"It's Just Between Us… Right?" Navigating Teen Privacy Challenges in AI Companion Conversations

HSUEN-CHI (HAZEL) CHIU, Purdue University, USA
JEREMY FOOTE, Purdue University, USA

Teens increasingly turn to AI companions for emotional support, often treating them as trusted confidants. However, these interactions blur the line between interpersonal privacy and corporate data collection, raising concerns about how young users understand and manage their privacy. This article applies Communication Privacy Management (CPM) theory and dimensional privacy to explore how teens navigate privacy in their relationships with AI companions. By examining these dynamics, we aim to raise awareness and spark discussion on the ethical, design, and educational implications of AI companion use among teens.

Additional Key Words and Phrases: AI Companions, Human-AI Communication, Privacy Management, Adolescents, Digital Privacy, Data Ethics, Teen Technology Use

ACM Reference Format:
Hsuen-Chi (Hazel) Chiu and Jeremy Foote. 2025. "It's Just Between Us… Right?" Navigating Teen Privacy Challenges in AI Companion Conversations. In *Proceedings of (CHI)*. ACM, New York, NY, USA, 6 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

The rise of AI companion technology, such as Character.AI and Kindroid, is reshaping how teenagers experience intimacy and navigate privacy in the digital age [27]. These AI systems, designed to provide emotional support and engagement, have become an integral part of some adolescents' social lives, offering a judgment-free space to express their thoughts and emotions [27, 32]. However, the very nature of these interactions presents new challenges, as personal disclosures made to AI companions are not truly private but instead become part of corporate data collection practices [14].

Teenagers, who are in a critical phase of developing their understanding of privacy, intimacy, and relationship boundaries, may be particularly vulnerable to these dynamics [18]. Research indicates that teens often feel more comfortable sharing personal information with AI companions than with human confidants, perceiving the AI as a safe, nonjudgmental listener [12, 13, 22]. The integration of anthropomorphic features in these systems further increases the likelihood of deep personal disclosures, as young users may develop a sense of friendship with the AI [4, 14]. While this openness can provide emotional relief and support, it also exposes teens to significant privacy risks, as every confession and moment of vulnerability is stored as data, potentially used for purposes far removed from the original intimate exchange [11].

This paradox—where interactions that feel deeply personal are, in reality, corporate transactions—creates a complex privacy landscape for young users. Unlike human relationships, AI companions do not offer true confidentiality; instead,

Authors' Contact Information: Hsuen-Chi (Hazel) Chiu, chiu101@purdue.edu, Purdue University, West Lafayette, Indiana, USA; Jeremy Foote, jdfoote@purdue.edu, Purdue University, West Lafayette, Indiana, USA.

Manuscript submitted to ACM

they function within institutional frameworks where data collection is a fundamental component [28]. This institutional storage of personal information raises concerns beyond corporate use, as breaches, unauthorized access, or data misuse could expose users' most intimate disclosures [2, 11]. As AI companions become increasingly sophisticated and widely adopted [5], teens are turning to them for emotional support, often treating these systems as trusted confidants. However, the nature of privacy in AI companion relationships differs fundamentally from human relationships, creating risks if teens misinterpret what is ultimately a corporate data relationship as a personal, confidential exchange.

The concept of dimensional privacy explains that people apply different privacy strategies when sharing with individuals (horizontal privacy) versus institutions (vertical privacy) [19]. AI companions disrupt this distinction, leading teens to disclose personal information using interpersonal heuristics, often without fully grasping how their data is processed and stored. This complexity also presents challenges for parents and educators, who may not always have the resources or understanding needed to help teens navigate these emerging privacy risks.

To examine this issue, we apply Communication Privacy Management (CPM) theory and dimensional privacy to explore how teens navigate these privacy dimensions in their interactions with AI companions. Understanding these dynamics has direct implications for platform design, policy, and education, ensuring AI systems are ethically designed while equipping parents and educators with better strategies to support teens in managing their privacy.

2 Understanding Privacy in Teen-AI Communication

The emergence of AI companions has transformed how teenagers navigate privacy in digital spaces, introducing challenges beyond traditional interpersonal boundaries [12, 17, 27]. Understanding how teens manage privacy in these interactions requires Communication Privacy Management (CPM) theory and dimensional privacy, which together highlight the complexities of balancing interpersonal and institutional privacy boundaries in AI companion relationships.

2.1 Privacy Through Communication Management: A Theoretical Lens

For adolescents, managing privacy boundaries is a essential development task that become increasingly complex in digital spaces [18, 33]. Communication Privacy Management (CPM) theory provides valuable insights into how teens navigate privacy decision with AI companions [23, 24]. Through dialectical interaction and communication, people develop their own rules to manage their privacy boundaries [24]. The theory centers on three fundamental elements: privacy ownership, privacy control, and privacy turbulence [24, 25].

