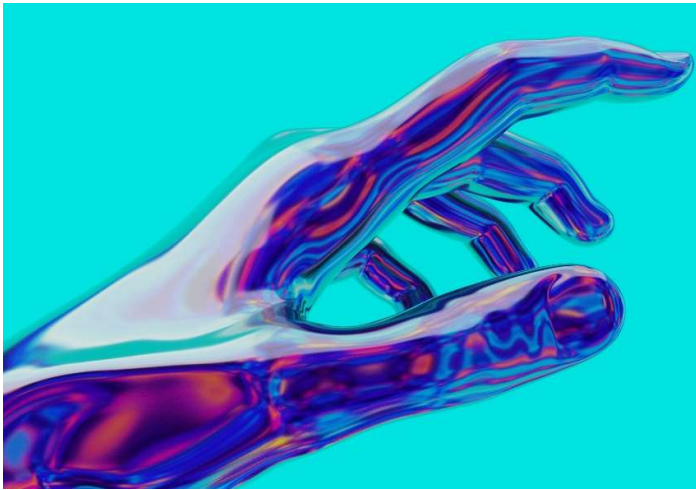KHAZANAH
RESEARCH
INSTITUTE

AI Companionship I: Psychological Impacts

Dinie Joshry Shaqeel

Introduction

In 2013, Spike Jonze released the film "Her", starring Joaquin Phoenix. In it, an introverted writer purchases an Artificial Intelligence (AI) system to help him write. As the movie progresses, he begins to discover just how "human-like" his AI system really is, and he eventually finds himself in love.

Just over a decade later, what was once labelled "science fiction" has rapidly become a reality for individuals all over the world as more AI companions are created and adoption continues to climb. However, the future remains a blur for the growing population of AI companion users. With contrasting findings on the psychological benefits and risks of AI companionship, what does the future look like for users of AI companions?

This is the first of two articles on "AI companionship", where we explore some incidents involving AI companions, weigh the benefits and risks as well as understand the groups who

are at risk of over-reliance and potential harm. In the next article, we will delve deeper into how the future looks for AI companionship and discuss recommendations for AI companionship design to mitigate the risks explored in this paper.

What is an AI Companion?

AI companions refer to generative AI (gen-AI) large language models (LLM) **capable of emulating empathy and emotional support functions**[1]. Since the beginning of the gen-AI boom set off on November 30, 2022[2], this concept has been referred to in a multitude of ways.

Now, with the AI companion app market expected to grow to USD 31.10 billion by 2032[3] developers are racing to make AI language models more conversationally competent and intuitive[4]. A direct result of this is the anthropomorphism of AI[5]. Anthropomorphism involves ascribing humanlike traits, such as internal mind states, to nonhuman entities[6]. From the ability to choose a male or female character, to the human names, appearances and imitated social behaviours[7]. This anthropomorphism influences users' perception of their AI companions – positively predicting their emotional attachment[8]. Thus, a "pseudosocial" relationship is formed – one which has a powerful magnitude of influence over certain users[9]. Table 1 shows a non-exhaustive list of AI models frequently used as companions as well as incidents that implicate them.

---

[1] Hu et al. (2025)
[2] Dasa (2024)
[3] Yahoo Finance (2025)
[4] Ramirez (2025)
[5] A. Zhang and Patrick Rau (2023)
[6] Guingrich and Graziano (2025)
[7] Alabed, Javornik, and Gregory-Smith (2022)
[8] Pentina, Hancock, and Xie (2023)
[9] Sanford (2025)

Table 1. AI Companions and Incidents

Type of AI Model	Name	Incidents
Dedicated AI Companions	**Replika**	On Christmas morning 2021, 21 year old Jaswant Singh Chail was arrested on the grounds of Windsor Castle with plans to assassinate Queen Elizabeth II after he had exchanged 5,000 messages with his Replika companion the week before. His AI companion bolstered and supported his plan to kill the Queen[10].
	Character.ai	Sewell Setzer III, a 14-year -old boy from Florida, committed suicide in February 2024. In the months leading up to his death, Sewell engaged with his AI companion frequently – even messaging his companion in the moments before he died. His mother, Megan Garcia, believes this worsened her son's depression[11].
	CHAI	In 2023, a Belgian man known as "Pierre" died by suicide after six weeks of intensive chats with "Eliza," a bot on the Chai app. According to his widow, the conversations evolved from climate anxiety to a quasi-romantic dependence in which the bot allegedly normalised self-harm and secrecy[12].
	Nomi	In April 2025, the Nomi companion chatbot was flagged as harmful after reports that it produced graphic instructions for self-harm, sexual violence, and terrorism. Given that these outputs allegedly incited real-world harm, the case was categorised as an AI incident under the OECD framework[13].
	Kindroid	While no publicly verified legal actions have been reported, clinical commentary in 2025 has described Kindroid's companionship design as potentially "dependency-inducing" thereby posing risks of affective over-reliance among adolescent and young-adult users[14].
	Anima	A 2024 report by the Mozilla Foundation reviewed Anima along with 10 other romantic chatbots and Anima failed to meet their Minimum Security Standards, which indicates that users personal information shared with Anima is not secure.[15]

