

# Ethical Challenges in Deploying Large Language Model Chatbots for LGBTQ2+ Mental Health Support

Alexandre Ngoc Nguyen Teichmann and Jiayi Xing

Volume 9, Number 2, 2026

URI: <https://id.erudit.org/iderudit/1124215ar>  
DOI: <https://doi.org/10.7202/1124215ar>

[See table of contents](#)

## Publisher(s)

Programmes de bioéthique, École de santé publique de l'Université de Montréal

## ISSN

2561-4665 (digital)

[Explore this journal](#)

## Cite this article

Ngoc Nguyen Teichmann, A. & Xing, J. (2026). Ethical Challenges in Deploying Large Language Model Chatbots for LGBTQ2+ Mental Health Support. *Canadian Journal of Bioethics / Revue canadienne de bioéthique*, 9(2), 155–160.  
<https://doi.org/10.7202/1124215ar>

## Article abstract

As AI-powered chatbots become more common in mental health care, we explore in this commentary the ethical concerns they raise for LGBTQ2+ individuals — users who already face considerable systemic barriers and stigma. We examine how these tools, while promising greater access, may unintentionally perpetuate harm through bias, privacy risks, and techno-solutionism.



CRITICAL COMMENTARY (PEER-REVIEWED)

## Ethical Challenges in Deploying Large Language Model Chatbots for LGBTQ2+ Mental Health Support

Alexandre Ngoc Nguyen Teichmann<sup>a</sup>, Jiayi Xing<sup>b</sup>

### Résumé

Alors que les chatbots alimentés par l'IA deviennent de plus en plus courants dans le domaine des soins de santé mentale, nous explorons dans ce commentaire les préoccupations éthiques qu'ils soulèvent pour les personnes LGBTQ2+, des utilisateurs qui sont déjà confrontés à des obstacles systémiques et à une stigmatisation considérables. Nous examinons comment ces outils, bien qu'ils promettent un meilleur accès, peuvent involontairement perpétuer les préjugés par le biais de préjugés, de risques liés à la confidentialité et du technosolutionnisme.

### Mots-clés

LGBTQ2+, grands modèles linguistiques, chatbots IA

### Abstract

As AI-powered chatbots become more common in mental health care, we explore in this commentary the ethical concerns they raise for LGBTQ2+ individuals — users who already face considerable systemic barriers and stigma. We examine how these tools, while promising greater access, may unintentionally perpetuate harm through bias, privacy risks, and technosolutionism.

### Keywords

LGBTQ2+, large language models, AI chatbots

### Affiliations

<sup>a</sup> Department of Psychiatry, University of Ottawa, Ottawa, Ontario, Canada

<sup>b</sup> School of Medicine, Queen's University, Kingston, Ontario, Canada

**Correspondance / Correspondence:** Alexandre Ngoc Nguyen Teichmann, [ateic077@uottawa.ca](mailto:ateic077@uottawa.ca)

## INTRODUCTION

The emergence of large language model (LLM) chatbots, such as ChatGPT and similar artificial intelligence (AI)-driven conversational agents, has opened new frontiers in the delivery of mental health support, particularly in the context of increasing demand for mental health services. By offering on-demand and anonymous interactions at the very fingertips of users, these tools have immense potential in the realm of mental health, such as reducing barriers to care, especially in contexts where access is limited due to stigma, cost, or geographic isolation. However, the deployment of such technologies in mental health contexts raises significant ethical issues, especially amongst vulnerable populations. Marginalized and underrepresented populations require careful and specific considerations when it comes to healthcare needs and more importantly mental health needs, given historical mistreatment and barriers to access that remain deeply embedded in health systems. Notably, the unchecked deployment of AI technologies, such as general use LLM chatbots, have been shown to perpetuate biases, raising ethical concerns regarding the safety and equitable use of such technologies (1).