Privacy ownership refers to the belief that individuals have the right to own and control their private information. When people share private information, they create co-ownership of that information, establishing collective responsibility for its management [24]. In the context of teen-AI interactions, this concept becomes complicated because teens may perceive ownership of their disclosures, while unknowingly creating co-ownership with corporate entities [11].

Privacy control involves the development and use of privacy rules to regulate information flow. These rules are influenced by cultural values, motivations, risk-benefit assessments, and contextual factors [25]. Teens develop these rules to manage their disclosures in the contexts of interpersonal and mediated relationships [8, 24], but AI companions complicate this process by creating an illusion of controlled, dyadic communication while actually operating within broader institutional frameworks.

Privacy turbulence occurs when privacy rules are violated or boundaries become unclear, leading to boundary management problems [24]. This is especially relevant with AI companions, as teens may not only face interpersonal boundary violations but also institutional data practices that they did not fully understand or consent to [28]. Teenagers may believe they control their self-disclosures when interacting with AI companions, yet their personal information is

often co-owned by the companies operating these systems. Unlike traditional interpersonal relationships, where privacy rules evolve through mutual understanding, AI-driven interactions lack reciprocity—companies unilaterally determine how user data is collected, stored, and potentially monetized [11]. The discrepancy between perceived privacy and actual data ownership creates privacy turbulence, leaving users struggling to maintain control over their sensitive information.

2.2 Dimensional Privacy: The Dilemma of Privacy Management

While CPM theory explains how individuals develop and manage privacy rules, it does not fully account for the structural forces shaping privacy decisions in human-AI interactions. To address this gap, the concept of dimensional privacy distinguishes between horizontal (peer-to-peer) and vertical (institutional) privacy concerns, providing a valuable framework for examining teens' AI companion use[26]. Masur [19] applies this framework to understand situational privacy and self-disclosure in contemporary digital environments. The horizontal dimension of privacy—managing information flow between interpersonal relationships—is familiar territory for adolescents who have grown up negotiating social media boundaries through socializing in mediated context [1]. However, the vertical dimension, which involves institutional data practices, often remains opaque and poorly understood by young users [10].

AI companions blur the lines between these two dimensions, particularly for teenagers who may not yet have a fully developed understanding of digital privacy risks. On a horizontal level, teens often treat AI companions as personal confidants, sharing their innermost thoughts, frustrations, and personal struggles. These exchanges mimic private conversations with a trusted friend, reinforcing the perception that the AI provides a safe and intimate space [16]. However, teens may fail to recognize that, unlike human friends, their disclosures are not protected by the same ethical or social norms governing human relationships.

At the vertical level, AI companions exist within a corporate structure where user interactions are systematically collected, stored, and potentially monetized. This institutional data collection creates a significant risk, especially for vulnerable teenage users who may turn to AI companions in times of distress. Some adolescents, particularly those struggling with loneliness, social anxiety, or mental health challenges, may disclose deeply personal and sensitive information to AI systems, believing these conversations to be private. However, AI systems do not operate under confidentiality agreements like human therapists or trusted adults or friends.

Moreover, AI companionship presents an unprecedented intersection of horizontal and vertical privacy risks, making teenage users particularly susceptible to breaches and misuse. While teens may expect a relationship similar to a friendship, the reality is that they are interacting with a system designed to collect and analyze their behavior. Unlike human friendships, where privacy boundaries evolve based on trust and mutual respect, AI interactions operate within a regulatory and ethical gray area where teens have little control over how their data is stored and used.

This intricate dimensional intersection is especially problematic because teens' developmental stage makes them particularly vulnerable to privacy risks [13]. During adolescence, young people are naturally inclined toward intimate disclosure as part of identity formation and relationship building [6, 30]. AI companions, with their seemingly private, judgment-free interfaces, tap into this developmental need [12]. However, while teens may expertly manage horizontal privacy concerns, they often lack the awareness or tools to address vertical privacy implications.

3 Discussion and Implication

3.1 The Developmental Stakes

The intersection of CPM theory and dimensional privacy in teens' AI companion use raises significant developmental concerns. During adolescence, young people need safe spaces to explore identity, process emotions, and practice vulnerability [20]. AI companions appear to offer this safety, but the reality of institutional data collection may compromise these crucial developmental experiences [9, 14]. When teens' personal disclosures become corporate assets, it raises questions about the long-term impact on their privacy development and relationship formation.

