

Human–AI Collaboration in Socioemotional Support: Empirical Evidence of the Spontaneous Use of LLMs

Isabelle Q. G. de Assis, Jéssyka Vilela

Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Recife, PE – Brasil

{iqga, jffv}@cin.ufpe.br

Abstract. *Context: The adoption of Large Language Models (LLMs) increased human–AI interaction in emotional support settings, but their collaborative role is still unclear. Objective: We investigate LLMs' use of socioemotional support, examining user motivations, perceived benefits, and risks. Method: A Multivocal Literature Review analyzed 22 academic studies and 113 gray literature sources published between 2021 and 2025. Results: Users seek privacy, and relief from loneliness, perceiving interactions as accessible and non-judgmental; however, risks include emotional dependency, inappropriate advice, and data vulnerabilities. Conclusions: LLMs are becoming mediators of informal care, showing collaborative potential and ethical risks.*

1. Introduction

The widespread dissemination of Large Language Models (LLMs), such as ChatGPT, has significantly expanded the possibilities of interaction between humans and computational systems [Nazir and Wang 2023] [Spallek et al. 2023]. ChatGPT, for instance, reached 100 million users within just two months of its launch, demonstrating a massive adoption that is reshaping how people interact with technology [Forbes Brasil 2023].

With advanced capabilities in text comprehension and generation, these models have been integrated into various everyday contexts, ranging from academic and technical task support to the simulation of conversations and social interactions [Nazir and Wang 2023] [Haensch 2025]. Among the multiple observed uses, a notable trend is the growing spontaneous use of these tools as spaces for emotional listening and support, particularly by individuals experiencing loneliness, anxiety, grief, or relational difficulties [Rousmaniere et al. 2025] [Giray 2024].

According to a survey conducted by [Talk Inc Research 2024], 42% of Brazilian youth aged 18 to 24 reported using AI in their personal lives. Specifically in the context of emotional support, 13% of Brazilian AI users—approximately 12 million people—turn to these systems as an “emotional friend/advisor.” Although not designed for clinical or affective purposes, these tools are increasingly appropriated as accessible forms of subjective support [Ma et al. 2024]. Interaction with these artificial intelligences is often perceived as an emotionally safe space, even when users are aware of the system’s limitations [Rousmaniere et al. 2025].

Given the growing appropriation of LLMs for such functions, it becomes necessary to understand not only their technical operation but also their role in users’ lives. In this context, the present study seeks to answer the following research question: *How do users spontaneously employ LLMs for therapeutic and socioemotional support*

purposes, and how are the causes, processes, and consequences of this use reported in academic literature, journalistic media, and social networks?

2. Objectives and Relevance

The general objective of this study is to analyze how individuals spontaneously use LLMs for therapeutic and socioemotional support purposes, based on reports from academic literature, media sources, and social networks. To achieve this, the following specific objectives are defined:

- To identify the motivations that lead users to adopt LLMs as tools for therapeutic and socioemotional support.
- To characterize the interaction dynamics between users and LLMs for these purposes.
- To examine the subjective benefits reported by users resulting from such interactions.
- To map the risks and emerging challenges associated with the unsupervised use of LLMs for mental health.
- To outline the profile of users who turn to these technologies for such support.

The relevance of this research lies in advancing the understanding of emerging human–AI collaborative practices in socioemotional support contexts. By integrating academic sources and gray literature (such as Reddit and TikTok), this study provides a comprehensive and contextualized analysis of a sociotechnical phenomenon that reshapes how individuals seek care, build trust, and construct support networks. The findings contribute to discussions on ethical design, responsibility, and governance in collaborative AI systems that increasingly mediate social and emotional interactions.

3. Related Work

The use of Large Language Models (LLMs) as tools for mental health and socioemotional support has attracted growing interest from both academia and the general public. Recent literature explores this phenomenon from multiple perspectives, ranging from qualitative case studies [Giray 2024] to the development of computational frameworks [Aggarwal et al. 2025]. To contextualize and justify the present research, three relevant studies are analyzed below, contrasting their approaches with the objectives and methodology of this work.

