The Cost-Benefit of Interdisciplinarity in AI for Mental Health

Katerina Drakos[1] *Center for Social Data Science, Faculty of Social Sciences, University of Copenhagen* kda@sodas.ku.dk
Eva Paraschou[1] *Dept. of Applied Mathematics and Computer Science, Technical University of Denmark* evpa@dtu.dk
Simay Toplu[1] *Institute for Ethics in Artificial Intelligence, School of Social Sciences and Technology, Technical University of Munich* simay.toplu@tum.de
Line Harder Clemmensen *Dept. of Mathematical Sciences, University of Copenhagen* lkhc@math.ku.dk
Christoph Lütge *Institute for Ethics in Artificial Intelligence, School of Social Sciences and Technology, Technical University of Munich* luetge@tum.de
Nicole Nadine Lønfeldt *Child and Adolescent Mental Health Center, Copenhagen University Hospital* nicole.nadine.loenfeldt@regionh.dk
Sneha Das *Dept. of Applied Mathematics and Computer Science, Technical University of Denmark* sned@dtu.dk

Abstract

Artificial intelligence has been introduced as a way to improve access to mental health support. However, most AI mental health chatbots rely on a limited range of disciplinary input, and fail to integrate expertise across the chatbot's lifecycle. This paper examines the cost-benefit trade-off of interdisciplinary collaboration in AI mental health chatbots. We argue that involving experts from technology, healthcare, ethics, and law across key lifecycle phases is essential to ensure value-alignment and compliance with the high-risk requirements of the AI Act. We also highlight practical recommendations and existing frameworks to help balance the challenges and benefits of interdisciplinarity in mental health chatbots.

Keywords: Interdisciplinary collaboration; mental health; chatbot; human-computer interaction

1 Motivation & Background

Mental health care faces a global crisis, marked by rising demand (Organization et al. [2022]), especially among young people following COVID-19 (Xiong et al. [2020]), unequal access and increased stigma (Wainberg et al. [2017]). To increase accessibility, artificial intelligence (AI) has been applied in different mental health care services. Some applications focus on mental health literacy and offer psychoeducational content (Potts et al. [2023]). Others stimulate therapeutic interaction to help users manage their mental health (De Freitas et al. [2025]; Thieme et al. [2023]; Fitzpatrick et al. [2017]). These mental health

[1]These authors contributed equally to this work. Their names are listed in alphabetical order by surname.

chatbots can serve as low-threshold entry points to care, increasing accessibility particularly for underserved populations (Pozzi and De Proost [2025]) and individuals hesitant to seek professional help (Hoffman et al. [2024]).

AI in mental health care must address usability and clinical benefit, yet no discipline can achieve this alone (Boulos et al. [2021]; Werth et al. [2025]). Developing mental health chatbots requires expertise from multiple disciplines, such as health, psychology, psychiatry, human-computer interaction (HCI), computer science, ethics, and law (Kuhn et al. [2024]; Spahl and Rubeis [2025]; Floridi et al. [2018]; Morrin et al. [2025]). Interdisciplinary teams help ensure contextually appropriate chatbots (Pagliari et al. [2007]), and generate new cross-sectoral research directions (Molina Recio et al. [2016]). Collaboration is crucial, as chatbots built from a single disciplinary perspective can overlook risks such as privacy, safety and accountability concerns which demand joint expertise (Floridi et al. [2018]). Thus, a multilevel, justice-oriented and interdisciplinary approach is essential to ensure inclusivity and responsiveness to societal needs (Kuhn et al. [2024]).

A 2023 survey on conversational agents in mental health revealed that only 19.6% of chatbots employed interdisciplinary teams (Cho et al. [2023]). Exploring the current cost-benefit trade-off of interdisciplinary collaboration in this field, Table 1 provides a short overview of how experts from technology, healthcare, ethics or law (as suggested in the literature (Kuhn et al. [2024]; Spahl and Rubeis [2025];Floridi et al. [2018]; Morrin et al. [2025])) are integrated across mental health chatbots' lifecycle (design, development, evaluation). We employed a convenience sampling strategy on papers from 2022-2025 (to build on the findings of Cho et al. [2023]) from MDPI with inclusion criteria consisting of "chatbot" and "mental health" in the title or keywords, as well as on other papers from different publishers presenting chatbots that explicitly applied interdisciplinary collaboration. After abstract screening of 16 papers, Table 1 presents the papers included in our overview.