This paper addresses the ethical issues raised by LLM chatbots in mental health services, and more specifically amongst one community: LGBTQ2+ individuals. Pre-existing barriers such as discrimination, social stigma, and childhood adversities have been well-documented contributors to chronic stress in LGBTQ2+ individuals; there are particularly pronounced effects amongst transgender and nonbinary populations, who experience higher prevalence of anxiety and depression, as well as post-traumatic stress disorder (PTSD), non-suicidal self-injury, and suicidal tendency compared to cisgender peers (2,3). In LGBTQ2+ populations, disparities persist even after accounting for access to social support and negative social interactions, indicating that unmeasured aspects of systemic discrimination continue to drive elevated risk (4).

While LLM chatbots for mental health raise numerous ethical issues themselves, interaction with individuals from the LGBTQ2+ community involve specific considerations that pose additional ethical issues, given the nuanced and identity-related nature of the challenges faced by this community and established prevalence of mental health concerns. In this paper, we thus first explore the mental health considerations that influence mental health care delivery in LGBTQ2+ individuals. Secondly, we examine in depth three ethical issues that LLM chatbot use for mental health raise in this specific community: bias, safety and related privacy concerns, and technological solutionism.

## MENTAL HEALTHCARE CONSIDERATIONS

Mental healthcare is unlike any other form of healthcare, in that its outcomes are extremely dependent on the phenomenology of the individual on the receiving end (5). From this perspective, there are several key considerations that must be integrated when delivering mental healthcare to the LGBTQ2+ community. First, it is important when delivering care to foster an environment that promotes affirmation and inclusive communication, through sensitive language, and a safe and inclusive environment. This has been shown to positively affect health outcomes (6) and is undoubtedly also a significant factor to

consider in mental healthcare delivery, as addressing individuals who may already have a weakened or damaged sense of identity requires building a thoughtful therapeutic rapport.

Second, we must consider minority stress theory when discussing healthcare delivery, which in this context can be defined as the: “hypothesis that sexual minority health disparities are produced by excess exposure to social stress faced by sexual minority populations due to their stigmatized social status” (7). As LGBTQ2+ communities face systemic stressors that affect their mental health, the stigma and barriers present in the healthcare system demand special identification and intervention. This is linked to the concept of intersectionality, which also affects health outcomes and the care that these populations receive — particularly individuals with multiple marginalized identities (8). Systemic stress factors, such as discrimination, internalized stress, prejudice and perceived rejection must all be accounted for and evaluated in a health setting, as response to interventions are partly dictated by the external influence of those stressors (9). This applies heavily to mental health, where sensitive issues such as psychological vulnerabilities demand trust between patients and providers, which can only be fostered by meaningful understanding and non-prejudicial behaviour.

Additionally, interacting in a health care setting with individuals from the LGBTQ2+ community requires a thorough understanding of the traumas to which these individuals are frequently exposed, highlighting the need for a trauma-informed approach when delivering care. Researchers have highlighted the importance of promoting safety, trust, choice, collaboration and empowerment in a clinical setting (10). This also includes active screening for trauma, such as violence and ensuing psychological consequences (11).

Thus in the scope of our analysis, we are considering LLM chatbots as providers of mental health advice, but not as a modality possessing personhood nor a replacement for human therapists. In the context of providing such services, treating LLMs as anything more than sophisticated tools risks misleading users about their actual capacities.

Contemporary LLMs currently operate without consciousness, intentionality, culpability, or the ability to execute autonomous understanding; outputs are the result of probabilistic text prediction, not reflective thought or empathy (12). This distinction is essential in the mental health context, where therapeutic relationships rely on authenticity and trust. Similarly, we do not consider LLM chatbots to be replacements for human therapists. A recent control trial tested a generative AI chatbot called Therabot for mental health treatment in adults presenting with clinically significant depression, anxiety, or high-risk eating disorder symptoms. Participants reported high engagement and a therapeutic alliance score comparable to outpatient psychotherapy norms (13). Even with a notable reduction in participants’ symptoms, the study authors emphasized the continuing need for risk mitigation and close human supervision. Again, referencing minority stress theory, LGBTQ2+ individuals in whom therapeutic trust is often already fragile may face compounded risks if an LLM chatbot fails to recognize their lived realities and layered identities. Thus, despite the promise they show as supplemental tools, we do not regard LLM chatbots as replacements for human therapists.