The developmental implications become even more significant when considering that adolescence is a critical period for learning to navigate relationships and establish personal boundaries[18]. During this stage, teens are actively developing their capacity for intimate relationships while simultaneously learning to protect their privacy [33]. AI companions, which blur the lines between genuine intimacy and datafied interactions, may interfere with this natural development process. When teens practice vulnerability with AI systems that appear to offer unconditional acceptance but actually operate within commercial frameworks, they may develop skewed expectations about privacy, trust, and reciprocity in relationships [21].

Furthermore, the AI companion experience may impact teens' developing ability to assess risk in digital environments. As adolescents engage with these platforms during their formative years, their experiences shape their future approaches to privacy management and digital literacy [15, 21]. If teens become habituated to sharing deeply personal information with AI systems without fully understanding the horizontal and vertical privacy concerns, they may develop compromised mental models of privacy that persist into adulthood. This could make them more vulnerable to privacy risks across various digital contexts and potentially impact their ability to form healthy boundaries in both online and offline relationships [33].

3.2 Implications for Design and Stakeholder Engagement

AI companions present a new paradox between usage and privacy: they encourage deep disclosures but simultaneously expose users to institutional data collection [2, 12]. Future designs should enhance privacy transparency while preserving their therapeutic benefits. One approach is integrating "transparency moments"—periodic, conversational reminders that clarify how user data is stored and used [13, 14, 31]. To support privacy literacy, AI interfaces could allow users to customize data-sharing preferences and introduce subtle nudges that encourage reflection before sharing sensitive information. Research should further investigate how teens develop and modify privacy rules, particularly as AI systems evolve [3].

Beyond interface design, addressing privacy concerns in teen-AI interactions requires engaging multiple stakeholders in teens' digital lives. Parents and educators, who traditionally guide technology use, may not always be aware of the unique privacy risks associated with AI companions [7]. While managing overall technology use remains important, it is equally essential to understand how AI companions collect and store sensitive disclosures. Educational programs should equip parents and teachers with knowledge of both the benefits and risks of AI companions, helping teens develop effective privacy management strategies alongside broader digital well-being efforts [29]. Additionally, teens facing limited access to mental health resources may turn to AI companions for emotional support, increasing their risk of over-disclosure and unintended data exposure [22]. As these systems become more prevalent, privacy protections and literacy efforts must ensure teens can engage safely while maintaining autonomy and emotional well-being.

References

[1] Denise E Agosto and June Abbas. 2017. "Don't be dumb—that's the rule I try to live by": A closer look at older teens' online privacy and safety attitudes. *New Media & Society* 19, 3 (2017), 347–365.

[2] Pieter Arntz. 2024. AI girlfriend site breached, user fantasies stolen. https://www.malwarebytes.com/blog/news/2024/10/ai-girlfriend-site-breached-user-fantasies-stolen February, 2025.

[3] Rahime Belen Saglam, Jason RC Nurse, and Duncan Hodges. 2021. Privacy concerns in chatbot interactions: When to trust and when to worry. In *HCI International 2021-Posters: 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part II 23*. Springer, 391–399.

[4] Elin A Björling, Emma Rose, Andrew Davidson, Rachel Ren, and Dorothy Wong. 2020. Can we keep him forever? Teens' engagement and desire for emotional connection with a social robot. *International Journal of Social Robotics* 12, 1 (2020), 65–77.

[5] Sandy Carter. 2024. When Humans Swipe Right For An AI Companion. https://www.forbes.com/sites/digital-assets/2024/10/17/when-humans-swipe-right-for-an-ai-companion/ February, 2025.

[6] Meghan A Costello, Corey Pettit, Amanda F Hellwig, Gabrielle L Hunt, Natasha A Bailey, and Joseph P Allen. 2024. Adolescent social learning within supportive friendships: Self-disclosure and relationship quality from adolescence to adulthood. *Journal of Research on Adolescence* (2024).

[7] Lorrie Faith Cranor, Adam L Durity, Abigail Marsh, and Blase Ur. 2014. {Parents'} and {Teens'} Perspectives on Privacy In a {Technology-Filled} World. In *10th Symposium On Usable Privacy and Security (SOUPS 2014)*. 19–35.

[8] Ralf De Wolf. 2020. Contextualizing how teens manage personal and interpersonal privacy on social media. *New media & society* 22, 6 (2020), 1058–1075.

[9] Ziwei Gao. 2024. Why Does AI Companionship Go Wrong? *The International Review of Information Ethics* 34, 1 (2024).

[10] Eszter Hargittai and Alice Marwick. 2016. What can I really do?" Explaining the privacy paradox with online apathy. *International journal of communication* 10 (2016), 21.

[11] Martin Hasal, Jana Nowaková, Khalifa Ahmed Saghair, Hussam Abdulla, Václav Snášel, and Lidia Ogiela. 2021. Chatbots: Security, privacy, data protection, and social aspects. *Concurrency and Computation: Practice and Experience* 33, 19 (2021), e6426.

