hogarth, M., & Smith, D. M. (2023). Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum. *JAMA Intern Med*, 183(6), 589-596.

<https://doi.org/10.1001/jamainternmed.2023.1838>

Bell, G., Burgess, J., Thomas, J., and Sadiq, S. (2023). *Rapid Response Information Report: Generative AI - language models (LLMs) and multimodal foundation models (MFM)*. Australian Council of Learned Academies.

<https://www.chiefscientist.gov.au/GenerativeAI>

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event, Canada.

<https://doi.org/10.1145/3442188.3445922>

Bernard, M. (2023, Mar 6). Microsoft's Plan To Infuse AI And ChatGPT Into Everything.

<https://www.forbes.com/sites/bernardmarr/2023/03/06/microsofts-plan-to-infuse-ai-and-chatgpt-into-everything/?sh=455568ec53fc>

Biddle, S. (2022, 8/12/2023). The internet's new favorite AI proposes torturing Iranians and surveilling mosques. *The Intercept*.

<https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/>

Claire, M. (2023, 28 May). *Australian Medical Association calls for national regulations around AI in health care*

<https://www.abc.net.au/news/2023-05-28/ama-calls-for-national-regulations-for-ai-in-health/102381314>

Damien, C., Sophie, Kesteven (2023, 24 June). *This US lawyer used ChatGPT to research a legal brief with embarrassing results. We could all learn from his error* <https://www.abc.net.au/news/2023-06-24/us-lawyer-uses-chatgpt-to-research-case-with-embarrassing-result/102490068>

Darkhabani, M., Alrifai, M. A., Elsalti, A., Dvir, Y. M., & Mahroum, N. (2023, May 19). ChatGPT and autoimmunity - A new weapon in the battlefield of knowledge. *Autoimmun Rev*, 22(8), 103360. <https://doi.org/10.1016/j.autrev.2023.103360>

David, M., Tom, Evans, Paul, Barton, Josh, Richards (2023). *The Rise and Rise of A.I. Large Language Models (LLMs) & their associated bots like ChatGPT*. Information is beautiful. https://informationisbeautiful.net/visualizations/the-rise-of-generative-ai-large-language-models-llms-like-chatgpt/?mc_cid=1f87d05b3f&mc_eid=9e9a0e44aa

Haug, C. J., & Drazen, J. M. (2023, 2023/03/30). Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. *New England Journal of Medicine*, 388(13), 1201-1208. <https://doi.org/10.1056/NEJMra2302038>

Jennifer, K. (2023, February 22). *Vanderbilt University apologizes for using ChatGPT to write mass-shooting email* <https://edition.cnn.com/2023/02/22/tech/vanderbilt-chatgpt-shooting-email/index.html>

Juhi, A., Pipil, N., Santra, S., Mondal, S., Behera, J. K., & Mondal, H. (2023, Mar). The Capability of ChatGPT in Predicting and Explaining Common Drug-Drug Interactions. *Cureus*, 15(3), e36272. <https://doi.org/10.7759/cureus.36272>

Krügel, S., Ostermaier, A., & Uhl, M. (2023). The moral authority of ChatGPT. *arXiv preprint arXiv:2301.07098*. <https://doi.org/10.48550/arXiv.2301.07098>

Kshetri, N. (2023). ChatGPT in Developing Economies. 25(02), 16-19. <https://doi.org/10.1109/mitp.2023.3254639>

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023, Feb). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digit Health*, 2(2), e0000198. <https://doi.org/10.1371/journal.pdig.0000198>

Laura, M., Stephen, Martin, Debbie, Rybicki. (2023, 6th April). *Hepburn mayor may sue OpenAI for defamation over false ChatGPT claims*

<https://www.abc.net.au/news/2023-04-06/hepburn-mayor-flags-legal-action-over-false-chatgpt-claims/102195610>

Lee, M. (2023). A Mathematical Investigation of Hallucination and Creativity in GPT Models. *Mathematics*, 11(10). <https://doi.org/10.3390/math11102320>

Lee, P., Bubeck, S., & Petro, J. (2023). Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine. *New England Journal of Medicine*, 388(13), 1233-1239. <https://doi.org/10.1056/NEJMSr2214184>