The study by [Giray 2024] investigates the role of ChatGPT in mental health support through a narrative case study based on Reddit reports. Similar to the present research, it uses user accounts to identify use cases, benefits, and risks. However, [Giray 2024] focuses on a single source and analyzes a sample of only seven posts, which differs substantially from this study’s aim of providing a broader and more multifaceted view of how the phenomenon is addressed across different spheres.

Complementarily, [Haensch 2025] examines discourse on TikTok. [Haensch 2025] conducts a large-scale analysis of more than 10,000 comments using supervised classification models. In contrast, the present study adopts a more qualitative approach

within the Multivocal Literature Review (MLR) framework, involving manual data collection and aiming to compare findings with other, less informal sources.

Finally, [Aggarwal et al. 2025] propose a technical framework to detect mental disorders in Reddit posts and generate automated therapeutic recommendations. Their primary objective is the creation and evaluation of a diagnostic and intervention system based on machine learning and Natural Language Processing (NLP). In contrast, this study conducts an exploratory and qualitative investigation into how users already appropriate LLMs for such purposes, rather than developing a tool to do so. A comparative analysis of the discussed studies is summarized in Table 1.

Table 1. Comparison of the objectives, methodologies, and sources of related works.

Criteria	Giray (2024)	Haensch (2025)	Aggarwal et al. (2025)	This work (2025)
Main Goal	To explore the emerging role of ChatGPT in mental and psychological health support in order to understand its potential and limitations.	To explore how users interact with LLMs as mental health tools in order to identify experiences, attitudes, and recurring themes.	To propose a framework for identifying and assessing mental health disorders and their severity, as well as generating therapeutic and behavior change recommendations.	To investigate how individuals use LLMs for therapeutic purposes by mapping motivations, usage patterns, effects, and user profiles.
Methodology	Narrative Case Study	Large-scale data analysis	Framework development	Multivocal Literature Review
Data Source	Reddit	TikTok	Reddit	Academic papers, news outlets, Reddit, TikTok

3. Research Methods

The methodology of this study adopts a Multivocal Literature Review (MLR), an approach chosen due to the emerging nature of the use of Large Language Models (LLMs) for socioemotional purposes. This method integrates the rigor of a Systematic Literature Review (SLR), focused on scientific articles, with the analysis of gray literature, including sources such as news outlets and social media discussions. The protocol was structured based on the guidelines of [Kitchenham and Charters 2007] for SLR and [Garousi et al. 2019] for gray literature, and was divided into three main phases: planning, execution, and reporting.

To classify and evaluate gray literature sources, this study adopted the “shades of gray” model discussed in [Garousi et al. 2019]. This model organizes gray literature into tiers according to the level of editorial control and the credibility and expertise of the content producer. Within this research scope, gray literature selection aimed to capture both public discourse and authentic user experiences. Therefore, sources of “medium gray” (Tier 2), characterized by moderate control and credibility, and “dark gray” (Tier 3), reflecting spontaneous ideas and personal accounts, were included. This approach enabled the capture of both structured media discourse and informal, direct user experiences.

The search strategy was designed to answer five Research Questions (RQs) addressing motivations, interaction dynamics, benefits, risks, and user profiles. Academic databases consulted included ACM Digital Library, Emerald Insight, IEEE Xplore, ScienceDirect, and Scopus. Gray literature sources were selected based on credibility levels, including established news outlets (BBC and The Guardian), specific Reddit communities (r/ChatGPT, r/mentalhealth, r/therapy, and r/selfhelp), and the TikTok platform. The inclusion of TikTok and Reddit as gray literature sources was specifically intended to capture the perspectives of younger demographics. These platforms are predominantly populated by digital natives and young adults, who are generally early adopters of new technologies for alternative purposes.

The selection process applied strict inclusion and exclusion criteria, including a temporal scope from 2021 to 2025 and the requirement of full and open access. In academic literature, 586 initial articles were filtered down to 31, which underwent a quality assessment scoring from 0 to 5, resulting in a final sample of 22 articles scoring above 2.5. For social media, the official Reddit API was used via a Python script to extract the 50 most relevant posts per scope, while 50 TikTok videos were collected manually.

