Tidsskriftet

Can AI be empathetic?

NIKOLAI OLAVI CZAJKOWSKI

n.o.czajkowski@psykologi.uio.no
Nikolai Olavi Czajkowski, associate professor at PROMENTA Research Centre, Department of Psychology, University of Oslo, and researcher at the Norwegian Institute of Public Health
The author has completed the ICMJE form and declares no conflicts of interest.

When healthcare personnel evaluate answers to questions about illness and symptoms, provided by either doctors or AI, they seem to prefer the latter.

The study by Mork et al. published in this edition of the Journal of the Norwegian Medical Association is based on 192 health-related questions from the website Studenterspør.no (1). GPT-4 was used to generate new answers to questions that had already been answered by doctors. A total of 344 respondents with a background in health care then evaluated the responses based on three criteria: empathy, quality of information and helpfulness, without knowing the source of the responses. The results showed that the responses generated by GPT-4 were considered more informative, helpful and empathetic than those from the doctors.

Many became aware of artificial intelligence (AI) through the launch of ChatGPT. Large language models like this are an example of generative AI, a technology that not only classifies and predicts data but also generates new content based on the examples it has been trained on. We are surrounded by AI-generated text, images, audio and video that are becoming increasingly difficult to distinguish from content produced by humans. The anticipated broad societal impact of generative AI was the driving force behind our national AI strategy, which was presented in January 2020 (2). It would be extremely difficult to find a more widely predicted technological revolution. Yet, even some of the developers have been surprised by how quickly the technology has evolved, driven by a huge increase in computing power and the availability of training data.

«It would be extremely difficult to find a more widely predicted technological revolution. Yet, even some of the developers have been surprised by how quickly the technology has evolved»

ChatGPT's architecture and the data used in its training are not publicly known, but a large portion of all text produced by humans is known to be used in the training. The fact that the model is well-informed is therefore not particularly surprising, and it has already been shown to perform at a high level in fields such as biology and law (3). What may be more remarkable, however, is that such models are also capable of providing answers that are perceived as helpful and, not least, empathetic.

This is also an important aspect of the revolution we are witnessing. AI has proven to be brilliant, not only in logic-based strategy games like chess and Go but also in navigating complex social situations (4). It has demonstrated an impressive ability to understand the perspectives of others, often better than most humans (5, 6). This may explain why GPT-4 is perceived as empathetic to people's concerns in the study by Mork et al. However, it is important to remember that ChatGPT does not actually feel genuine empathy, even though it may appear that way. How this apparent empathy will affect patients is currently unknown. Can it improve communication, or are we at risk of blurring the line between simulated and real care?

«The perception that ChatGPT is informative, helpful and empathetic in health counselling and advice does not mean that the technology is ready to replace healthcare personnel in patient care»

The power of these models brings with it significant ethical challenges. AI models like ChatGPT are known to 'hallucinate', generating inaccurate information that is often presented in such a fluent and eloquent way that it is highly convincing. This is particularly problematic in the field of medicine, where misinformation can have serious consequences. The perception that ChatGPT is informative, helpful and empathetic in health counselling and advice does not, therefore, mean that the technology is ready to replace healthcare personnel in patient care. Mork et al. are also clear on this: their study explores how generative AI can serve as an advisory tool, not as a substitute for medical expertise.

Who is responsible if AI provides incorrect information? How do we ensure that patients maintain trust in healthcare personnel when they know that the answers might be generated by AI? And how do we prevent AI from reinforcing existing biases in health information or access to health care? The secrecy surrounding ChatGPT's training data makes it difficult to determine what the model's responses are based on and to what extent it is influenced by source bias.

AI cannot be held to account for its recommendations, and the responsibility for medical advice and decisions must still lie with healthcare personnel. Nevertheless, the study by Mork et al. adds to an increasing body of research

indicating that such models can serve as a valuable support tool for healthcare personnel, providing assistance in the face of growing demands for efficiency. With responsible use and a clear focus on patient safety and ethical principles, this technology can be an important contributor to a safe and efficient health service.

REFERENCES

1. Mork ET, Mjøs HG, Jammer I et al. Kunstig intelligens og legers svar på helsespørsmål. Tidsskr Nor Legeforen 2025; 145. doi: 10.4045/tidsskr.24.0402. [CrossRef]

2. Regjeringen. Nasjonal strategi for kunstig intelligens. https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/ Accessed 19.1.2025.

3. Achiam J, Adler S, Agarwal S et al. Gpt-4 technical report. arXivorg. Preprint 15.3.2023. https://arxiv.org/abs/2303.08774 Accessed 21.1.2025.

4. Meta Fundamental AI Research Diplomacy Team (FAIR)†. Human-level play in the game of Diplomacy by combining language models with strategic reasoning. Science 2022; 378: 1067–74. [PubMed][CrossRef]

5. Kosinski M. Evaluating large language models in theory of mind tasks. Proc Natl Acad Sci U S A 2024; 121. doi: 10.1073/pnas.2405460121. [PubMed] [CrossRef]

6. Strachan JWA, Albergo D, Borghini G et al. Testing theory of mind in large language models and humans. Nat Hum Behav 2024; 8: 1285–95. [PubMed] [CrossRef]