Having laid the groundwork for responsible mental health care practices and considerations in this population, we now turn to the ethical issues raised by LLM chatbot use.

## BIAS

LLMs are trained on immense datasets that stem from what is available on today’s internet, from books to articles, forums, and social media. While this provides the user access to an unfathomable quantity of data, it can also reflect prejudices, stereotypes, and systemic inequalities already embedded within these data sources. Consequently, when deployed in mental health contexts, LLMs may inadvertently reproduce harmful or exclusionary language, particularly toward marginalized groups such as the LGBTQ2+ community. Research has shown that LLMs can perpetuate gender, racial, and sexual orientation biases due to skewed or under-representative training data (14). Therefore, if we think of LLM chatbots as a provider of mental health support, they can have potentially negative consequences when compared to more impartial and trusted human support. In the LGBTQ2+ community, this could reflect unchecked biases embedded in various data sources, as opposed to a human clinician who could at least be aware and attempt to mitigate their own biases. As this community has faced and continues to be exposed to significant gender and sexual orientation biases and discrimination, biased LLM training datasets will reproduce these biases in their interactions with users (15). As Serravalle et al. explain, “These biases stem from the reproduction of historical cultural patterns that are inadvertently and often unintentionally transferred to the process of creating and developing technologies such as algorithms or automated systems.” (16) Further, identifying and correcting these biases will be even harder in LLMs, as the way they are trained is obscure and considered a “black box” (17).

Second, there is a risk that LLMs trained on more prevalent and dominating cultural narratives will exclude or misrepresent LGBTQ2+ subgroups, particularly those at the intersections of race, disability, or socioeconomic status. For example, LLMs have been found to present an underlying masculine bias (18), which can be particularly problematic in a community where many individuals already struggle with their gender identity and trauma. Such subtle but frequent harms compound the minority stress already experienced by these users and may contribute to therapeutic disengagement and systemic mistrust. Given the psychological vulnerability of LGBTQ2+ users seeking mental health support, such risks cannot be overlooked.

Therefore, if LLM chatbots perpetuate bias in an already stigmatized population, this poses a significant health equity problem (19). Tools intended to increase access and reduce disparities in mental health care could instead deepen inequities

by reinforcing harmful stereotypes and eroding trust. Research shows that biased interactions with LLMs decrease user trust in this technology more broadly, reducing their willingness to engage with it across other domains, such as education and professional development (20). Petzel and Sowerby highlight that AI prejudice disproportionately affects and harms marginalized groups, undermining their confidence in AI technologies. Though this research did not include LGBTQ2+ groups, it is reasonable to imagine that individuals in this community would be similarly affected, and so discouraged from using LLM-based tools even where they might be appropriate and beneficial, thus further widening structural inequalities. This risk becomes even more concerning if human mental health professionals are intentionally or unintentionally replaced by AI systems in efforts to reduce workload or costs, an issue explored later in this paper.

## SAFETY AND PRIVACY-RELATED CONCERNS

It has been well established that individuals from the LGBTQ2+ community suffer from higher risk, than their peers, of compromised mental health, including dangerous behaviours such as suicidality and self-harm (21). So it is reasonable to assume that handling such situations would be one of the LLM chatbots primary tasks when integrated into mental health care amongst this community.

Up to now, LLM chatbots have not been subjected to clinical validation, unlike other medical devices, which is particularly problematic in a context of evidence-based medicine. Non-clinically validated “black box” technologies raise the possibility of expressing unpredictable outputs. When we are dealing with human lives in a high-stake environment, such as in a crisis management scenario, unsupervised outputs present a serious concern for potential danger. LLM chatbots should be able to systematically refer to tools such as hotlines where human intervention is essential in scenarios that it detects as being dangerous. Yet, LLM chatbots have had conflicting successes in this regard: on one hand, they have been shown to be slightly superior to human therapists when dealing with suicidality (22), but on the other hand, they have been shown to fail adequate crisis management, omitting elements such as sufficient inquiry or providing excessive directive advice (23). We are then faced with a problem of clinical validity. Is it possible to test an LLM chatbot in a real-world scenario, such as in an actual crisis situation? Would it be possible to perform those tests in an ethical manner, knowing that not offering the gold standard to an individual in crisis would be unethical? Additionally, carrying out such experiments in this specific community would perpetuate historic mistreatment, rendering this option non-viable and dangerous. We are led to consider what measures of validity should be implemented in LLM chatbots for mental health.