Data extraction and analysis were conducted using Thematic Analysis, enabling the identification of recurring patterns through qualitative and quantitative lenses. Artificial intelligence tools, such as NotebookLM PRO and Gemini PRO, supported evidence categorization and scoring under researcher supervision and manual validation. The final corpus of 135 publications was coded to ensure traceability of findings regarding motivations and consequences of the studied phenomenon.

4. Results and Discussion

The analysis reveals that the use of Large Language Models (LLMs) for emotional support represents an emerging form of human–AI collaboration within sociotechnical networks. Users engage with LLMs not merely as tools, but as interactive agents that mediate emotional expression, reflection, and coping. Demand for these systems is driven by accessibility, anonymity, and the absence of social judgment—conditions that foster disclosure of experiences often silenced in human settings.

A notable tension emerges between academic discourse and lived user experience. While scientific literature emphasizes risks such as algorithmic hallucinations, ethical gaps, and data security vulnerabilities, users in digital communities frequently frame these interactions as collaborative exchanges that provide immediate emotional relief. The predominant user profile includes young adults facing structural barriers to traditional mental health services, positioning LLMs as complementary actors within broader support ecosystems.

Accessibility functions as the core enabler of this human–AI collaboration. Unlike conventional care infrastructures constrained by cost and availability, LLMs offer continuous, on-demand interaction. Users actively shape these interactions through prompt engineering, assigning therapeutic roles (e.g., cognitive behavioral therapist) and engaging in anthropomorphic framing. This co-construction process highlights the participatory and relational nature of the interaction, but also exposes fragilities when safety mechanisms trigger refusals or generic responses.

Users frequently apply prompt engineering techniques to “personify” the AI, requesting it to act according to specific therapeutic approaches (e.g., “act as a cognitive behavioral therapist”). A high degree of anthropomorphization was observed, with some users developing emotional dependence and reporting that the AI’s “presence” mitigates feelings of isolation. However, gray literature sources (e.g., Reddit forums) reveal that this connection is fragile and often leads to frustration when models generate generic responses or safety refusals in sensitive contexts.

Users report that tools such as ChatGPT provide emotional validation and help structure thoughts, functioning as a form of therapeutic writing that facilitates reflection and the processing of difficult experiences. Nevertheless, important risks remain, including potential “clinical hallucinations,” superficial agreement responses, and concerns about privacy and sensitive data usage.

One of the most relevant findings from Reddit posts was the emotional ambivalence present in user narratives: individuals express both relief and gratitude, as well as shame or guilt for relying on machines for emotional disclosure [Haensch 2025]. This tension suggests that socioemotional use of LLMs remains taboo in many contexts, including online mental health communities that often discourage or censor such practices due to more conservative perspectives. The contrast between technical and affective discourse is also evident: technology communities celebrate innovation and autonomy, whereas mental health spaces adopt a cautious and critical stance [Moylean and Doherty 2025]. This dichotomy reflects broader value disputes about what constitutes “care” and “support” in the age of artificial intelligence.

In summary, the findings show that while LLMs can provide valuable preliminary support in situations of crisis or loneliness, their role as collaborative actors in socioemotional care remains largely unregulated and without professional oversight. The effectiveness perceived by users contrasts with unresolved technical and ethical risks, highlighting tensions within emerging human–AI support networks. These results reinforce the need for responsible integration of LLMs into mental health ecosystems, emphasizing ethical design, transparency, and governance to ensure safe and trustworthy collaborative practices.

5. Conclusions and Future Works

This study investigated the phenomenon of the spontaneous use of Large Language Models (LLMs) for therapeutic purposes, a behavior that emerges outside official clinical recommendations. Through a Multivocal Literature Review, the established objectives were achieved, revealing that ease of access, free availability, and the impersonal nature of AI are the primary drivers of adoption.

The results show that, although users report immediate subjective benefits—such as reduced feelings of loneliness and support in organizing thoughts—this use occurs

within a zone of ethical and clinical risk. The dissonance between user optimism and the caution expressed by mental health communities on Reddit, for example, indicates that traditional healthcare systems and AI platforms still lack governance mechanisms to address this emerging paradigm of care.

