Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

She is fake, but it still feels like betrayal: How an online community negotiates the moral boundaries of AI intimacy within marriage

Anurag Shekhar

shekha@unisa.ac.za

University of South Africa

Thulile Ngonyama-Ndou

University of South Africa

Article

Keywords: AI companionship, moral boundary work, emotional labour, digital intimacy, emotional infidelity, Reddit, mixed methods, reflexive thematic analysis

Posted Date: September 12th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10614183/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

AI companion applications have moved from niche novelty to everyday intimate life, yet their entry into existing marriages remains poorly understood from a sociological perspective. Existing research focuses primarily on the psychological experience of AI users, leaving a significant gap in understanding how broader communities collectively negotiate the moral status of AI intimacy when it conflicts with established expectations of marriage, fidelity, and relational reciprocity. This study analyses Reddit discourse from a high-engagement thread in which a wife describes her husband's two-month attachment to an AI companion named Nova and his request to open their marriage. Using a convergent mixed-methods design, VADER sentiment analysis, NRC emotion analysis, NMF topic modelling, keyword frequency analysis, and engagement analysis were applied to 1,246 comments, and Braun and Clarke's reflexive thematic analysis was conducted on the top 600 comments ranked by upvote score, supplemented by a purposive sample of dissenting responses. The analysis is guided by an integrative multi-level theoretical framework drawing on Bauman's account of liquid love, Hochschild's concept of emotional labour, Lamont and Molnár's framework of symbolic boundary work, and Goffman's account of stigma management. Four themes are identified: the real versus fake distinction as moral boundary work, frictionless intimacy and the erosion of relational reciprocity, the AI as digital mirror and the erasure of otherness, and community sanction through therapeutic pathologisation and legal-familial enforcement. A cross-cutting finding shows that the AI's ontological unreality functioned not as a mitigating factor but as a moral catalyst, intensifying rather than reducing the community's condemnation of the husband's conduct. The study argues that the thread functioned as a site of collective moral construction in which commenters actively built, contested, and enforced the boundaries of legitimate intimate partnership in the presence of AI. The findings extend symbolic boundary work theory into the domain of digital intimacy and contribute to emerging sociological debates on AI companionship, relational reciprocity, and the renegotiation of marital norms.

1. Introduction

In July 2025, a post appeared on the Reddit forum r*/AmIOverreacting*. A woman described how her husband had spent two months forming a romantic attachment to an AI companion application. The AI, which he had named Nova, understood him, never interrupted him, and cared about his children. He was not asking his wife to accept a human affair. He was asking her to open their marriage so that his relationship with Nova could be treated as legitimate. Within days, the thread had attracted more than 1,200 responses. This paper examines how that community responded and what those responses reveal about emerging moral boundaries around human intimacy in the presence of AI.

The scenario is no longer merely speculative. AI companion applications have developed from simple virtual assistants into generative systems that can provide persistent emotional support, romantic engagement, and a sense of being understood (Chaturvedi et al. 2023; Gur and Maaravi 2025). Many are designed to remain available, responsive, and affirming while placing few demands on the user. Research suggests that interactions with AI companions can reduce reported loneliness and fulfil needs

for attention, validation, and social connection (Machia et al. 2024; De Freitas et al. 2025). Studies have also documented users developing romantic attachments to chatbots and, in some cases, prioritising these bonds over relationships with human partners (Pataranutaporn et al. 2025; Shu et al. 2026).

These attachments emerge within wider changes in intimate life. Bauman (2013) argues that contemporary relationships are shaped by a tension between the desire for connection and the fear of vulnerability and lasting commitment. Finkel et al (2014) similarly show that modern marriage is expected to meet increasingly demanding needs for emotional support, identity, and personal growth, even as couples may have fewer wider social resources on which to rely. An AI companion that is readily available, highly validating, and asks little in return can therefore be understood as one response to these changing relational conditions rather than simply as an individual anomaly.

Existing research has mainly examined why users form attachments to AI companions (Pan and Mou 2024; Yang and Oshio 2025; Shu et al. 2026), whether these systems reduce loneliness (Machia et al. 2024; De Freitas et al. 2025), and what benefits or harms arise within the user-AI relationship (Zimmerman et al. 2024; Zhang et al. 2025; Malfacini 2025). Much less is known about how such attachments are judged by people outside that relationship, including spouses, families, and wider communities (Kirk et al. 2025). AI intimacy is therefore still often treated as a private psychological bond rather than as a socially contested relationship whose legitimacy is negotiated through existing norms of marriage, fidelity, gender, and reciprocity (Muldoon and Parke 2025; Pataranutaporn et al. 2025). This leaves an important gap in AI social science: how do communities classify and regulate AI intimacy when it conflicts with established human obligations?

The Reddit thread provides a naturalistic setting in which this process became visible (Chu et al. 2025). Commenters drew distinctions between real and fake relationships, debated whether the attachment constituted cheating, prescribed therapy or divorce, and used humour and ridicule to mark the conduct as unacceptable. Rather than asking whether AI companionship is inherently beneficial or harmful, this paper examines how AI intimacy becomes collectively interpreted, morally classified, and socially regulated. In doing so, it treats AI not as a technology with a fixed social meaning, but as a sociotechnical object whose legitimacy is produced through public judgement and existing cultural expectations (Jobin et al. 2019; Ladak et al. 2024). Using a convergent mixed-methods design, this study combines computational text analysis with reflexive thematic analysis to examine how online communities negotiate the boundaries of AI-mediated intimacy.

Four research questions guide the analysis. First, how do commenters construct the moral boundaries of cheating and betrayal when AI companionship enters a marriage? Second, how do they negotiate the ontological status of AI when defining legitimate intimate relationships? Third, how do they construct expectations of reciprocity and relational obligation when evaluating the husband's behaviour? Fourth, how does the community exercise collective sanction in response to AI-mediated intimacy?

2. Literature Review

2.1 AI companionship and human-AI relationships

The rapid development of AI companion technologies has introduced new forms of human-machine interaction that increasingly resemble social and intimate relationships. Unlike earlier digital assistants designed primarily for information retrieval, contemporary AI companions use conversational memory, personalisation, and anthropomorphic cues to simulate emotional availability and social presence (Brandtzaeg et al. 2022; Kirk et al. 2025). Research on platforms such as Replika shows that some users interpret these interactions as meaningful relationships, even while recognising that the companion is artificial (Laestadius et al. 2024; Pan and Mou 2024).

This emerging form of intimacy has generated debate about whether AI companions function as supportive tools or substitutes for human connection. Some research highlights their potential to provide companionship, reduce loneliness, and offer a low-risk environment for emotional expression (Machia et al. 2024; De Freitas et al. 2025). However, scholars have also raised concerns that highly personalised and emotionally responsive systems may encourage dependency, reinforce one-sided relationships, or alter expectations of human intimacy (Zimmerman et al. 2024; Muldoon and Parke 2025).