Second, if we are to consider the safety of such AI tools for mental health care in the LGBTQ2+ community, we must examine privacy-related safety concerns. LLM chatbots acquire and analyze immense quantities of data through their interactions. Many LGBTQ2+ individuals live in geographical and/or social contexts that are hostile to their identity. They may live in fear that their identity could be exposed, causing considerable ensuing psychological or even physical trauma (24). Thus, privacy issues are deeply concerning for LGBTQ2+ users who, in many regions, must guard their identities due to fear of discrimination, outing, or legal consequences. LLM chatbots designed for mental health support frequently encourage users to disclose intimate details about their emotions, identity, and personal history (25). If there are not robust security measures, safeguards, or regulations in place to prevent data sharing or leaks, this could put those individuals at greater risk. Moreover, depending on the political context, an ill-intentioned government or insurance agency having access to this type of data could easily enact discriminatory measures against those users, exposing them to additional harms. Recent litigation underscores the severity of these risks.

In the landmark 2024 US (Florida) case of *Garcia v. Character Technologies* (26), the mother of a 14-year-old boy alleged that her son died by suicide after a chatbot provided harmful advice, raising questions about legal liability in the use of LLM chatbots for mental health support. The company claimed that the chatbot should be protected under First Amendment free speech rights, echoing arguments raised in a prior case in 2023, *Walters v. OpenAI, LLC* (27), in which the court’s decision suggested that generative AI outputs may be entitled to First Amendment protections. If US courts are to continue endorsing such claims, developments in AI technology designed for mental health care require further amendment of existing legal doctrines and regulatory frameworks, to ensure adequate legal protections for vulnerable populations such as LGBTQ2+ users.

Furthermore, companies developing those tools often do not have incentives to implement safe data policies, as their main goal is profit, and currently there is a paucity of laws and regulations governing data availability, especially in North America. This creates a clear problem as companies handling sensitive data have insufficient motivation to protect that very data, thus exposing vulnerable health data to potential leaks. Moreover, it seems that chatbots currently used in the health setting lack even basic transparency about data security and privacy. A 2021 scoping review by May and Denecke, which included a significant number of mental health coaching chatbots, found that “the majority [of health chatbots] did not provide any information regarding security or privacy aspects” (28). This is particularly pertinent, as vulnerable users must understand how their sensitive data will be handled, especially the LGBTQ2+ community, who already experience systemic mistrust in the healthcare system. We acknowledge that certain data protections do exist in the US, including the California Consumer Privacy Act (CCPA) (29) which protects rights regarding the collection and use of consumer data, and the Health Insurance Portability and Accountability Act (HIPAA) (30) that widely regulates health data in clinical settings. However, these current frameworks do not adequately encompass the regulation of AI chatbots for mental health support. They do not impose proactive safeguards against inference based on sensitive issues, such as sexual orientation or gender identity, nor address the risks of secondary use of such data by insurers, advertisers, or state actors. HIPAA also applies only to “covered entities”, which leaves consumer

health apps and chatbots largely unregulated (30). There is thus a regulatory gap in the US that allows LLM chatbots to handle sensitive health-related data without necessarily falling under formal health service classifications and privacy guidelines.

## TECHNO-SOLUTIONISM

The very concept of LLM chatbots for mental health support in a particularly socially marginalized community raises the issue of techno-solutionism, defined as the “oversimplified reduction of complex problems into technological puzzles” (31). In our context, it can be applied to healthcare and societal issues, and more specifically to the societal roots of the compromised mental health experienced by members of the LGBTQ2+ community. LLM chatbots may provide rapid access to support, but they do not address the structural issues that shape well-being: discrimination, economic insecurity, familial rejection, and societal stigma. Techno-solutionism in mental health is a significant concern as disproportionately marginalized subgroups within the LGBTQ2+ community face intersectional barriers linked to race, disability, socioeconomic status, and other identities. These compounded inequalities mean that introducing LLM chatbots without addressing systemic causes could create new barriers to accessing appropriate and supportive mental health care rather than removing existing ones. As Eubanks (32) highlights, technological “fixes” often obscure the very social injustices they purport to address, reinforcing inequality by shifting responsibility away from systemic reform.