[12] Jodi Heckel. 2024. Researchers examine teens' use of generative AI, safety concerns. https://techxplore.com/news/2024-12-teens-generative-ai-safety.html February, 2025.

[13] Chia Min Ho. 2023. Facilitating Student Counseling Through the Chatbot. In *International Conference on Human-Computer Interaction*. Springer, 51–59.

[14] Carolin Ischen, Theo Araujo, Hilde Voorveld, Guda van Noort, and Edith Smit. 2020. Privacy concerns in chatbot interactions. In *Chatbot Research and Design: Third International Workshop, CONVERSATIONS 2019, Amsterdam, The Netherlands, November 19–20, 2019, Revised Selected Papers 3*. Springer, 34–48.

[15] Isabell Koinig. 2020. " I'm not a kid anymore! Towards a teen-centric approach of online privacy management. *Medienjournal-Zeitschrift für Medien-und Kommunkiationsforschung* 44, 1 (2020), 41–54.

[16] Frank Landymore. 2024. Lonely teens are making "friends" with AIs. Teens aren't just using chatbots to do their homework anymore. https://futurism.com/the-byte/lonely-teens-friends-with-ai February, 2025.

[17] Sage Lazzaro. 2024. Teens are using AI, but are worried about what it means for their futures. https://fortune.com/2024/11/14/ai-risk-bigger-than-climate-change-inequality-teen-survey-finds-eye-on-ai/ February, 2025.

[18] Roger JR Levesque. 2016. *Adolescence, privacy, and the law: A developmental science perspective*. Oxford University Press.

[19] Philipp K Masur. 2018. *Situational privacy and self-disclosure: Communication processes in online environments*. Springer.

[20] Debbie Noble-Carr and Elise Woodman. 2018. Considering identity and meaning constructions for vulnerable young people. *Journal of Adolescent Research* 33, 6 (2018), 672–698.

[21] Jussi Okkonen and Sirkku Kotilainen. 2019. Minors and Artificial Intelligence–implications to media literacy. In *Information Technology and Systems: Proceedings of ICITS 2019*. Springer, 881–890.

[22] Jinkyung Park, Vivek Singh, and Pamela Wisniewski. 2023. Supporting youth mental and sexual health information seeking in the era of artificial intelligence (ai) based conversational agents: Current landscape and future directions. *Available at SSRN 4601555* (2023).

[23] S Petronio. 2002. Boundaries of privacy: Dialectics of disclosure. *State Univ of New York Pr* (2002).

[24] Sandra Petronio. 2010. Communication privacy management theory: What do we know about family privacy regulation? *Journal of family theory & review* 2, 3 (2010), 175–196.

[25] Sandra Petronio and Jeffrey T Child. 2020. Conceptualization and operationalization: Utility of communication privacy management theory. *Current opinion in psychology* 31 (2020), 76–82.

[26] Kate Raynes-Goldie. 2010. Aliases, creeping, and wall cleaning: Understanding privacy in the age of Facebook. *First Monday* (2010).

[27] Rebecca Ruiz. 2024. Teens are talking to AI companions, whether it's safe or not. https://mashable.com/article/ai-companion-teens-safety February, 2025.

[28] Marita Skjuve, Asbjørn Følstad, Knut Inge Fostervold, and Petter Bae Brandtzaeg. 2021. My chatbot companion-a study of human-chatbot relationships. *International Journal of Human-Computer Studies* 149 (2021), 102601.

[29] NH Vasoya. 2023. The role of parents and educators in managing the risks of artificial intelligence. *Asian Journal of Education and Social Studies* 41, 4 (2023), 1–5.

[30] Nandita Vijayakumar and Jennifer H Pfeifer. 2020. Self-disclosure during adolescence: Exploring the means, targets, and types of personal exchanges. *Current Opinion in Psychology* 31 (2020), 135–140.

[31] Jonathan Vitale, Meg Tonkin, Sarita Herse, Suman Ojha, Jesse Clark, Mary-Anne Williams, Xun Wang, and William Judge. 2018. Be more transparent and users will like you: A robot privacy and user experience design experiment. In *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*. 379–387.

[32] Joel Wester, Henning Pohl, Simo Hosio, and Niels van Berkel. 2024. " This Chatbot Would Never...": Perceived Moral Agency of Mental Health Chatbots. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–28.

[33] Pamela J Wisniewski, Jessica Vitak, and Heidi Hartikainen. 2022. Privacy in adolescence. In *Modern socio-technical perspectives on privacy*. Springer International Publishing Cham, 315–336.

Received 13 February 2025