Ethical, Legal, and Social Implications of Generative AI (GenAI) in Healthcare

Publication Date: Feb 8, 2024

Collection Editor(s):

Kristin Kostick-Quenet, PhD
Assistant Professor, Center for Medical Ethics and Health Policy, Baylor College of Medicine

Introduction

The co-evolution of computational processing power and neural network models has made revolutionary developments in generative artificial intelligence (GenAI) possible. One type of GenAI, large language models (LLMs), are disrupting a wide range of industries, including healthcare. LLMs are trained on large corpora of natural human language to predict and generate text (in chatbot form) that persuasively conveys contextual and semantic understandings. They are also being used to discover patterns in other types of data, such as genomic data, and enable the integration of multiple data types in ways that far surpass our previous capabilities.

Although few hypothesized uses are ready for application, GenAI suggests many potential benefits for healthcare, including clinical decision support, enhanced patient communication and engagement, streamlined clinical documentation and reduced administrative workloads for healthcare professionals, assisted candidate gene prioritization and selection, and accelerated drug discovery. While these emerging tools may improve care, their ethical, legal, and social implications (ELSI) remain unclear.

At the time of writing, no GenAI systems have been reviewed by the US Food and Drug Administration (FDA) and regulatory frameworks are unequipped to address two novel features of GenAI: 1) its ability to adapt and improve performance over time or in response to changing conditions and 2) its capacity for continuous learning through unsupervised, "autodidactic" self-teaching. These features make GenAI a "moving target" that potentially requires new regulatory approaches.

The use of GenAI in the healthcare setting raises numerous ethical concerns. Notably, LLMs are known to "hallucinate," or generate outputs not grounded in fact, logic, or their training data. Accurate or not, results are conveyed with a tone of confidence that may easily be mistaken by humans as fact. Integrating unverified outputs into healthcare scenarios could negatively impact patient safety as well as expose practitioners to liability. Further, the quality and reliability of responses vary in line with the quality of user's prompts, an observation that has spawned a new focus on "prompt engineering" and hybrid intelligence, which is a broad set of approaches that aim to optimize human-machine interactions in ways that capitalize on both human and AI capacities.

A defining feature of GenAI technology is its capacity to generate new content *in the style of* a select portion of the training data. While this allows for the ability to generate novel outputs that emulate established styles (e.g., a new sonnet in the style of Shakespeare) or to tailor outputs for specific users or use cases, it can also reveal sensitive or personal information contained in training data sets. This is because model parameters represent specific features of data sets. The greater the

number of parameters, the more features are inextricably embedded in models, and the more closely outputs mirror the training data. This capacity has drawn criticism from artists and other content creators who call for greater consent, credit, and compensation for the inclusion of their copyrighted information in GenAI training data. In healthcare settings, the potential for unintended data security breaches raises significant concerns about patient privacy and ability to consent to this use of their information.

GenAI models are also difficult to audit, in part because they are often proprietary rather than open source. Companies are reluctant to disclose which data was used to train their models. This lack of transparency has raised concerns about bias, generalizability, and fair attribution related to model outputs. Studies have revealed that LLMs reproduce racial, ethnic, gender and other biases found in human language and can produce outputs that can perpetuate "race-based" medicine or other forms of discrimination in healthcare. Further, given that internet data overrepresent younger users and those from Western, English-speaking countries, LLMs have been criticized for propagating dominant rather than diverse viewpoints. Potential injustices for marginalized groups are compounded by the fact that non-Western nations may be more likely to pay the price for climate changes caused by the exploding energy requirements (rising over 300,000x in just six years) of LLMs and other GenAI systems.

Some researchers are asking, "how big is *too* big?" and calling for greater transparency and more careful curation of training data sets for GenAI. This collection offers an introduction to the ELSI of GenAI that have been identified so far and highlights the need for participatory and anticipatory pathways to GenAI policy and regulation. It will be updated with new literature as it is published. You are invited to email your suggestions to info@elishub.org.

Bias, Error and Hallucinations in GenAI & LLMs

- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623.
- Alkuraya, I. F. (2023) Is artificial intelligence getting too much credit in medical genetics? *American Journal of Medical Genetics Part C: Seminars in Medical Genetics, 193*(3), Article e32062.
- Omiye, J. A., Lester, J. C., Spichak, S. Rotemberg, V., & Daneshjou, R. (2023). Large language models propagate race-based medicine. *npj Digital Medicine, 6*, Article 195.

Privacy & Consent

- Blease, C. (2023). Open AI meets open notes: Surveillance capitalism, patient privacy and online record access. *Journal of Medical Ethics, 50*, 84–89.
- Allen, J. W., Earp, B. D., Koplin, J., & Wilkinson, D. (2024). Consent-GPT: Is it ethical to delegate procedural consent to conversational AI? *Journal of Medical Ethics, 50*, 77–83.
- Minssen, T., Vayena, E., & Cohen, I. G. (2023). The challenges for regulating medical use of ChatGPT and other large language models. *JAMA, 330*(4), 315–316.

Accountability and Liability

- Mann, S. P., Earp, B. D., Nyholm, S., Danaher, J., Møller, N., Bowman-Smart, H., Hatherley, J., Koplin, J., Plozza, M., Rodger, D., Treit, P. V., Renard, G., McMillan, J., & Savulescu, J.

(2023) [Generative AI entails a credit−blame asymmetry](). *Nature Machine Intelligence, 5*, 472−475.

- Duffourc, M., & Gerke, S. (2023). [Generative AI in health care and liability risks for physicians and safety concerns for patients](). *JAMA, 330*(4), 313−314.
- Spirling, A. (2023). [Why open-source generative AI models are an ethical way forward for science](). *Nature, 616*(7957), Article 413.
- Van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). [ChatGPT: Five priorities for research](). *Nature, 614*(7947), 224−226.
- Meskó, B., & Topol, E. J. (2023). [The imperative for regulatory oversight of large language models (or generative AI) in healthcare](). *npj Digital Medicine, 6*, Article 120.

GenAI, Bioethics, and Medical Ethics Education

- Cohen, I. G. (2023). [What should ChatGPT mean for bioethics?]() *The American Journal of Bioethics, 23*(10), 8-16.
- Rahimzadeh, V., Kostick-Quenet, K., Blumenthal Barby, J., & McGuire, A. L. (2023). [Ethics education for healthcare professionals in the era of ChatGPT and other large language models: Do we still need it?]() *The American Journal of Bioethics, 23*(10), 17−27.
- Chen, J., Cadiente, A., Kasselman, L. J., & Pilkington, B. (2023). [Assessing the performance of ChatGPT in bioethics: A large language model's moral compass in medicine](). *Journal of Medical Ethics, 50*, 97−101.

Suggested Citation

Kostick-Quenet, K. (2024). Ethical, legal, and social implications of generative AI (GenAI) in healthcare. In *ELSIhub Collections*. Center for ELSI Resources and Analysis (CERA).