Regarding the specific objectives, the study characterized the interaction dynamic as a process of active anthropomorphization, in which users shape AI systems to fill emotional support gaps not effectively addressed by state or market structures. However, risks such as hallucinations and the lack of clinical depth remain critical barriers.

We conclude that the use of LLMs for socioemotional support represents an established and expanding form of human–AI collaboration within emerging sociotechnical care networks, demanding urgent attention from the information systems and collaborative systems communities. Future research should prioritize the establishment of specific legislation and robust AI governance frameworks in healthcare, promoting transparency, trust, and responsible AI mediation in emotionally sensitive contexts. Longitudinal studies are needed to monitor the long-term impacts of sustained human–AI interaction. Furthermore, it is crucial to investigate the potential benefits of collaborative models in which mental health professionals and AI platforms work in tandem to enhance therapeutic support both inside and outside of clinical settings (whether physical or virtual).

To address current literature gaps, we propose conducting a future survey-based study. This approach will enable the collection of complementary data across diverse age groups, particularly capturing populations that may exhibit greater technological resistance. Finally, as the technological landscape rapidly evolves, subsequent investigations must broaden their scope to evaluate other LLM platforms that surged in popularity throughout 2025—such as Gemini, Grok, and DeepSeek—comparing how these distinct architectures and their respective guardrails perform in emotionally critical scenarios.

References

- [Aggarwal 2024] AGGARWAL, Vaishali et al. Leveraging LLMs for Mental Health: Detection and Recommendations from Social Discussions. In: 2024 IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT). IEEE, 2024. p. 350-354.
- [Forbes Brasil 2023] FORBES BRASIL. ChatGPT tem recorde de crescimento da base de usuários. Forbes Brasil, [s.l.], 01 fev. 2023. Disponível em: <https://forbes.com.br/forbes-tech/2023/02/chatgpt-tem-recorde-de-crescimento-da-base-de-usuarios/>.
- [Garousi et al. 2019] GAROUSI, Vahid; FELDERER, Michael; MÄNTYLÄ, Mika V. Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and software technology*, v. 106, p. 101-121, 2019.
- [Giray 2025] GIRAY, Louie. Cases of using ChatGPT as a mental health and psychological support tool. *Journal of Consumer Health on the Internet*, v. 29, n. 1, p. 29-48, 2025.

- [Haensch 2025] HAENSCH, Anna-Carolina. " It Listens Better Than My Therapist": Exploring Social Media Discourse on LLMs as Mental Health Tool. arXiv preprint arXiv:2504.12337, 2025.
- [Kitchenham e Charters 2007] KITCHENHAM, Barbara et al. Guidelines for performing systematic literature reviews in software engineering. 2007.
- [Ma et al. 2024] MA, Zilin et al. Evaluating the experience of LGBTQ+ people using large language model based chatbots for mental health support. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. 2024. p. 1-15. NAZIR, Anam;
- [Moylan e Doherty 2025] MOYLAN, Kayley; DOHERTY, Kevin. Expert and Interdisciplinary Analysis of AI-Driven Chatbots for Mental Health Support: Mixed Methods Study. Journal of Medical Internet Research, v. 27, p. e67114, 2025.
- [Rousmaniere et al. 2025] ROUSMANIERE, Tony et al. Large language models as mental health resources: Patterns of use in the united states. 18 mar. 2025.
- [Spallek et al. 2023] SPALLEK, Sophia et al. Can we use ChatGPT for mental health and substance use education? Examining its quality and potential harms. JMIR Medical Education, v. 9, n. 1, p. e51243, 2023
- [Talk Inc Research 2024] TALK INC RESEARCH. Inteligência Artificial na Vida Real. Talk Digital, [s.l.], 2024. Disponível em: <https://talkdigital.co/ianavidareal/index.html>.
- [Wang et al. 2023] WANG, Ze. A comprehensive survey of ChatGPT: advancements, applications, prospects, and challenges. Meta-radiology, v. 1, n. 2, p. 100022, 2023.
- [Zheng et al. 2025] ZHENG, Xi et al. Customizing emotional support: How do individuals construct and interact with LLM-powered chatbots. In: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. 2025. p. 1-20.