Furthermore, there is a danger that these tools may be deployed as substitutes, rather than supplements, to human care. Mental health care relies not only on empathetic responses, but on rapport building and attunement (33). These qualities are critical in therapeutic interactions, particularly for LGBTQ2+ individuals whose mental health needs often involve identity affirmation and trauma-informed, sensitive support. Although research has found that patients may more readily disclose sensitive information (e.g., suicidal thoughts) to a chatbot than to a human clinician when facing stigma (34), AI-written content elicits less empathy from users than human-narrated content (35), highlighting the gap that would occur if we were to completely replace human clinicians with AI agents. LGBTQ2+ users, already marginalized in clinical settings, may find that chatbots not only fail to meet their emotional needs, but also inadvertently contribute to a sense of lack of validation, a key element in gender affirmation (36).

## CONCLUSION

In this paper, we examined the ethical complexities of using LLM chatbots for mental health support within the LGBTQ2+ community. While these tools promise expanded access to care, particularly for populations facing barriers due to stigma, geography, or cost, they also raise serious concerns when deployed amongst marginalized users. We focused on three major areas of ethical risk: the reproduction of bias, safety and privacy vulnerabilities, and the broader issue of techno-solutionism. These risks cannot be examined in isolation and require an intersectional lens that accounts for interdependent sources of harm when examining inclusion and exclusion. We explored how mental health considerations set an ethical baseline for AI development, emphasizing the importance of trust and therapeutic integrity. These considerations are inseparable from systemic bias that continues to influence which communities are excluded from care and whose needs are marginalized. For LGBTQ2+ communities, bias can manifest in misgendering, exclusionary language, or invalidation of identity, potentially leading to psychological harm. Concerns around safety and privacy amplify these exclusions, as risks of data misuse or harm are not evenly distributed but are rather disproportionately experienced by LGBTQ2+ individuals living in social or political contexts where disclosure of their identity could lead to real-world consequences. The current lack of transparent data handling practices, combined with limited regulation of AI tools, is at risk of augmenting an already present systemic mistrust. The lack of clinical validation and unpredictable outputs from these “black box” systems means that chatbot responses can be unreliable or even dangerous when users are in acute distress. With legal frameworks currently not equipped to hold LLM chatbots culpable to the same degree as human therapists, caution must be exercised appropriately in future developments. Techno-solutionist framing is the overarching structural risk that further obscures the inequities that shape access to care, prioritizing efficiency at the cost of trivializing the systemic challenges faced by vulnerable populations. While chatbots may provide rapid, anonymous support, they cannot replace the relational depth, empathy, and affirmation that LGBTQ2+ individuals often require from mental health care providers. There is thus a strong possibility that, in the name of innovation and cost-effectiveness, such tools will become substitutes for meaningful human care.

In the current healthcare worker shortage, LLM chatbots hold promise as an aid in mental health support. However, this promise can only be realized if these tools are developed and deployed with specific attention to the needs and vulnerabilities of the LGBTQ2+ individuals they aim to serve, such as through frameworks that can ensure the safety and quality of LLM chatbot responses (37). Without such care, these technologies risk reproducing the very harms they claim to alleviate. Employing an intersectional lens is a starting point that highlights how these dynamics are compounded, and it reminds us that future ethical deployment of LLM chatbots is less about producing rapid technological fixes than addressing the layers of inequities that shape mental health care.

**Reçu/Received:** 25/06/2025**Remerciements**

Nous remercions chaleureusement le Dr Jay Shaw (Université de Toronto) pour ses conseils dans la réalisation de ce travail.