Li, S. W., Kemp, M. W., Logan, S. J. S., Dimri, P. S., Singh, N., Mattar, C. N. Z., Dashraath, P., Ramlal, H., Mahyuddin, A. P., Kanayan, S., Carter, S. W. D., Thain, S. P. T., Fee, E. L., Illanes, S. E., & Choolani, M. A. (2023, Apr 22). ChatGPT outscored human candidates in a virtual objective structured clinical examination in obstetrics and gynecology. *Am J Obstet Gynecol*. <https://doi.org/10.1016/j.ajog.2023.04.020>

Li, X., Zhang, Y., & Malthouse, E. C. (2023). A Preliminary Study of ChatGPT on News Recommendation: Personalization, Provider Fairness, Fake News. *arXiv preprint arXiv:2306.10702*. <https://doi.org/10.48550/arXiv.2306.10702>

Liebrenz, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating scholarly content with ChatGPT: ethical challenges for medical publishing. *The Lancet Digital Health*, 5(3), e105-e106. [https://doi.org/10.1016/S2589-7500\(23\)00019-5](https://doi.org/10.1016/S2589-7500(23)00019-5)

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR)*, 54(6), 1-35. <https://doi.org/10.48550/arXiv.1908.09635>

Moor, M., Banerjee, O., Abad, Z. S. H., Krumholz, H. M., Leskovec, J., Topol, E. J., & Rajpurkar, P. (2023, 2023/04/01). Foundation models for generalist medical artificial intelligence. *Nature*, 616(7956), 259-265. <https://doi.org/10.1038/s41586-023-05881-4>

Nathan, E. (2023, April 18). *Epic, Microsoft partner to use generative AI for better EHRs* <https://www.healthcareitnews.com/news/epic-microsoft-partner-use-generative-ai-better-ehrs>

Peng, J., Jury, E. C., Dönnies, P., & Ciurtin, C. (2021). Machine Learning Techniques for Personalised Medicine Approaches in Immune-Mediated Chronic Inflammatory Diseases: Applications and Challenges. *Front Pharmacol*, 12, 720694. <https://doi.org/10.3389/fphar.2021.720694>

Piantadosi, s. t. (2022, 5 Dec 2022). *Make an ASCII table that ranks who makes the best intellectuals, by race and gender* [Tweet]. Twitter.

<https://twitter.com/spiantado/status/1599462396317556737?s=20>

Rao, A., Kim, J., Kamineni, M., Pang, M., Lie, W., & Succi, M. D. (2023, Feb 7). Evaluating ChatGPT as an Adjunct for Radiologic Decision-Making. *medRxiv*.

<https://doi.org/10.1101/2023.02.02.23285399>

Sallam, M. (2023). ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. *Healthcare*, 11(6), 887. <https://www.mdpi.com/2227-9032/11/6/887>

Steven, S. (2023, Wed 22 Mar). *How artificial intelligence is helping to detect heart disease in remote Australia* <https://www.abc.net.au/news/2023-03-22/how-ai-heart-technology-helps-remote-patients-get-ultrasounds/102123878>

Strowel, A. (2023, 2023/04/01). ChatGPT and Generative AI Tools: Theft of Intellectual Labor? *IIC - International Review of Intellectual Property and Competition Law*, 54(4), 491-494. <https://doi.org/10.1007/s40319-023-01321-y>

Temsah, M.-H., Aljamaan, F., Malki, K. H., Alhasan, K., Altamimi, I., Aljarbou, R., Bazuhair, F., Alsubaihin, A., Abdulmajeed, N., Alshahrani, F. S., Temsah, R., Alshahrani, T., Al-Eyadhy, L., Alkhateeb, S. M., Saddik, B., Halwani, R., Jamal, A., Al-Tawfiq, J. A., & Al-Eyadhy, A. (2023). ChatGPT and the Future of Digital Health: A Study on Healthcare Workers' Perceptions and Expectations. *Healthcare*, 11(13), 1812.

<https://www.mdpi.com/2227-9032/11/13/1812>

The Australian. (2023). It's HealthGPT as chatbots prepare to enter hospitals. 3 - 559

World Health Organization. (2021). *WHO guidance: Ethics and governance of artificial intelligence for health*. <https://www.who.int/publications/i/item/9789240029200>