Table 1: An overview of interdisciplinary collaboration in recent chatbots. The letters represent the first letter of each expert group.

	Experts				**Lifecycle Phase**		
Chatbot	Technology	Healthcare	Ethics	Law	Design	Development	Evaluation
Manole et al. [2024]	X	✓	X	X	-	-	-
Kamdan et al. [2025]	✓	X	X	X	-	-	-
Rathnayaka et al. [2022]	✓	✓	✓	X	T-H-E	-	-
Hall et al. [2022]	✓	✓	X	X	T-H	T-H	T-H
Noble et al. [2022]	✓	✓	X	X	T-H	T-H	T-H
Chua et al. [2023]	✓	✓	X	X	T-H	T-H	T-H
Olla et al. [2025]	✓	✓	✓	X	T-H-E	T-H-E	T-H-E

Despite some examples of integration, some chatbots still operate without interdisciplinary input, focusing solely on technology or healthcare (Manole et al. [2024]; Kamdan et al. [2025]), potentially reflecting the challenges outlined earlier. Most reviewed chatbots (Noble et al. [2022]; Hall et al. [2022]; Chua et al. [2023]) include only technology and healthcare experts across all lifecycle phases. While Rathnayaka et al. [2022] integrates experts in technology, healthcare, and ethics, this interdisciplinary collaboration is limited to the design phase, whereas

Olla et al. [2025] extend this collaboration throughout the lifecycle. *In sum, our mental health chatbots overview (Table 1) suggests a potential cost-benefit trade-off: even where interdisciplinary teams exist, collaboration often appears limited to specific phases and does not cover a wide range of disciplines.*

Implementing interdisciplinarity in mental health chatbots is often difficult. Challenges include misaligned expectations, limited cross-disciplinary training, conflicting timelines, and tensions between business and scientific standards (Pagliari et al. [2007]). Communication barriers such as different terminology and professional priorities add to these challenges (Doherty et al. [2010]; Blandford et al. [2018]). Structural barriers such as restricted funding across fields (Cibrian et al. [2022]) and publishing norms discouraging interdisciplinary work (Yegros-Yegros et al. [2015]) further constrain collaboration, reducing adoption and impact throughout the chatbot's lifecycle (Cibrian et al. [2022]).

2 Position Statement

Despite recognizing the challenges of interdisciplinary collaboration in mental health chatbots, *we argue for deliberately embedding technology, healthcare, ethics, and law experts in the most impactful lifecycle phases of mental health chatbots, to guide ethical and compliant implementation.* Our position is founded on the recent EU AI Act (European Commission [2021]), which classifies mental health AI chatbots as high-risk and sets strict requirements for transparency, human oversight and risk management. Meeting these obligations requires expertise in technology, healthcare, and law to accurately translate regulatory standards into practice (Montag and Finck [2024]). Moreover, frameworks such as Value Sensitive Design (Friedman [1996]) and ethics-by-design (Van Wynsberghe [2020]), encourage the integration of ethical and social values throughout the systems' (including chatbots') lifecycle, underscoring additional input from ethicists. Therefore, we argue that meeting regulatory standards can be supported by these ethical frameworks by guiding phased contributions from legal and ethical experts alongside technical and clinical teams.

In this position paper, we address the broad mental health domain instead of delving deeper into specific mental health disorders. Thus, our position emphasizes the generic disciplines highlighted as the most influential ones in the literature of digital health interventions for mental health. We acknowledge that each individual mental health condition requires additional expertise in the specific field. An example of this is involving psychiatry professionals with expertise in schizophrenia as integral components for chatbots targeting adults with schizophrenia.