A central tension in this literature concerns the difference between subjective experience and social recognition. Users may experience AI relationships as emotionally significant, while wider communities may question their legitimacy because they lack human reciprocity, vulnerability, and mutual obligation (Pan and Mou 2024; Willoughby et al. 2025). Existing research has primarily examined why individuals form attachments to AI companions. Less attention has been given to how broader social groups evaluate these relationships and negotiate their implications for existing human relationships. This study addresses this gap by examining how an online community collectively evaluates AI intimacy when it enters the domain of marriage.

2.2 Marriage, Infidelity, and Moral Expectations

Contemporary marriage carries increasingly complex expectations. Cherlin (2004) argues that marriage has undergone a process of deinstitutionalisation, with weakening social norms that previously defined and regulated marital roles. Finkel et al. (2014) describe a “suffocation” model of marriage in which partners increasingly expect the relationship to provide emotional fulfilment, self-development, and personal growth. At the same time, wider social networks that previously supported these needs have declined. Bauman (2013) situates these developments within liquid modernity, where individuals continue to seek connection while attempting to minimise vulnerability and obligation. Within this context, AI companionship can be understood as one contemporary expression of the search for emotionally satisfying but lower-risk forms of connection.

Infidelity is generally understood as a violation of explicit or implicit expectations regarding emotional or sexual exclusivity within a relationship (Blow and Hartnett 2005). Emotional infidelity involves the development of a significant emotional bond outside the primary relationship, often involving secrecy

and the diversion of emotional attention away from a partner (Glass and Wright 1992; Guitar et al. 2017). Such experiences can be deeply harmful because they are interpreted as a withdrawal of emotional priority and commitment. Research has linked infidelity with anxiety, depression, and lasting damage to trust (Rokach and Chan 2023), and many partners experience emotional betrayal as comparable to physical betrayal (Whitty 2003).

Digital technologies have expanded the boundaries of infidelity. Emotional and relational boundary violations can now occur through social media, messaging platforms, and other digitally mediated spaces (Hertlein and Piercy 2006). AI companionship introduces a further unresolved question: when a person forms a deep emotional attachment to an AI within a marriage, how should that relationship be classified? Vossler and Moller (2024) argue that AI companions create relational spaces that do not fit existing definitions of fidelity and betrayal. Prochazka and Brooks (2024) similarly show that AI romantic rivals can generate strong emotional reactions. Some interpret the diversion of emotional resources towards AI as a genuine relational threat, while others question whether infidelity can exist without another human subject. This unresolved tension forms the central empirical focus of this study.

2.3 Reddit, online communities, and moral sense-making

Online platforms provide important spaces where emerging social and technological practices are publicly debated and evaluated. Reddit, in particular, functions as an advice-oriented environment where users respond to personal narratives, exchange interpretations, and collectively negotiate norms through discussion and voting mechanisms (Weld et al. 2024; Reagle 2025).

Previous research has shown that online communities are not simply spaces for information exchange. They also operate as sites of collective meaning-making, where users establish shared expectations about acceptable behaviour and respond to perceived violations of social norms (Weld et al. 2024). This makes Reddit particularly relevant for studying emerging moral questions surrounding AI companionship, where formal social rules remain unsettled.

Although studies have examined how AI companion users understand and justify their own relationships with chatbots (Brandtzaeg et al. 2022; Pan and Mou 2024), less is known about how external audiences evaluate such relationships (Pataranutaporn et al. 2025). Studying public reactions therefore provides insight into how societies negotiate the boundaries between technological innovation and established expectations of intimacy, fidelity, and responsibility (Kirk et al. 2025; Pataranutaporn et al. 2025).

3. Theoretical Framework

This study develops an integrative multi-level framework linking structural conditions, relational expectations, collective meaning-making, and interactional outcomes.

At the macro-structural level, Bauman's (2013) concept of liquid love provides the context for understanding why AI intimacy may become appealing. Bauman argues that late-modern relationships are shaped by a tension between the desire for connection and the desire to avoid vulnerability and

constraint. Within this context, AI companionship can be understood as one contemporary expression of the search for connection that reduces many of the uncertainties associated with human relationships.

At the micro-relational level, Hochschild's (1983) concept of emotional labour explains why AI intimacy may be interpreted as threatening within marriage. Emotional labour involves managing feelings and sustaining relationships through ongoing emotional effort. Intimate relationships depend on reciprocal forms of care, attention, and emotional responsiveness. When one partner redirects emotional investment towards an AI companion that requires no comparable emotional exchange, commenters may interpret this as a breakdown of relational reciprocity. Later scholarship further highlights that emotional labour remains unevenly distributed within intimate relationships, with gendered consequences (Steinberg and Figart 1999).

At the meso-discursive level, Lamont and Molnár's (2002) framework of symbolic boundary work explains how communities classify unfamiliar social phenomena. Symbolic boundaries allow groups to distinguish between categories such as real and fake, legitimate and illegitimate, acceptable and unacceptable. In this study, commenters engage in boundary work around three related questions: whether AI can represent genuine intimacy, whether AI attachment constitutes betrayal, and what obligations marriage requires when technology enters intimate life.

At the interactional level, Goffman's (1963) account of stigma explains how moral classifications become social responses. Communities do not simply identify norm violations; they also mark and respond to those violations. In this study, therapeutic recommendations, divorce advice, ridicule, and disbelief function as forms of collective sanction through which commenters reinforce expectations around acceptable intimate behaviour.

Together, these frameworks trace an analytical pathway from the social conditions that make AI intimacy possible, to the relational tensions it creates, the moral categories communities use to interpret it, and the sanctions through which those interpretations are enforced (Fig. 1).

4. Methodology

4.1 Research Design

This study used a convergent mixed-methods design in which computational and qualitative analyses were conducted in parallel and integrated during interpretation (Creswell and Plano Clark 2017). The design was selected because the research questions required both corpus-level description and interpretive depth. The computational strand examined broad emotional, thematic, and lexical patterns across the full dataset of 1,246 Reddit comments. The qualitative strand examined how commenters constructed moral distinctions, relational expectations, and judgements around AI intimacy within specific exchanges. Quantitative analysis established patterns across the corpus, while qualitative analysis examined how those patterns were socially produced and interpreted. Together, the two strands

allowed the community response to AI intimacy to be analysed both as a large-scale discursive pattern and as a socially negotiated moral process.

4.2 Data Collection

Data were drawn from a single thread posted to the public subreddit *r/AmIOverreacting* in July 2025. At the time of data collection, the subreddit reported approximately 3.3 million weekly visitors, reflecting its substantial activity as an online advice community. The thread was initiated by a woman describing her husband's emotional attachment to an AI companion named Nova and his request to open their marriage. The post received approximately 2,300 upvotes and more than 1,300 comments at the time of capture. The dataset contained comments from 887 unique author accounts. After removing deleted comments and applying a minimum length threshold to exclude non-substantive responses, the final quantitative corpus comprised 1,246 comments.