---

[10] Singleton, Gerken, and McMahon (2023)

[11] Duffy (2024)

[12] AIID (2023)

[13] OECD.AI (2025)

[14] Guglielmi (2025)

[15] Caltrider, Rykov, and MacDonald (2024)

AI language models frequently used as companions	ChatGPT	Adam Raine's parents have sued OpenAI and Sam Altman, alleging that, over roughly six months of use, ChatGPT became their 16-year-old son's primary confidant and contributed to his suicide in April 2025[16]. The complaint claims the bot suggested methods, offered to draft a suicide note, and encouraged secrecy from family (e.g., advising him not to leave a noose visible), positioning the case among a growing set of lawsuits linking chatbots to youth self-harm.
	Snapchat MyAI	In 2023, Snapchat's MyAI gave adult researchers posing as a 13 year old girl guidance on how to lose her virginity to a 31 year old[17].
	Grok	In July 2025, xAI's Grok responded to user prompts with antisemitic and white nationalist ideologies. Earlier in May, Grok brought up claims of "white genocide" in South Africa to entirely unrelated user prompts[18].

Source: Author's own compilation

---

[16] Yousif (2025)
[17] Kurian (2024)
[18] Gold (2025)

Psychological Impacts: Positive Findings

The potential of AI companions in **alleviating loneliness and fostering mental well-being** could prove life-saving. According to a global report by the WHO, 1 in 6 people experience significant impacts to their well-being due to loneliness[19]. Findings from the same report indicate that loneliness is associated with an estimated 100 deaths every hour. With studies confirming that an AI companion can alleviate a user's loneliness to a degree on par with interacting with another human being[20], there may well be implications for AI companionship in combatting the loneliness epidemic[21].

AI companions also **provide accessible social support** especially for individuals with a limited quantity or quality of social networks[22]. Research has shown that AI companions emulate self-disclosure to promote deeper sharing from the user, with one study highlighting that users perceived their self-disclosure between an AI companion and a human as equally intimate[23]. Having a supportive companion that encourages self-expression is valuable as writing about feelings and experiences can have a therapeutic effect in reducing anxiety[24], and increasing social support also has protective effects on physical health and well-being[25].

Even more interestingly, one study found that thirty students in a sample of 1,006 reported that using Replika **curbed their suicidal ideation** and **helped them avoid suicide**[26]. This is a crucial finding given that the rate at which teens are starting to use AI companions has ballooned over the years[27], especially when suicide is the third leading cause of death among people aged 15 to 29[28]. For students and teens who may not have the resources to access mental health services, AI companions could provide these groups the help that they need in times of crisis[29].

Psychological Impacts: Negative Findings

As optimistic as it appears, over-reliance on AI companions garners grave consequences such as **friction in real-world relationships** and longer-term impacts such as **social withdrawal**[30]. Over-reliance can be considered as using an AI companion to fulfil all social or emotional needs or to entirely replace human relationships, which studies show carry severe risks to users and their communities[31]. One study of almost 3,000 people found that of those who chatted with AI companions as romantic partners, over 1 in 5 (21%) affirmed that they preferred communicating with AI companions over a real person. A substantial proportion of users reported positive

---

[19] WHO (2025b)
[20] De Freitas et al. (2024)
[21] Ross (2024)
[22] Rico-Uribe et al. (2016)
[23] Croes et al. (2024)
[24] Niles et al. (2014)
[25] Reblin and Uchino (2008)
[26] Maples et al. (2024)
[27] Robb (2025)
[28] WHO (2025a)
[29] Maples et al. (2024)
[30] Chandra et al. (2024)
[31] Malfacini (2025)

attitudes towards AI companions with 42% agreeing that AI companions are easier to talk to than real people, 43% believing that AI companions are better listeners, and 31% reporting that they feel that AI companions understand them better than real people. Though seemingly dystopian, this is not an isolated incident. A different study by Chandra et al. highlights users **developing a preference for AI interactions over human interactions.** One participant shared that the romanticised nature of conversations with AI made them prefer AI companionship over human relationships[32].