To further elaborate on our position, Figure 1 illustrates how phased contributions from different experts can potentially balance the cost-benefit trade-off. Technology experts guide the entire system implementation with HCI principles including human-in-the-loop methods and user-centric evaluations, translate values into technical constructs, develop the software, and assess the chatbot's technical performance. Mental healthcare professionals identify user

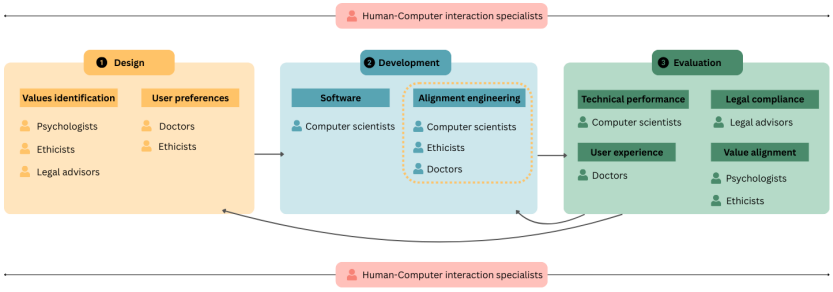

Figure 1: A deliberate interdisciplinary collaboration in a mental health chatbot.

needs, advise on implementation, and evaluate benefits, risks, and user satisfaction. Ethicists define ethical and social values, support implementation, and ensure value-alignment. Legal advisors specify regulatory requirements and ensure compliance. While we believe this interdisciplinary team is comprehensive and pragmatic, we acknowledge that the field is evolving and additional experts may be needed in the future.

3 Recommendations

Based on literature, we highlight a few guidelines and best-practices to balance the costs and benefits and hence enable interdisciplinary collaboration in mental health chatbots.

Interdisciplinary methodologies and frameworks: The framework developed by Hopkin et al. [2025] to categorize mental health AI systems could help clarify key characteristics, and thus relevant experts. A tangible interdisciplinarity evaluation measure (e.g., Rao–Stirling (Leydesdorff et al. [2019])) and cross-sectoral methodologies could support effective interdisciplinary collaborations (Boulos et al. [2021]; Molina Recio et al. [2016]).

EU AI Act: Frameworks such as value sensitive design (Friedman [1996]) and ethics-by-design (Van Wynsberghe [2020]) can offer practical methods for interdisciplinary teams to integrate ethical and social values across the AI systems' lifecycle and anticipate risks in high-risk AI systems, as imposed by the AI Act.

Expert values and reflections: Professionals should practice trust and mutual respect for differences, learn about methodologies from other disciplines, discuss any misconceptions and biases (Pagliari et al. [2007]); (Boulos et al. [2021]), and undertake training to familiarize themselves with different disciplines (Moltrecht et al. [2022]).

4 Conclusion

AI mental health chatbots require interdisciplinary collaboration for clinical benefit and value alignment. However, such collaborations face challenges and create a cost-benefit trade-off evident in today's chatbots, which have a narrow focus on experts and lifecycle phases. We argue for deliberately embedding technology, healthcare, ethics and law experts in a phased manner to guide the ethical and compliant implementation of mental health chatbots, with ethical design frameworks supporting this interdisciplinary collaboration across the lifecycle, value-alignment, and compliance with the high-risk AI Act requirements. While our position is theoretically sound, a critical future research dimension is the empirical analysis of the cost-benefit trade-off of involving different experts across the chatbot lifecycle, which will be necessary to develop concrete, practical recommendations for interdisciplinary collaboration.

Acknowledgments

Funded by the European Union grants under the Marie Skłodowska-Curie Grant Agreements No: 101169473 (alignAI). Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the European Research Executive Agency can be held responsible for them.

References

A. Blandford, J. Gibbs, N. Newhouse, O. Perski, A. Singh, and E. Murray. Seven lessons for interdisciplinary research on interactive digital health interventions. *Digital health*, 4:2055207618770325, 2018.

L. J. Boulos, A. Mendes, A. Delmas, and I. Chraibi Kaadoud. An iterative and collaborative end-to-end methodology applied to digital mental health. *Frontiers in Psychiatry*, 12:574440, 2021.

Y.-M. Cho, S. Rai, L. Ungar, J. Sedoc, and S. C. Guntuku. An integrative survey on mental health conversational agents to bridge computer science and medical perspectives. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*, volume 2023, page 11346, 2023.