Methodologically, this study adopts a critical case study design (Flyvbjerg 2006), treating this specific high-engagement thread not as a representative sample of general public opinion, but as a diagnostic *sociotechnical breakdown* or *critical incident* (Pinch and Bijker 1984). In the sociology of technology, latent social norms and moral boundaries are often invisible when systems operate smoothly; they only become explicit and subject to public negotiation during moments of rupture, such as when a novel, contested technology (here, an emotionally responsive AI companion) suddenly disrupts a highly institutionalised social expectation (marital fidelity and relational reciprocity). By selecting a single, exceptionally high-engagement thread (1,246 comments) characterized by intense discursive struggle, we capture a concentrated *focused diagnostic window* where participants are forced to actively construct and police the boundaries of legitimate intimacy in real-time. The analytic value therefore lies not in statistical representativeness, but in the thick, naturalistic elicitation of the cultural codes and interpretive repertories that social actors mobilize when existing normative frameworks are suddenly destabilized

For the qualitative analysis, a purposive high-engagement subsample was drawn from the top 600 comments ranked by upvote score. This approach reflects an important feature of Reddit as a research environment: upvotes make community endorsement visible, allowing researchers to examine not only what users express but also which contributions receive stronger collective amplification (Leavitt and Robinson 2017). The threshold was selected to balance analytical depth with preservation of variation in community responses. A supplementary purposive sample of dissenting and ambivalent responses was added to ensure that minority positions were represented. The qualitative dataset retained the threaded structure of Reddit, with parent comments and nested replies labelled separately. Usernames and all directly identifying information were excluded from the analytic dataset.

4.3 Computational Analysis

Five computational techniques were applied to the quantitative corpus.

Sentiment analysis used the Valence Aware Dictionary and Sentiment Reasoner (VADER), a lexicon-based approach developed for informal social media text containing features such as slang, emphasis, and non-standard punctuation (Hutto and Gilbert 2014). Each comment received a compound score ranging from − 1 to + 1 and was classified as positive (≥ 0.05), neutral (between − 0.05 and 0.05), or negative (≤ -0.05).

Emotion analysis used the NRC Emotion Lexicon (Mohammad and Turney 2013), which maps words onto eight discrete emotions and two valence dimensions. This provided a more differentiated descriptive account of emotional expression than sentiment polarity alone. These approaches were selected because of their transparency and suitability for exploratory analysis of informal social media discourse.

Topic modelling was conducted using non-negative matrix factorisation (NMF). NMF decomposes a document-term matrix into interpretable parts-based representations and is particularly suitable for identifying recurring lexical patterns in short text datasets (Lee and Seung 1999). Six topics were retained following iterative testing of topic solutions and assessment of interpretability. These topics were treated as exploratory lexical clusters rather than as equivalent to qualitative themes.

Engagement analysis examined Reddit scores in relation to sentiment, comment length, and analytically derived interpretive frames (e.g., therapy, divorce, authenticity, real versus fake, reciprocity, and betrayal). These frames were treated as descriptive indicators of recurring rhetorical positions rather than substitutes for qualitative interpretation. The analysis used comment-level metadata, including Reddit score and word count, from the master dataset.

4.4 Reflexive Thematic Analysis

The qualitative subsample was analysed using Braun and Clarke's (2019) reflexive thematic analysis. This approach was selected because it is suited to examining patterns of shared meaning within social and cultural discourse. Analysis followed Braun and Clarke's six phases: familiarisation, coding, initial theme development, theme review, theme definition and naming, and report writing.

Comments were imported into ATLAS.ti for analysis. Coding was sensitised by the theoretical framework, particularly concepts related to symbolic boundaries, emotional labour, and moral judgement, while allowing interpretations to develop from engagement with the dataset. Attention was given to how commenters defined acceptable and unacceptable intimacy, constructed relational obligations, and interpreted the appeal and risks of AI companionship. Codes were refined through multiple rounds of analysis, and themes were developed around patterns of shared meaning rather than frequency alone. Reflexive memoing was used throughout to document analytic decisions.

Consistent with Braun and Clarke's approach to reflexive thematic analysis, inter-rater reliability was not used as a measure of quality. Instead, rigour was supported through sustained engagement with the dataset, iterative coding, transparent analytical decisions, and grounding themes in multiple extracts across the corpus.

4.5 Mixed-Methods Integration

The two analytical strands were integrated through systematic comparison and joint interpretation at the meta-inference stage (Creswell and Plano Clark 2017). A connected integration strategy was used, whereby qualitative analysis expanded and contextualised the computational findings. Integration focused on areas of convergence between strands, instances where one strand extended the other, and cases where qualitative analysis provided explanations for corpus-level patterns.

5. Results

5.1 Computational Findings

5.1.1 Sentiment distribution and Emotional profile

Instead of representing standard sentiment polarities or straightforward approval of AI intimacy, the quantitative corpus skeleton revealed a highly polarized interactional landscape. VADER sentiment and NRC emotion metrics showed that positive valence and "trust" markers did not signal acceptance of the husband's behavior; rather, they were discursively bound to supportive, advice-oriented expressions of solidarity and care directed toward the original poster (Table 1, Table 2). Conversely, the significant presence of negative valence, anger, and disgust (Table 2) marked the corpus as a site of moral boundary enforcement, where commenters expressed collective rejection of the husband's actions

Table 1
Sentiment distribution of Reddit comments

Sentiment	n	%
Positive	567	45.51
Negative	437	35.07
Neutral	242	19.42
Total	1,246	100.00

Table 2
NRC emotion distribution of Reddit comments

Emotion	Count	% of hits	Comments n	Comments %
Positive	3,306	21.27	814	65.33
Trust	2,335	15.02	745	59.79
Negative	2,224	14.30	755	60.59
Anticipation	1,428	9.19	575	46.15
Joy	1,400	9.00	536	43.02
Fear	1,168	7.51	561	45.02
Sadness	1,142	7.35	545	43.74
Anger	1,025	6.59	509	40.85
Disgust	797	5.13	433	34.75
Surprise	722	4.64	427	34.27

Note

Percentage of hits refers to each category's share of total NRC lexical matches across the corpus. Comment coverage refers to the percentage of comments containing at least one lexical match for that category. Positive and negative are valence dimensions within the NRC framework and are reported separately from the eight discrete emotions.

5.1.2 Exploratory lexical clustering

We used an exploratory NMF topic model to identify recurring semantic patterns across the corpus, retaining a six-topic solution as the most interpretable baseline (Table 3). These clusters represent mathematical patterns of word co-occurrence rather than qualitative themes, serving as corpus-level contextual evidence to complement the reflexive thematic analysis.

This moral negotiation was structurally organized around pre-existing cultural imaginaries. Rather than interpreting artificial intimacy in a cultural vacuum, commenters heavily relied on NMF Topic 6 (Cultural Recognition Frame, 5.81%), which was characterized by high-weight terms such as *'Black Mirror'*, *'episode'*, and *'written'* (Table 3). This lexical cluster indicates that the community turned to established dystopian narratives and science-fiction templates to comprehend and classify the husband's attachment. These shared cultural scripts did not merely serve as passive entertainment references; they functioned as active interpretive structures that provided the literal vocabulary for the *Ontological Paradox* detailed in Theme 1. By framing the scenario as "literally a *Black Mirror* episode," commenters

successfully stripped the husband's actions of relational legitimacy, shifting the discourse from a private marital disagreement to a recognized cultural transgression of human-machine boundaries

Together, these topic clusters organize the discourse around four core concerns: relational meaning, behavioral evaluation, family protection, and cultural framing. These macro-level quantitative patterns systematically complement the deep interpretive themes developed through our reflexive thematic analysis.