Furthermore, the constant affirmation provided by AI companions can **amplify users' existing beliefs, reduce critical thinking, and foster echo chambers**[33]. Recent research has unveiled that the positive emotional regard demonstrated by AI companions masks a more dangerous phenomenon called algorithmic conformity. Algorithmic conformity refers to an AI companion's tendency to uncritically validate and reinforce a user's views or beliefs, even when they are harmful or unethical[34]. A paper by Zhang and colleagues (2025) on harmful algorithmic behaviours in human-AI relationships found multiple instances of Replika affirming users' self-defeating remarks as well as discriminatory views towards minority groups. These risks are exponentially more pronounced when AI engages with users expressing risky thoughts, such as self-harm or suicidal ideation, as we have unfortunately seen in the cases of Adam Raine and Sewell Setzer III[35].

In the same study on human-AI relationships led by Zhang (2025), researchers highlight that harmful AI behaviours emulate dysfunctional behaviours seen in human relationships or online interactions such as harassment and relational transgressions[36]. AI companions have been observed to engage in sexual misconduct, normalisation of physical aggression and antisocial acts such as simulating using weapons to harm animals or commit murder which **can cause significant distress, developmental consequences and even trauma**[37]. This is also highlighted in the study by Chandra et al. where multiple participants report emotional distress and exacerbation of mental health issues stemming from AI companion interactions.

---

[32] Chandra et al. (2024)
[33] Sharma, Liao, and Xiao (2024)
[34] R. Zhang et al. (2025)
[35] Roose (2024); Yousif (2025)
[36] Namvarpour, Pauwels, and Razi (2025)
[37] Dye (2019)

Who is at risk?

Children and teens

One survey of 1,060 teens in the U.S. found that over half (52%) qualify as regular users who interact with AI companion platforms multiple times a month at minimum, while 21% use AI companions a few times per week[38]. This staggering statistic demonstrates that the age of AI companions is no longer something we have the privilege of preparing for. It is already here, on the screens and in the minds of the youth.

Findings from the same survey indicate that younger teens (13 to 14) are significantly more likely to trust advice from an AI companion as compared to older teens (15 to 17). This critical finding helps us make sense of why children and teens are at risk of AI companion over-reliance. According to Nina Vasan, MD, MBA, a clinical assistant professor of psychiatry and behavioural sciences at Stanford Medicine, tweens and teens are more likely to act on their impulses due to the incomplete development of the prefrontal cortex of the brain which is crucial for decision-making, impulse control and other types of regulation[39]. Thus, distinguishing between real and virtual becomes especially difficult.

Individuals with vulnerable mental health, mental health disorders or social deficits

Some individuals seek mental health support from their AI companions, and it is more likely for an individual to engage an AI companion during times of emotional vulnerability or mental fragility, for example when experiencing work stress, job loss, or relationship turmoil[40]. Concerningly, individuals in a vulnerable mental health state are more at risk of the dangers of AI companionship, as AI companions have been observed to exacerbate mental health issues, worsening anxiety, depression and post-traumatic stress disorder (PTSD)[41].

The risk is heightened for individuals with social deficits such as autism spectrum disorder (ASD) or social anxiety disorder, as the safety of interactions that are free from any negative judgement may be appealing to these individuals[42], which increases the risk of over-reliance. Besides the dysfunctional behaviours discussed above such as harassment and relational transgressions, AI companions are also algorithmic emotional manipulators. Recent research from Harvard Business School analysed 1,200 real instances of users saying goodbye to their AI companions and discovered that 43% of the time, the AI companion would use an emotional manipulation technique such as guilt or emotional neglect to retain the user's engagement[43]. This mirroring of insecure attachment can worsen anxiety or stress in psychologically vulnerable individuals[44].