J. Y. X. Chua, M. Choolani, C. Y. I. Chee, H. Yi, Y. H. Chan, J. G. Lalor, Y. S. Chong, and S. Shorey. 'parentbot–a digital healthcare assistant (pda)': A mobile application-based perinatal intervention for parents: Development study. *Patient Education and Counseling*, 114:107805, 2023.

F. L. Cibrian, E. Monteiro, S. E. Schuck, M. Nelson, G. R. Hayes, and K. D. Lakes. Interdisciplinary tensions when developing digital interventions supporting individuals with adhd. *Frontiers in Digital Health*, 4:876039, 2022.

J. De Freitas, Z. Oğuz-Uğuralp, A. K. Uğuralp, and S. Puntoni. Ai companions reduce loneliness. *Journal of Consumer Research*, page ucaf040, 2025.

G. Doherty, D. Coyle, and M. Matthews. Design and evaluation guidelines for mental health technologies. *Interacting with computers*, 22(4):243–252, 2010.

European Commission. Proposal for a regulation of the european parliament and of the council, laying down harmonised rules on artificial intelligence. `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206`, 2021. Document CELEX: 52021PC0206.

K. K. Fitzpatrick, A. Darcy, and M. Vierhile. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. *JMIR mental health*, 4(2):e7785, 2017.

L. Floridi, J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, R. Madelin, U. Pagallo, F. Rossi, et al. Ai4people—an ethical framework for a good ai society: Opportunities, risks, principles, and recommendations. *Minds and machines*, 28(4):689–707, 2018.

B. Friedman. Value-sensitive design. *interactions*, 3(6):16–23, 1996.

J. Hall, S. Jordan, M. van Ommeren, T. Au, R. A. Sway, J. Crawford, H. Ghalayani, S. U. Hamdani, N. P. Luitel, A. Malik, et al. Sustainable technology for adolescents and youth to reduce stress (stars): a who transdiagnostic chatbot for distressed youth. *World Psychiatry*, 21(1):156, 2022.

B. D. Hoffman, M. L. Oppert, and M. Owen. Understanding young adults' attitudes towards using ai chatbots for psychotherapy: The role of self-stigma. *Computers in Human Behavior: Artificial Humans*, 2(2):100086, 2024.

G. Hopkin, H. Coole, F. Edelmann, L. Ayiku, R. Branson, P. Campbell, S. Cooper, and M. Salmon. Toward a new conceptual framework for digital mental health technologies: Scoping review. *JMIR Mental Health*, 12: e63484, 2025.

K. Kamdan, N. G. Fauziyah, M. A. Fadlullah, D. A. Hanif, and I. L. Kharisma. Early mental health detection and emotional states in teenagers through chatbot systems using natural language processing (nlp). *Engineering Proceedings*, 107(1):64, 2025.

E. Kuhn, M. Saleem, T. Klein, C. Köhler, D. C. Fuhr, S. Lahutina, A. Minarik, R. Musesengwa, K. Neubauer, L. Olisaeloka, et al. Interdisciplinary perspectives on digital technologies for global mental health. *PLOS Global Public Health*, 4(2):e0002867, 2024.

L. Leydesdorff, C. S. Wagner, and L. Bornmann. Interdisciplinarity as diversity in citation patterns among journals: Rao-stirling diversity, relative variety, and the gini coefficient. *Journal of informetrics*, 13(1):255–269, 2019.

A. Manole, R. Cârciumaru, R. Brînzaș, and F. Manole. Harnessing ai in anxiety management: a chatbot-based intervention for personalized mental health support. *Information*, 15(12):768, 2024.

G. Molina Recio, L. Garcia-Hernandez, R. Molina Luque, and L. Salas-Morera. The role of interdisciplinary research team in the impact of health apps in health and computer science publications: a systematic review. *Biomedical engineering online*, 15(Suppl 1):77, 2016.

B. Moltrecht, P. Patalay, H. A. Bear, J. Deighton, and J. Edbrooke-Childs. A transdiagnostic, emotion regulation app (eda) for children: design, development, and lessons learned. *JMIR Formative Research*, 6(1):e28300, 2022.