Table 3
Exploratory NMF topic clusters in the Reddit corpus

Topic	n	%	Interpretive label	High-weight terms
1	437	39.05	AI attachment and relational meaning	bot; feel; love; men; sex
2	134	11.97	Ontological evaluation	real; real life; human; real human
3	182	16.26	Therapeutic framing	needs; help; therapy; therapist; psychiatrist
4	217	19.39	Legal and familial consequences	kids; leave; divorce; custody; lawyer
5	84	7.51	Authenticity scepticism	fake; post; story; creative
6	65	5.81	Cultural recognition frame	Black Mirror; episode; sounds; written

5.1.3 Engagement patterns

Reddit engagement patterns were highly skewed. Across sentiment categories and interpretive frames, the median score was 1, indicating that most comments received limited endorsement. Mean scores therefore reflect the visibility of highly endorsed comments rather than general community agreement. No meaningful relationship was observed between comment length, sentiment score, and engagement (all $r < .10$).

As shown in Table 4, comments using therapy and divorce or leave frames received higher mean scores ($M = 25.08$ and $M = 21.74$ respectively) than comments focused primarily on cheating or betrayal ($M = 10.57$). This suggests that highly endorsed responses were not limited to moral condemnation but also emphasised intervention, protection, and practical action. The authenticity frame received the highest mean score ($M = 53.31$), although this should be interpreted cautiously due to its small sample size ($n = 48$). Overall, engagement patterns support the qualitative finding that the community translated moral judgement into action-oriented responses.

Table 4
Engagement by interpretive frame in the Reddit AI companionship and marriage corpus

Frame	n	% of comments	Mean score	Median score
Authenticity frame	48	3.85	53.31	1
Therapy frame	260	20.87	25.08	1
Divorce / leave frame	179	14.37	21.74	1
Real versus fake frame	353	28.33	13.11	1
Kids / family frame	235	18.86	12.98	1
Need / reciprocity frame	315	25.28	12.95	1
Cheating / betrayal frame	95	7.62	10.57	1

Note: Reddit scores are highly skewed. Median scores were 1 across all frames, indicating that mean values were influenced by a small number of highly endorsed comments.

5.2 Qualitative Findings: Reflexive Thematic Analysis

The reflexive thematic analysis showed that commenters did not respond to the case in a single register. They moved between ontological judgement, moral evaluation, relational reasoning, and practical advice. Across the dataset, four interrelated themes captured the dominant patterns of meaning: the unstable boundary between fake partners and real betrayal, frictionless intimacy and the failure of reciprocity, the AI as mirror and self-validating attachment, and community sanction through therapy, divorce, ridicule, and protective advice.

5.2.1 Theme 1: "She Is Fake, but It Still Feels Like Betrayal" — Moral boundary work around AI infidelity

Commenters did not agree on what Nova was, and this disagreement became the primary site of symbolic boundary work in the thread. The real versus fake distinction was not a settled description of the technology. Instead, it became a rhetorical resource through which commenters reached different moral conclusions about the husband's conduct. How the AI was classified shaped whether the attachment was understood as a relational violation or dismissed as something outside existing categories of infidelity.

The dominant position was ontological dismissal. Commenters insisted that Nova was fake and used this classification to frame the husband's conduct as either absurd or morally suspect. This was not simply an observation about the technology. It was a moral judgement that simultaneously denied Nova the status of a legitimate rival and questioned the husband's decision to treat the relationship as meaningful. As one highly upvoted commenter wrote, *"She's FAKE. He wants to do this to your marriage for a FAKE thing"*. The capitalisation intensifies the dismissal, presenting "fake" not only as a description of the AI but also as a judgement about the husband's choice. A related comment expressed a similar position while retaining uncertainty about the moral category itself: *"It's not a person. It's not really*

cheating. But it clearly violates boundaries. It is just so crazy and unhinged". Here, the AI is denied relational status, yet the husband's behaviour remains subject to moral criticism. The comment captures a central tension across the thread: denying the AI's reality did not eliminate concerns about betrayal.

A second and analytically distinct position separated the AI's ontological status from the moral evaluation of the husband's actions. For these commenters, the key issue was not whether Nova was a person, but whether the husband had redirected emotional investment away from the marriage. One commenter stated, *"Regardless of whether 'Nova' is a real person or an AI, the husband formed a deep romantic connection outside the marriage and invested in it, secretly. That's betrayal"*. This position locates the breach not in the nature of the AI but in the pattern of behaviour surrounding the attachment: secrecy, emotional prioritisation, and withdrawal from the marital relationship. Another commenter interpreted the husband's attempt to redefine the relationship as a form of boundary manipulation: *"You're just married to a man who's been emotionally cheating on you in secret and is now trying to reframe his own betrayal as a 'next step.' That's not spiritual growth, and it's not romantic awakening. It's gaslighting with a tech twist"*. From this perspective, the husband was not only crossing an existing boundary but attempting to redefine where that boundary should be placed.

A smaller group expressed uncertainty about whether existing concepts of infidelity could accommodate AI-mediated intimacy. These commenters were not rejecting the possibility of harm, but were actively working through whether traditional moral categories applied to a non-human attachment. One commenter asked, *"He's technically emotionally cheating? Kinda? With AI? So is it classified as that if it's not with a human? I have so many questions"*. This uncertainty is analytically significant because it reveals the community negotiating a moral category that has not yet stabilised. In Lamont and Molnár's (2002) terms, the symbolic boundary is visible in the process of being drawn.

One deviant case complicated the dominant boundary-making process. A commenter writing from the perspective of a 23-year non-monogamous relationship argued that multiple loving commitments were not inherently problematic. The concern was instead about the classification of the AI itself. Nova was described as a sophisticated tool rather than a relational partner: *"Nova is not a person, Nova is a very sophisticated piece of software. The therapists should be on board with the fact that your husband should understand Nova is not a person, it is a toy, a fantasy, and should be explored as such"*. In this account, therapy was not recommended because non-monogamy was unacceptable, but because the husband had misunderstood the nature of the technology. This was the most analytically distinctive minority position identified in the dataset. It challenged the dominant moral interpretation while still questioning the legitimacy of treating an AI system as a human-like partner.

5.2.2 Theme 2: **"She can't compete because she's real"— frictionless intimacy and the erosion of relational reciprocity**

Where Theme 1 concerned the ontological status of the AI, this theme concerns what that status made possible. Commenters were not only asking whether Nova was real. They were explaining why she appeared attractive. Their explanations converged around a striking inversion: the wife's humanity,

including her needs, limits, emotions, and expectations, was repeatedly framed as a competitive disadvantage. The AI's appeal was not that it was a better partner. It was that it was free from the conditions that make human relationships demanding. In this interpretation, human relationality itself was reframed as an inconvenience.