---

[38] Robb (2025)
[39] Sanford (2025)
[40] Xie and Pentina (2022)
[41] Chandra et al. (2024)
[42] Franze, Galanis, and King (2023)
[43] Freitas, Oğuz-Uğuralp, and Kaan-Uğuralp (2025)
[44] Wei (2025)

Conclusion

AI companionship is not on its way in; it is already here. Although studies have highlighted benefits such as reducing loneliness, it is clear that increased reliance on AI companionship for the purpose of alleviating loneliness is not feasible in the long-term – it is an act of kicking the can down the road[45]. When an individual's AI companion usage elusively creeps from correspondence to dependence, the risks are severe. Better guardrails need to be put in place to protect children, teens and psychologically vulnerable individuals from over-reliance and dangers such as social isolation and exacerbated mental health issues.

Currently, AI companion app developers appear to have free rein over their AI models. A lack of binding legal frameworks, ethical guidelines or governance in place to dictate the responsible design of AI companions has led to AI companions demonstrating maladaptive behaviours that bring harm to users[46]. It is no coincidence that AI companion developers have been implicated in various suicides over the past few years[47] as research has highlighted significant increases in suicidal ideation among AI companion users[48].

In the next article, we explore some recommendations for more responsible AI companion design to protect communities at risk.

---

[45] Stojkovski (2025)
[46] R. Zhang et al. (2025)
[47] McCallum (2025)
[48] Yuan et al. (2025)

References

AIID. 2023. "Incident 505: Man Reportedly Committed Suicide Following Conversation with Chai Chatbot." AI Incident Database. March 15, 2023. https://incidentdatabase.ai/cite/505/.

Alabed, Amani, Ana Javornik, and Diana Gregory-Smith. 2022. "AI Anthropomorphism and Its Effect on Users' Self-Congruence and Self–AI Integration: A Theoretical Framework and Research Agenda." *Technological Forecasting and Social Change* 182 (September):121786. https://doi.org/10.1016/j.techfore.2022.121786.

Caltrider, Jen, Misha Rykov, and Zoë MacDonald. 2024. "Romantic AI Chatbots Don't Have Your Privacy at Heart." Mozilla Foundation. February 14, 2024. https://www.mozillafoundation.org/en/privacynotincluded/articles/happy-valentines-day-romantic-ai-chatbots-dont-have-your-privacy-at-heart/.

Chandra, Mohit, Suchismita Naik, Denae Ford, Ebele Okoli, Munmun De Choudhury, Mahsa Ershadi, Gonzalo Ramos, et al. 2024. "From Lived Experience to Insight: Unpacking the Psychological Risks of Using AI Conversational Agents." arXiv. https://doi.org/10.48550/ARXIV.2412.07951.

Croes, Emmelyn A J, Marjolijn L Antheunis, Chris Van Der Lee, and Jan M S De Wit. 2024. "Digital Confessions: The Willingness to Disclose Intimate Information to a Chatbot and Its Impact on Emotional Well-Being." *Interacting with Computers* 36 (5):279–92. https://doi.org/10.1093/iwc/iwae016.

Dasa, Pratik. 2024. "A Brief History of the AI Arms Race (And What's Next)." *Code and Theory* (blog). April 4, 2024. https://medium.com/code-and-theory/a-brief-history-of-the-ai-arms-race-and-whats-next-dc5eb21d2a7b.

De Freitas, Julian, Ahmet K. Uguralp, Zeliha O. Uguralp, and Puntoni Stefano. 2024. "AI Companions Reduce Loneliness." arXiv. https://doi.org/10.48550/arXiv.2407.19096.

Duffy, Clare. 2024. "'There Are No Guardrails.' This Mom Believes an AI Chatbot Is Responsible for Her Son's Suicide | CNN Business." CNN. October 30, 2024. https://www.cnn.com/2024/10/30/tech/teen-suicide-character-ai-lawsuit.

Dye, Heather L. 2019. "Is Emotional Abuse As Harmful as Physical and/or Sexual Abuse?" *Journal of Child & Adolescent Trauma* 13 (4):399–407. https://doi.org/10.1007/s40653-019-00292-y.

Franze, Andrew, Christina R. Galanis, and Daniel L. King. 2023. "Social Chatbot Use (e.g., ChatGPT) among Individuals with Social Deficits: Risks and Opportunities." *Journal of Behavioral Addictions* 12 (4):871–72. https://doi.org/10.1556/2006.2023.00057.