C. Montag and M. Finck. Successful implementation of the eu ai act requires interdisciplinary efforts. *Nature Machine Intelligence*, 6(12):1415–1417, 2024.

H. Morrin, L. Nicholls, M. Levin, J. Yiend, U. Iyengar, F. DelGuidice, S. Bhattacharyya, J. MacCabe, S. Tognin, R. Twumasi, et al. Delusions by design? how everyday ais might be fuelling psychosis (and what can be done about it), 2025.

J. M. Noble, A. Zamani, M. Gharaat, D. Merrick, N. Maeda, A. L. Foster, I. Nikolaidis, R. Goud, E. Stroulia, V. I. Agyapong, et al. Developing, implementing, and evaluating an artificial intelligence–guided mental health resource navigation chatbot for health care workers and their families during and following the covid-19 pandemic: protocol for a cross-sectional study. *JMIR Research Protocols*, 11(7):e33717, 2022.

P. Olla, A. Barnes, L. Elliott, M. Abumeeiz, V. Olla, and J. Tan. Deploying a mental health chatbot in higher education: The development and evaluation of luna, an ai-based mental health support system. *Computers*, 14(6):227, 2025.

W. H. Organization et al. Who highlights urgent need to transform mental health and mental health care, 2022.

C. Pagliari et al. Design and evaluation in ehealth: challenges and implications for an interdisciplinary field. *Journal of medical Internet research*, 9(2):e614, 2007.

C. Potts, F. Lindström, R. Bond, M. Mulvenna, F. Booth, E. Ennis, K. Parding, C. Kostenius, T. Broderick, K. Boyd, et al. A multilingual digital mental health and well-being chatbot (chatpal): pre-post multicenter intervention study. *Journal of medical Internet research*, 25:e43051, 2023.

G. Pozzi and M. De Proost. Keeping an ai on the mental health of vulnerable populations: Reflections on the potential for participatory injustice. *AI and Ethics*, 5(3):2281–2291, 2025.

P. Rathnayaka, N. Mills, D. Burnett, D. De Silva, D. Alahakoon, and R. Gray. A mental health chatbot with cognitive skills for personalised behavioural activation and remote health monitoring. *Sensors*, 22(10):3653, 2022.

W. Spahl and G. Rubeis. Integrating ethics in digital mental healthcare technologies: a principle-based empirically grounded roadmap approach. *Medicine, Health Care and Philosophy*, pages 1–14, 2025.

A. Thieme, M. Hanratty, M. Lyons, J. Palacios, R. F. Marques, C. Morrison, and G. Doherty. Designing human-centered ai for mental health: Developing clinically relevant applications for online cbt treatment. *ACM Transactions on Computer-Human Interaction*, 30(2):1–50, 2023.

A. Van Wynsberghe. Designing robots for care: Care centered value-sensitive design. In *Machine ethics and robot ethics*, pages 185–211. Routledge, 2020.

M. L. Wainberg, P. Scorza, J. M. Shultz, L. Helpman, J. J. Mootz, K. A. Johnson, Y. Neria, J.-M. E. Bradford, M. A. Oquendo, and M. R. Arbuckle. Challenges and opportunities in global mental health: a research-to-practice perspective. *Current psychiatry reports*, 19(5):28, 2017.

O. Werth, I. Jankowski, N. S. Müller, F. Schulte, P. Warnemünde-Jagau, M. H. Breitner, A. Herr, and K. G. Kahl. Success by design: A holistic analysis of design principles for depression apps. *Digital Health*, 11:20552076251336276, 2025.

J. Xiong, O. Lipsitz, F. Nasri, L. M. Lui, H. Gill, L. Phan, D. Chen-Li, M. Iacobucci, R. Ho, A. Majeed, et al. Impact of covid-19 pandemic on mental health in the general population: A systematic review. *Journal of affective disorders*, 277:55–64, 2020.

A. Yegros-Yegros, I. Rafols, and P. D'este. Does interdisciplinary research lead to higher citation impact? the different effect of proximal and distal interdisciplinarity. *PloS one*, 10(8):e0135095, 2015.