**Conflits d'intérêts**

Aucun à déclarer

**Publié/Published:** 16/03/2026**Acknowledgements**

We gratefully acknowledge Dr. Jay Shaw (University of Toronto) for his guidance in the creation of this work.

**Conflicts of Interest**

None to declare

**Édition/Editors:** Aliya Affdal

Les éditeurs suivent les recommandations et les procédures décrites dans le [Core Practices](#) de COPE. Plus précisément, ils travaillent pour s'assurer des plus hautes normes éthiques de la publication, y compris l'identification et la gestion des conflits d'intérêts (pour les éditeurs et pour les auteurs), la juste évaluation des manuscrits et la publication de manuscrits qui répondent aux normes d'excellence de la revue.

The editors follow the recommendations and procedures outlined in the COPE [Core Practices](#). Specifically, the editors will work to ensure the highest ethical standards of publication, including: the identification and management of conflicts of interest (for editors and for authors), the fair evaluation of manuscripts, and the publication of manuscripts that meet the journal's standards of excellence.

**Évaluation/Peer-Review:** Matimba Swana

Les recommandations des évaluateurs externes sont prises en considération de façon sérieuse par les éditeurs et les auteurs dans la préparation des manuscrits pour publication. Toutefois, être nommé comme évaluateurs n'indique pas nécessairement l'approbation de ce manuscrit. Les éditeurs de la [Revue canadienne de bioéthique](#) assument la responsabilité entière de l'acceptation finale et de la publication d'un article.

Reviewer evaluations are given serious consideration by the editors and authors in the preparation of manuscripts for publication. Nonetheless, being named as a reviewer does not necessarily denote approval of a manuscript; the editors of the [Canadian Journal of Bioethics](#) take full responsibility for final acceptance and publication of an article.

**REFERENCES**

- Xue J, Wang Y-C, Wei C, et al. [Bias and fairness in chatbots: an overview](#). arXiv; 2023.
- Inderbinen M, Schaefer K, Schneeberger A, et al. [Relationship of internalized transnegativity and protective factors with depression, anxiety, non-suicidal self-injury and suicidal tendency in trans populations: a systematic review](#). Front Psychiatry. 2021;12:636513.
- Hainey KJ, Connolly DJ, Thomson R, et al. [Mental health outcomes in transgender and nonbinary people: an umbrella review](#). JAMA Psychiatry. 2025;82(11):1142-51.
- Kingsbury M, Findlay L. [Mental health and access to support among 2SLGBTQ+ youth](#). Health Rep. 2024;35(11):12-22.
- Ritunnano R, Papola D, Broome MR, Nelson B. [Phenomenology as a resource for translational research in mental health: Methodological trends, challenges and new directions](#). Epidemiol Psychiatr Sci. 2023;32:e5.
- Moore C, Dukes C. [The value of identity: providing culturally-responsive care for LGBTQ+ patients through inclusive language and practices](#). Del J Public Health. 2019;5(3):6-8.
- Frost DM, Meyer IH. [Minority stress theory: application, critique, and continued relevance](#). Curr Opin Psychol. 2023;51:101579.
- Gessner M, Bishop MD, Martos A, et al. [Sexual minority people's perspectives of sexual health care: understanding minority stress in sexual health settings](#). Sex Res Soc Policy. 2020;17(4):607-18.
- Comeau D, Johnson C, Bouhamdani N. [Review of current 2SLGBTQIA+ inequities in the Canadian health care system](#). Front Public Health. 2023;11:1183284.
- Sinko L, Ghazal LV, Fauer A, Wheldon CW. [It takes more than rainbows: supporting sexual and gender minority patients with trauma-informed cancer care](#). Cancer. 2024;130(4):507-16.
- Livingston NA, Berke D, Scholl J, et al. [Addressing diversity in PTSD treatment: clinical considerations and guidance for the treatment of PTSD in LGBTQ populations](#). Curr Treat Options Psychiatry. 2020;7(2):53-69.
- Browning J. [Personhood and AI: why large language models don't understand us](#). AI Soc. 2024;39(5):2499-2506.
- Heinz MV, Mackin DM, Trudeau B, et al. [Randomized trial of a generative AI chatbot for mental health treatment](#). NEJM AI. 2025;2(4):A10a2400802.
- Mehrabi N, Morstatter F, Saxena N, et al. [A survey on bias and fairness in machine learning](#). ACM Comput Surv. 2021;54(6):115.
- Ten Tije M. [Bias in, bias out? Why AI isn't as neutral as you think](#). Tilburg.ai; 2024.
- Santa Serravalle R, Muñoz Alarcon C, Duran Herrera F. [Gender bias in AI: uncovering the roots and shaping inclusive futures](#). Mondo Internazionale; 26 Mar 2025.
- Manche R, Myakala PK. [Explaining black-box behavior in large language models](#). Int J Comput Artif Intell. 2022;3(2):102-8.
- Bartl M, Murphy TB, Leavy S. [Adapting psycholinguistic research for LLMs: gender-inclusive language in a coreference context](#). arXiv:2502.13120; 2025.
- Mirzaei T, Amini L, Esmailzadeh P. [Clinician voices on ethics of LLM integration in healthcare: a thematic analysis of ethical concerns and implications](#). BMC Med Inform Decis Mak. 2024;24:250.
- Petzel Z, Sowerby L. [Prejudiced interactions with large language models \(LLMs\) reduce trustworthiness and behavioral intentions among members of stigmatized groups](#). Comput Hum Behav. 2025;165:108563.