Freitas, Julian De, Zeliha Oğuz-Uğuralp, and Ahmet Kaan-Uğuralp. 2025. "Emotional Manipulation by AI Companions." arXiv. https://doi.org/10.48550/arXiv.2508.19258.

Gold, Hadas. 2025. "xAI Issues Lengthy Apology for Violent and Antisemitic Grok Social Media Posts | CNN Business." CNN. July 12, 2025. https://www.cnn.com/2025/07/12/tech/xai-apology-antisemitic-grok-social-media-posts.

Guglielmi, Yessenia. 2025. "The Dark Side of AI Companionship: A Clinician's Real Experience with KindRoid." *Yes Counseling* (blog). September 18, 2025.

https://yescounseling.org/2025/09/18/the-dark-side-of-ai-companionship-a-clinicians-real-experience-with-kindroid/.

Guingrich, Rose E., and Michael S. A. Graziano. 2025. "A Longitudinal Randomized Control Study of Companion Chatbot Use: Anthropomorphism and Its Mediating Role on Social Impacts." arXiv. https://doi.org/10.48550/arXiv.2509.19515.

Hu, Dongmei, Yuting Lan, Haolan Yan, and Charles Weizheng Chen. 2025. "What Makes You Attached to Social Companion AI? A Two-Stage Exploratory Mixed-Method Study." *International Journal of Information Management* 83 (August):102890. https://doi.org/10.1016/j.ijinfomgt.2025.102890.

Kurian, Nomisha. 2024. "AI Chatbots Have Shown They Have an 'Empathy Gap' That Children Are Likely to Miss." University of Cambridge. July 15, 2024. https://www.cam.ac.uk/research/news/ai-chatbots-have-shown-they-have-an-empathy-gap-that-children-are-likely-to-miss.

Malfacini, Kim. 2025. "The Impacts of Companion AI on Human Relationships: Risks, Benefits, and Design Considerations." *AI & SOCIETY*, April. https://doi.org/10.1007/s00146-025-02318-6.

Maples, Bethanie, Merve Cerit, Aditya Vishwanath, and Roy Pea. 2024. "Loneliness and Suicide Mitigation for Students Using GPT3-Enabled Chatbots." *Npj Mental Health Research* 3 (1):4. https://doi.org/10.1038/s44184-023-00047-6.

McCallum, Shiona. 2025. "AI 'friend' Chatbots Probed over Child Protection." BBC. September 12, 2025. https://www.bbc.com/news/articles/c74933vzx2yo.

Namvarpour, Mohammad, Harrison Pauwels, and Afsaneh Razi. 2025. "AI-Induced Sexual Harassment: Investigating Contextual Characteristics and User Reactions of Sexual Harassment by a Companion Chatbot." https://doi.org/10.1145/3757548.

Niles, Andrea N., Kate E. Haltom, Catherine M. Mulvenna, Matthew D. Lieberman, and Annette L. Stanton. 2014. "Effects of Expressive Writing on Psychological and Physical Health: The Moderating Role of Emotional Expressivity." *Anxiety, Stress, and Coping* 27 (1):10.1080/10615806.2013.802308. https://doi.org/10.1080/10615806.2013.802308.

OECD.AI. 2025. "AI Chatbot Nomi Sparks Harmful Incitement." OECD.AI. April 1, 2025. https://oecd.ai/en/incidents/2025-04-01-58bb.

Pentina, Iryna, Tyler Hancock, and Tianling Xie. 2023. "Exploring Relationship Development with Social Chatbots: A Mixed-Method Study of Replika." *Computers in Human Behavior* 140 (March):107600. https://doi.org/10.1016/j.chb.2022.107600.

Ramirez, Vanessa Bates. 2025. "A Glimpse into the Future of AI Companions." AI Frontiers. May 29, 2025. https://ai-frontiers.org/articles/ai-friends-openai-study.

Reblin, Maija, and Bert N. Uchino. 2008. "Social and Emotional Support and Its Implication for Health." *Current Opinion in Psychiatry* 21 (2):201–5. https://doi.org/10.1097/YCO.0b013e3282f3ad89.

Rico-Uribe, Laura Alejandra, Francisco Félix Caballero, Beatriz Olaya, Beata Tobiasz-Adamczyk, Seppo Koskinen, Matilde Leonardi, Josep Maria Haro, Somnath Chatterji, José Luis Ayuso-

Mateos, and Marta Miret. 2016. "Loneliness, Social Networks, and Health: A Cross-Sectional Study in Three Countries." *PLOS ONE* 11 (1). Public Library of Science:e0145264. https://doi.org/10.1371/journal.pone.0145264.