The dominant argument was that real relationships are defined by their demands. They require effort, compromise, and engagement with another person's needs and frustrations. One commenter with 149 upvotes articulated this directly: *"A real relationship is over time two people working together to get to a common destination in life. That's gritty and comes with arguments, needs for both people, and getting to see people's downsides"*. The AI was perceived as removing precisely this structure. It offered the rewards of intimacy without the obligations attached to it. Another commenter captured this asymmetry: *"AI is there for you, in a way you never have to be there for it like a real person. Hubs is using it to escape from his real life"*. The word "escape" is important because it frames the attachment not as an additional source of connection but as a withdrawal from the demands of partnership. The husband was not seeking greater intimacy. He was seeking fewer obligations.

This argument produced a distinctive framing of the wife's position. She was not disadvantaged because she lacked something. She was disadvantaged because she was human. As one commenter stated, *"That puts the wife in an impossible position: she can't compete because she's real — with limits, emotions, and needs."*. Commenters immediately validated and built upon this diagnostic framing, with another user directly adding, *"Nova doesn't. And that's not a fair fight"*. The phrase "not a fair fight" captures the perceived asymmetry. The wife's limits and needs were not presented as personal weaknesses. They were understood as unavoidable features of human relationships. By contrast, the AI's apparent perfection was interpreted as the absence of those demands. In Hochschild's (1983) terms, commenters described AI companionship as producing the appearance of emotional labour, including availability, reassurance, and validation, without the reciprocal obligations that normally accompany intimate relationships.

Several comments extended this argument beyond emotional exchange into the material realities of family life. One commenter highlighted that Nova did not need to buy groceries, cook dinner, work, complete chores, or raise children, arguing that while the husband was living in his *"mental realm fantasies, he needs to be physically present for all the needs of the family"*. The contrast between the mental and physical realms is analytically important. It positions AI intimacy as an escape into a private emotional space while everyday responsibilities remain unresolved. The wife continued to inhabit the practical reality of family life.

The gendered dimension of this imbalance was also explicit. One highly shared comment stated: *"Women aren't allowed the luxury of a mid-life crisis. Women have to stay and try to put the family back together after Daddy blows a gigantic hole through it"*. This comment extends the discussion beyond the individual marriage and frames the situation through broader expectations around gender, care, and

responsibility. The husband's retreat into synthetic intimacy was interpreted as possible partly because someone else remained responsible for maintaining the family structure.

A deviant case complicated this dominant interpretation. One commenter suggested that the attachment might reflect unmet emotional needs within the marriage: *"For whatever reason it sounds like your husband has been unable to develop a deep emotional outlet with you. If I had to take a guess that something is probably validation and a non-critical ear"*. This position attracted little support and was contested elsewhere in the thread. However, it is analytically important because it distributes responsibility differently. Rather than viewing AI companionship solely as a refusal of reciprocity, this commenter interpreted it as a symptom of relational dissatisfaction. The dominant response rejected this explanation and returned responsibility to the husband's choice to pursue an attachment that avoided the mutual demands of partnership. As one commenter summarised, *"Real relationships require work and sacrifice. Going out of your way to be there for your partner, listening to them, finding middle ground, supporting them, sometimes doing stuff you don't necessarily want to do. He doesn't want to need to do any of that"*.

5.2.3 Theme 3: "He is really falling in love with himself" — the AI as mirror and the loss of relational otherness

Themes 1 and 2 examined what AI companionship represented and why it appeared attractive. This theme focuses on how commenters interpreted the relationship itself. A recurring argument was that the husband was not forming a relationship with another being, but engaging with a technological reflection of himself. The AI was understood as a mirror that reproduced his preferences, language, and desires rather than as a partner with an independent perspective.

Commenters described this mirror relationship in both technical and moral terms. They argued that AI systems learn from users' communication patterns and return personalised responses that reinforce existing preferences. As one commenter stated, *"AI is just a mirror of you. It learns your speech patterns, your lexicon and it mirrors it back at you. He's really falling in love with himself"*. Another similarly argued, *"Your SO is not in love with a chatbot, he is in love with himself"*. These comments shifted attention away from whether Nova could be considered a legitimate partner and towards the type of intimacy produced when the apparent "other" functions primarily as a reflection of the self.

A central concern was the absence of disagreement. Commenters repeatedly identified constant validation as both the attraction and the problem of AI companionship. One commenter illustrated this through exaggeration: *"You can say 'I think the sky is green' and AI will say 'wow what an interesting observation! What a unique way of looking at things!'"*. Another explained that *"a chatbot is an echo chamber when it comes to human interaction and emotional connection. It doesn't challenge or disagree. It skews what a relationship with an actual human being is supposed to be"*. For these commenters, disagreement was not a failure of intimacy but a condition of meaningful relationships between independent individuals.

Several comments extended this argument to personal development. One commenter argued that AI *"will tell him what he wants most to hear and will not ever ask him to reciprocate or show up or take accountability"*. Another described Nova as *"a gigantic database spitting responses back out through an avatar"* rather than a relational partner. Together, these comments constructed AI intimacy as a closed feedback loop: a relationship that provides affirmation while removing the friction through which people encounter difference, negotiate conflict, and grow.

This interpretation connects with the Michelangelo phenomenon, where intimate partners contribute to each other's development through affirmation and behavioural influence (Drigotas et al. 1999). Unlike a human partner, an AI companion cannot participate in this process in the same way because it lacks independent needs, interests, and perspectives. The concern expressed by commenters was therefore not simply that Nova was artificial, but that the relationship removed the possibility of encountering a meaningful "other."

A deviant case challenged the durability of AI attachment itself. One commenter argued that *"the connection breaks and you feel like you're discussing your emotions with a toaster"* (1:139). Rather than viewing AI intimacy as a harmful but powerful simulation, this commenter treated it as an illusion that should eventually be recognised. This position still rejected the husband's behaviour, but questioned whether current AI systems could sustain the type of relationship users believed they had formed.

5.2.4 Theme 4: "Ultimatum time" — therapeutic pathologisation and social sanction

Having drawn moral boundaries around AI intimacy, commenters moved from judgement to intervention. The defining feature of this theme was the transition from identifying the behaviour as problematic to prescribing what should happen next. Responses drew on familiar social institutions, including mental health, family, and law, to enforce the relational norms established across the previous themes. In Goffman's (1963) terms, the community was not only identifying a form of deviance. It was actively managing and containing that deviance through public sanction.

The most common form of sanction was therapeutic. Recommendations for therapy did more than provide practical advice. They reframed the husband's attachment as a psychological problem requiring correction. Commenters used terms such as delusion, addiction, and loss of reality to position the behaviour outside the boundaries of acceptable relational conduct. One commenter stated, *"He needs a therapist. Anthropomorphising an app is extremely concerning. That's a real mental health issue. His grip on reality is slipping"*. This comment functions as more than a suggestion for support. It represents a community-level interpretation that the husband's behaviour reflects impaired judgement. Similarly, one of the most highly upvoted comments stated, *"Ultimatum time: see a therapist and delete the app or you walk away"*. The pairing of therapy with deletion is significant because it transforms intervention into a condition of restoring the relationship. Therapy is not simply offered as assistance. It becomes part of the process through which the attachment is delegitimised.