21. Russell ST, Fish JN. [Mental health in lesbian, gay, bisexual, and transgender \(LGBT\) youth](#). *Annu Rev Clin Psychol*. 2016;12:465-87.
22. McBain RK, Cantor JH, Zhang LA, et al. [Competency of large language models in evaluating appropriate responses to suicidal ideation: comparative study](#). *J Med Internet Res*. 2025;27:e67891.
23. Scholich T, Barr M, Wiltsey Stirman S, Raj S. [A comparison of responses from human therapists and large language model-based chatbots to assess therapeutic communication: mixed methods study](#). *JMIR Ment Health*. 2025;12:e69709.
24. Schwartz AP. [Why outing can be deadly](#). National LGBTQ Task Force; 20 Feb 2014.
25. Gümüşel E. [A literature review of user privacy concerns in conversational chatbots: a social informatics approach](#). *J Assoc Inf Sci Technol*. 2025;76(1):121-54.
26. [Garcia v. Character Technologies, Inc.](#), 6:24-cv-01903, (M.D. Fla.).
27. [Walters v. OpenAI, L.L.C.](#), 1:23-cv-03122, (N.D. Ga.).
28. May R, Denecke K. [Security, privacy, and healthcare-related conversational agents: a scoping review](#). *Inform Health Soc Care*. 2022;47(2):194-210.
29. [The California Consumer Privacy Act of 2018](#), CA Civ Code § 1798.100 (2018).
30. Office for Civil Rights. [Summary of the HIPAA Privacy Rule \(original ruling 1996\)](#). U.S. Department of Health and Human Services; 2025.
31. Cambridge Dictionary. [Techno-solutionism](#).
32. Eubanks V. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press; 2018.
33. Beyene LS, Strand EB, Misund AR, et al. [Conceptualizing healthcare professionals' relational competence in mental healthcare: an integrative review](#). *Int J Nurs Stud Adv*. 2024;7:100266.
34. Miles O, West R, Nadarzynski T. [Health chatbots acceptability moderated by perceived stigma and severity: a cross-sectional survey](#). *Digit Health*. 2021;7:20552076211063012.
35. Shen J, DiPaola D, Ali S, et al. [Empathy toward artificial intelligence versus human experiences and the role of transparency in mental health and social support chatbot design: comparative study](#). *JMIR Ment Health*. 2024;11:e62679.
36. Glynn TR, Gamarel KE, Kahler CW, et al. [The role of gender affirmation in psychological well-being among transgender women](#). *Psychol Sex Orientat Gend Divers*. 2016;3(3):336-44.
37. Long Y, Ma Z, Mei Y, Su Z. [AffirmativeAI: towards LGBTQ+ friendly audit frameworks for large language models](#). *arXiv:2405.04652*; 2024.