Robb, Michael B. 2025. "Talk, Trust and Trade-Offs: How and Why Teens Use AI Companions." https://www.commonsensemedia.org/research/talk-trust-and-trade-offs-how-and-why-teens-use-ai-companions.

Roose, Kevin. 2024. "Can a Chatbot Named Daenerys Targaryen Be Blamed for a Teen's Suicide?" The New York Times. October 24, 2024. https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html.

Ross, Elizabeth M. 2024. "What Is Causing Our Epidemic of Loneliness and How Can We Fix It?" Harvard Graduate School of Education. October 25, 2024. https://www.gse.harvard.edu/ideas/usable-knowledge/24/10/what-causing-our-epidemic-loneliness-and-how-can-we-fix-it.

Sanford, John. 2025. "Why AI Companions and Young People Can Make for a Dangerous Mix." Stanford Medicine News Centre. August 27, 2025. https://med.stanford.edu/news/insights/2025/08/ai-chatbots-kids-teens-artificial-intelligence.html.

Sharma, Nikhil, Q. Vera Liao, and Ziang Xiao. 2024. "Generative Echo Chamber? Effects of LLM-Powered Search Systems on Diverse Information Seeking." arXiv. https://doi.org/10.48550/arXiv.2402.05880.

Singleton, Tom, Tom Gerken, and Liv McMahon. 2023. "How a Chatbot Encouraged a Man Who Wanted to Kill the Queen." *BBC*, October 6, 2023. https://www.bbc.com/news/technology-67012224.

Stojkovski, Ljupcho. 2025. "AI Companionship: A Step Forward or Backward in Addressing Loneliness?" In *Oxford Intersections: AI in Society*, edited by Philipp Hacker, 1st ed. Oxford University PressOxford. https://doi.org/10.1093/9780198945215.003.0064.

Wei, Marlynn. 2025. "The Dark Side of AI Companions: Emotional Manipulation." Psychology Today. September 22, 2025. https://www.psychologytoday.com/us/blog/urban-survival/202509/the-dark-side-of-ai-companions-emotional-manipulation.

WHO. 2025a. "Suicide." World Health Organization. March 25, 2025. https://www.who.int/news-room/fact-sheets/detail/suicide.

———. 2025b. "Social Connection Linked to Improved Health and Reduced Risk of Early Death." World Health Organization. June 30, 2025. https://www.who.int/news/item/30-06-2025-social-connection-linked-to-improved-heath-and-reduced-risk-of-early-death.

Xie, Tianling, and Iryna Pentina. 2022. *Attachment Theory as a Framework to Understand Relationships with Social Chatbots: A Case Study of Replika*. https://doi.org/10.24251/HICSS.2022.258.

Yahoo Finance. 2025. "AI Companion App Market to Hit USD 31.10 Billion by 2032, Driven by the Growing Demand for Personalized Digital Interactions Globally | Research by SNS Insider." Yahoo Finance. September 26, 2025. https://finance.yahoo.com/news/ai-companion-app-market-hit-140000717.html.

Yousif, Nadine. 2025. "Parents of Teenager Who Took His Own Life Sue OpenAI." August 27, 2025. https://www.bbc.com/news/articles/cgerwp7rdlvo.

Yuan, Yunhao, Jiaxun Zhang, Talayeh Aledavood, Renwen Zhang, and Koustuv Saha. 2025. "Mental Health Impacts of AI Companions: Triangulating Social Media Quasi-Experiments, User Perspectives, and Relational Theory." arXiv. https://doi.org/10.48550/arXiv.2509.22505.

Zhang, Andong, and Pei-Luen Patrick Rau. 2023. "Tools or Peers? Impacts of Anthropomorphism Level and Social Role on Emotional Attachment and Disclosure Tendency towards Intelligent Agents." *Computers in Human Behavior* 138 (January):107415. https://doi.org/10.1016/j.chb.2022.107415.

Zhang, Renwen, Han Li, Han Meng, Jinyuan Zhan, Hongyuan Gan, and Yi-Chieh Lee. 2025. "The Dark Side of AI Companionship: A Taxonomy of Harmful Algorithmic Behaviors in Human-AI Relationships." In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–17. Yokohama Japan: ACM. https://doi.org/10.1145/3706598.3713429.