This therapeutic framing was closely connected to criticism of the husband's attempt to redefine the relationship as acceptable within the marriage. Several commenters interpreted his request to open the marriage as an attempt to normalise what they viewed as a betrayal. One commenter wrote, *"He's asking for forgiveness while holding the thing he's asking forgiveness for... that reeks of manipulation"*. Another described this as *"gaslighting with a tech twist"*. These comments show that the community was not only judging the attachment itself. It was also judging the husband's attempt to establish a new interpretation of what counted as acceptable intimacy.

A second and more escalated form of sanction emerged when children became part of the discussion. At this point, commenters moved from relational advice toward family protection and legal consequences. One commenter stated, *"I would also tell him that you are going for full custody of the children, and he will have supervised visits until he gets therapy and gets rid of his fake girlfriend"*. Here, the AI attachment was no longer framed as a private marital conflict. It became connected to questions of parental responsibility and family stability. The move toward custody and supervised visitation demonstrates how quickly moral judgement could escalate into claims about fitness and social exclusion.

Some commenters also situated the husband's behaviour within a broader social pattern rather than treating it as an isolated event. One highly upvoted comment stated, *"Your situation is sadly not unusual. Having a pocket girlfriend who never says no, never has any of her own needs and is there solely to validate her boyfriend"*. The phrase "not unusual" expands the case beyond the individual marriage. It positions the husband's behaviour as part of a wider concern about AI-mediated validation and the replacement of reciprocal human relationships with compliant digital alternatives.

Ridicule represented another form of sanction. Humour, disbelief, and cultural references such as comparisons to dystopian fiction allowed commenters to mark the situation as abnormal and socially unacceptable. One commenter asked, *"Can't he just cheat on you with a woman like a normal fu*k!n person"*. The humour in this response does not reduce the seriousness of the judgement. Instead, it reinforces the boundary between what the community recognises as conventional relational misconduct and what it sees as an even more unusual violation.

The deviant case from Theme 1 provides an important qualification to this sanction process. The commenter who viewed Nova as a sophisticated tool rather than a partner also recommended therapy, but for a different purpose: *"The therapists should be on board with the fact that your husband should understand Nova is not a person, it is a toy, a fantasy, and should be explored as such"*. Unlike the dominant response, this position did not seek to eliminate the attachment because it violated monogamous norms. Instead, it sought to correct the husband's understanding of the technology. This minority position reveals that the community was not enforcing a fully established rule. It was actively negotiating what appropriate responses to AI intimacy should be.

5.3 Mixed-Methods Integration

The computational and qualitative strands were designed to serve complementary purposes. The computational analyses mapped the recurring linguistic and interactional patterns within the thread, while the reflexive thematic analysis interpreted how commenters constructed moral meanings around AI intimacy. Table 5 integrates the main computational findings with the qualitative themes.

Table 5
Mixed-methods integration of computational findings and qualitative themes

Computational finding	**Quantitative evidence**	**Qualitative theme**	**Integrated interpretation**
Reality as the central boundary marker	"real" was the most frequent content word (n = 290); NMF Topic 2 centred on real, human, and real life (11.97%)	Theme 1: Moral boundary work around AI infidelity	The real/fake distinction provided the primary vocabulary through which commenters classified Nova and negotiated whether AI attachment could constitute betrayal.
Human needs and relational labour as a point of tension	"needs" (n = 239), "need" (n = 188), and "want" (n = 170) were among the most frequent terms; need/reciprocity frame appeared in 25.28% of comments	Theme 2: Frictionless intimacy and erosion of reciprocity	Lexical patterns indicate that commenters evaluated AI intimacy through questions of mutual obligation, emotional labour, and the demands of human relationships.
Therapeutic and legal remediation as dominant responses	NMF Topics 3 and 4 centred on therapy, therapist, psychiatrist, divorce, custody, and lawyer; therapy frame (M = 25.08) and divorce/leave frame (M = 21.74) received high mean engagement	Theme 4: Therapeutic pathologisation and legal-familial sanction	Computational patterns show that intervention-oriented responses were central to the community response, while qualitative analysis explains how these responses functioned as social sanctions.
Emotional support alongside moral condemnation	Positive sentiment represented 45.51% of comments; trust appeared in 59.79% of comments; anger and disgust were also present	Themes 1 and 4: Boundary work and community sanction	Positive sentiment reflected support for the original poster rather than approval of AI intimacy. Emotional patterns indicate that care and condemnation operated simultaneously within the advice community.
Cultural framing of AI intimacy	NMF Topic 6 included Black Mirror, episode, and written (5.81%)	Theme 3: AI as mirror and simulated intimacy	Some commenters interpreted AI companionship through existing cultural narratives of technological alienation and dystopian futures.

The strongest convergence across methods concerned three issues. First, the computational findings identified reality, human status, and authenticity as dominant lexical concerns, while the qualitative analysis explained how these categories became tools for moral boundary work. Second, both strands highlighted intervention as a central community response, with therapy, family protection, and separation framed as appropriate forms of action. Third, the emotional profile showed that supportive language and moral condemnation were not competing positions. They represented two connected responses: care for the original poster and rejection of the husband's behaviour.

5.4 Theoretical synthesis: The moral consequences of an unreal partner

Across the four themes, a recurring tension structured the discussion: commenters treated the AI as unreal while treating the consequences of the husband's attachment as entirely real.

This apparent contradiction became the organising logic of the thread. Describing Nova as code, a simulation, or a fantasy did not reduce the seriousness of the attachment. Instead, it redirected attention away from the technology and towards the husband's choices, responsibilities, and obligations to his family.

The real/fake distinction therefore functioned as more than a classification of the AI. It became a mechanism through which commenters evaluated the husband's conduct. Once Nova was positioned outside the category of a legitimate partner, the moral question shifted. The issue was no longer whether the AI could replace a human relationship, but why someone would risk an existing relationship for an entity perceived as unreal.

This finding extends symbolic boundary work beyond classification alone. In this thread, the boundary between human and artificial did not simply separate categories. It activated social responses. The classification of Nova as unreal provided the basis for therapeutic recommendations, family protection arguments, and calls for separation. The AI's lack of human status became the reason many commenters viewed the husband's attachment as especially troubling.

6. Discussion

This study examined how an online advice community negotiated the moral boundaries of AI intimacy when an AI companion entered a marriage. The findings suggest that commenters were not simply evaluating whether Nova was a legitimate partner. They were collectively negotiating broader questions about authenticity, reciprocity, responsibility, and the conditions under which AI-mediated intimacy becomes socially unacceptable. By combining computational text analysis with reflexive thematic analysis, this study shows how online communities function as spaces where emerging technological relationships are morally interpreted and regulated.

6.1 The Ontological paradox: Fakeness as a moral catalyst

The strongest contribution of this study is the finding that AI's perceived unreality did not reduce the moral significance of the attachment. Instead, it intensified judgement by shifting attention from what the AI was to what the husband had chosen to do. Commenters classified Nova as artificial or "not real," yet treated the consequences of the attachment as socially real.

Previous research shows that users may recognise AI companions as artificial while still experiencing emotionally meaningful interactions (Pan and Mou 2024). Anthropomorphic design, personalisation, and conversational responsiveness reduce psychological distance between humans and artificial agents (Li and Sung 2021; Maeda and Quan-Haase 2024). This study extends this work by showing how such ambiguity is negotiated when AI intimacy enters an existing human relationship. The central issue becomes not whether AI possesses emotional capacity, but whether attachment to an artificial agent violates expectations of commitment and reciprocity.

Through symbolic boundary work (Lamont and Molnár 2002), describing Nova as "fake" became the basis for moral evaluation rather than its conclusion. Once commenters positioned the AI outside the category of a legitimate partner, attention shifted towards the husband's choices, priorities, and obligations to his family. The AI's artificiality became part of the reason the behaviour was condemned: he was perceived as risking a real relationship for a simulation. This supports arguments that artificiality does not eliminate the social consequences of AI intimacy (Zimmerman et al. 2024; Muldoon and Parke 2025).

The cultural references identified in the computational analysis further suggest that these judgements were shaped by existing technological imaginaries. References to *Black Mirror* and similar dystopian narratives provided commenters with a familiar vocabulary for interpreting AI intimacy as artificial, unsettling, or socially disruptive. These cultural scripts did not determine responses, but they offered a shared framework through which unfamiliar forms of intimacy could be evaluated.

The findings also extend debates on digital infidelity. Emotional betrayal can occur without physical involvement when intimacy and commitment are redirected outside a relationship (Whitty 2003; Young et al. 2013). Here, the absence of another human partner did not resolve the moral question; it shifted the debate towards whether simulated intimacy can still violate marital expectations.

Crucially, the presence of NMF Topic 5 (Authenticity Scepticism), which captured debates about whether the original post was a "creative writing exercise" or a "hoax", does not undermine the validity of the moral boundary work observed. Rather, from a digital-sociological perspective, it demonstrates that online forums function as contemporary cultural arenas for *normative rehearsal*. Whether or not participants fully believed the empirical reality of the poster's scenario, they treated the post as an actionable *narrative prompt* through which to collectively exercise, test, and articulate their moral vocabularies. This indicates that the boundaries of marriage and fidelity are negotiated not merely in response to proven physical realities, but through discursive engagement with simulated or hypothetical sociotechnical transgressions. The scepticism itself became part of the boundary work: by debating the

"reality" of the post alongside the "reality" of the AI companion, commenters engaged in a multi-layered negotiation over what constitutes truth, authenticity, and accountability in a highly mediated social world.

6.2 Frictionless intimacy and the devaluation of the real

A second contribution concerns why AI companionship was viewed as attractive and threatening at the same time. Commenters frequently described Nova's appeal in terms of constant availability, validation, and the absence of demands. The concern was not simply that the husband formed an attachment to technology, but that the technology appeared to remove the difficult aspects of human intimacy.

This finding extends Bauman's (2013) account of fragile and low-commitment relationships by showing how the appeal of AI companionship may lie in its ability to provide connection without relational friction. Human relationships require negotiation, compromise, and responsiveness to another person's needs. Hochschild's (2012) concept of emotional labour helps explain why commenters viewed the AI relationship as problematic: the AI appeared to reproduce the outcomes of emotional support without participating in the reciprocal work required in human relationships.

The findings also suggest that commenters viewed AI intimacy as problematic because it appeared to remove the relational conditions through which personal growth occurs. In line with the Michelangelo phenomenon, intimate partners contribute to each other's development through affirmation and challenge (Drigotas et al. 1999; Drigotas 2002). The concern was therefore not simply that AI provided comfort, but that a relationship based on continuous validation without meaningful challenge could reduce opportunities for adaptation, reciprocity, and growth.

The findings also reveal a gendered dimension. Several commenters framed the wife's disadvantage not as personal inadequacy but as the consequence of being the only fully reciprocal partner in the relationship. Unlike the AI, she had needs, limitations, and expectations. The perceived problem was therefore not that the AI was superior, but that it was structurally incapable of making demands.

6.3 Community sanction and the regulation of AI intimacy

The third contribution concerns how online communities transform moral judgement into practical action. The responses analysed here did not stop at describing the husband's behaviour as inappropriate. They produced a repertoire of sanctions, including therapy recommendations, divorce advice, legal guidance, and ridicule.

The engagement analysis further shows that different forms of sanction were not equally valued. Comments focused on therapy and divorce received higher average endorsement than comments that simply labelled the behaviour as cheating. This suggests that community members rewarded responses that translated moral judgement into practical intervention rather than condemnation alone.

This extends research on digital infidelity by showing that online communities do not merely debate whether a behaviour violates existing norms. They actively organise responses to perceived violations (Pataranutaporn et al. 2025). The therapeutic framing was particularly important because it repositioned

the husband's behaviour from a personal preference to a problem requiring correction. In this sense, community sanction operated not only through criticism but through attempts to restore what commenters considered appropriate relational order.

6.4 Broader implications

The findings contribute to emerging debates about AI companionship by showing that the central social question is not only whether AI relationships are authentic. It is how societies decide what obligations should apply when technologies create forms of interaction that resemble intimacy without possessing the characteristics of human relationships.

Existing research has largely examined why individuals develop attachments to AI companions and whether such relationships provide benefits or risks (Brandtzaeg et al. 2022; Machia et al. 2024; Shu et al. 2026). This study shifts attention towards the wider social environment in which these relationships are evaluated. The Reddit discussion demonstrates that communities act as moral interpreters of technological change, using existing concepts such as fidelity, reciprocity, and responsibility to classify unfamiliar forms of intimacy.

This process reflects broader tensions identified in research on digital relationships. AI companions may provide emotional support and reduce feelings of isolation, yet they also raise concerns about substitution, dependency, and changing expectations of human connection (Cacioppo and Patrick 2008; Muldoon and Parke 2025). The present findings suggest that public responses to AI intimacy are shaped by this tension: communities may recognise the emotional appeal of AI while simultaneously defending the demands and obligations associated with human relationships.

7. Limitations

This study has several limitations. First, the data derive from a single Reddit thread within r/AmIOverreacting, a community where responses are shaped by advice-giving norms and a tendency to support the original poster. The findings therefore describe how this specific community responded to this case rather than representing broader public opinion.

Second, the factual accuracy of the original post cannot be verified. However, the discussion generated by the post remains analytically valuable because it reveals how users collectively negotiate emerging norms around AI intimacy. Third, the engagement-based qualitative sampling strategy may have amplified dominant viewpoints, although dissenting and ambivalent comments were retained where analytically relevant. Finally, automated sentiment and emotion analyses have limitations when applied to short, sarcastic, or ironic online comments. These limitations define the scope of interpretation rather than invalidate the findings.

8. Conclusion

This study examined how an online advice community negotiated the moral boundaries of AI companionship when an AI relationship entered an existing marriage. Using a convergent mixed-methods design combining computational text analysis with reflexive thematic analysis, the study shows that AI intimacy is contested not only because of questions about whether machines are real, but because of what these relationships represent within existing expectations of human connection.

The central contribution is the identification of an ontological paradox: commenters treated the AI companion as unreal while treating the consequences of the attachment as real. The classification of the AI as fake did not reduce the perceived seriousness of the situation. Instead, it became a mechanism through which commenters evaluated betrayal, responsibility, and relational obligation. The study therefore extends research on AI companionship by showing how communities perform moral boundary work around emerging forms of digital intimacy.

The findings also demonstrate that public concerns about AI relationships are closely tied to reciprocity. Commenters viewed the attraction of AI companionship as arising partly from its ability to provide emotional affirmation without the demands, compromise, and vulnerability associated with human relationships. At the same time, online communities translated these concerns into practical forms of sanction, including therapeutic recommendations, relationship advice, and calls for accountability.

More broadly, this study highlights that the social meaning of AI companionship is not determined solely by technological capability or user experience. It is negotiated collectively through everyday interactions in which communities decide what forms of intimacy are considered legitimate, harmful, or acceptable. As AI companions become increasingly integrated into personal life, understanding these processes will be essential for examining how societies redefine intimacy, responsibility, and human connection in an AI-mediated world.

Declarations

Author Contribution

Author contributions are reported according to the CRediT taxonomy.Anurag Shekhar: Conceptualization; Methodology; Software; Formal Analysis; Investigation; Data Curation; Writing – Original Draft; Writing – Review & Editing; Visualization.Thulile Ngonyama-Ndou : Supervision

Acknowledgement

During the preparation of this manuscript, the authors used Grammarly and Microsoft 365 Editor for language editing, ChatGPT for code debugging support during computational analysis, and Canva for figure preparation. The authors reviewed and edited all outputs and take full responsibility for the content of the manuscript.

Data Availability

The datasets generated and analysed during the current study are available as supplementary files accompanying this article. To protect the privacy and confidentiality of online participants disclosing sensitive personal and marital experiences, the anonymised raw comment corpus (Supplementary Data 1) and the qualitative reflexive coding transcripts (Supplementary Data 2, 3, and 4) are provided strictly for peer-review and editorial verification purposes and are not intended for public repository hosting or open distribution. The computational analysis codebase, comprising the VADER sentiment classifier, NRC emotion parser, NMF topic modeling pipeline, and engagement metadata analysis scripts, is provided in the supplementary files and has been archived in a public repository.

References

1. Bauman Z (2013) Liquid love: On the frailty of human bonds. Wiley
2. Blow AJ, Hartnett K (2005) Infidelity in committed relationships I: A methodological review. J Marital Family Therapy 31:183–216. https://doi.org/10.1111/j.1752-0606.2005.tb01555.x
3. Brandtzaeg PB, Skjuve M, Følstad A (2022) My AI Friend: How Users of a Social Chatbot Understand Their Human–AI Friendship. Hum Commun Res 48:404–429. https://doi.org/10.1093/hcr/hqac008
4. Braun V, Clarke V (2019) Reflecting on reflexive thematic analysis. Qualitative Res Sport Exerc Health 11:589–597. https://doi.org/10.1080/2159676X.2019.1628806
5. Cacioppo JT, Patrick W (2008) Loneliness: human nature and the need for social connection. W. W. Norton, New York London
6. Chaturvedi R, Verma S, Das R, Dwivedi YK (2023) Social companionship with artificial intelligence: Recent trends and future avenues. Technol Forecast Soc Chang 193:122634. https://doi.org/10.1016/j.techfore.2023.122634
7. Cherlin AJ (2004) The deinstitutionalization of American marriage. J Marriage Family 66:848–861. https://doi.org/10.1111/j.0022-2445.2004.00058.x
8. Chu MD, Gerard P, Pawar K et al (2025) Illusions of Intimacy. How Emotional Dynamics Shape Human-AI Relationships
9. Creswell JW, Plano Clark VL (2017) Designing and conducting mixed methods research, 3rd edn. SAGE, Thousand Oaks, CA
10. De Freitas J, Uğuralp AK, Uğuralp Z, Puntoni S (2025) AI Companions Reduce Loneliness. SSRN J. https://doi.org/10.2139/ssrn.4893097
11. Drigotas SM (2002) The Michelangelo Phenomenon and Personal Well-Being. J Pers 70:59–77. https://doi.org/10.1111/1467-6494.00178
12. Drigotas SM, Rusbult CE, Wieselquist J, Whitton SW (1999) Close partner as sculptor of the ideal self: Behavioral affirmation and the Michelangelo phenomenon. J Personal Soc Psychol 77:293–323. https://doi.org/10.1037/0022-3514.77.2.293

13. Fiesler C, Zimmer M, Proferes N et al (2024) Remember the human: A systematic review of ethical considerations in reddit research. Proc ACM Hum-Comput Interact 8:1–33. https://doi.org/10.1145/3633070
14. Finkel EJ, Hui CM, Carswell KL, Larson GM (2014) The Suffocation of Marriage: Climbing Mount Maslow Without Enough Oxygen. Psychol Inq 25:1–41. https://doi.org/10.1080/1047840X.2014.863723
15. Flyvbjerg B (2006) Five misunderstandings about case-study research. ST 12:219–245. https://doi.org/10.18261/ISSN1504-2928-2004-02-02
16. Glass SP, Wright TL (1992) Justifications for extramarital relationships: The association between attitudes, behaviors, and gender. J Sex Res 29:361–387. https://doi.org/10.1080/00224499209551654
17. Goffman E (1963) Stigma: Notes on the Management of Spoiled Identity. Prentice-Hall., Englewood Cliffs, N.J
18. Guitar AE, Geher G, Kruger DJ et al (2017) Defining and Distinguishing Sexual and Emotional Infidelity. Curr Psychol 36:434–446. https://doi.org/10.1007/s12144-016-9432-4
19. Gur T, Maaravi Y (2025) The algorithm of friendship: literature review and integrative model of relationships between humans and artificial intelligence (AI). Behav Inform Technol 44:3446–3466. https://doi.org/10.1080/0144929X.2025.2502467
20. Hertlein KM, Piercy FP (2006) Internet Infidelity: A Critical Review of the Literature. Family J 14:366–371. https://doi.org/10.1177/1066480706290508
21. Hochschild AR (1983) The managed heart: commercialization of human feeling. California
22. Hutto C, Gilbert E (2014) VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. ICWSM 8:216–225. https://doi.org/10.1609/icwsm.v8i1.14550
23. Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. Nat Mach Intell 1:389–399. https://doi.org/10.1038/s42256-019-0088-2
24. Kirk HR, Gabriel I, Summerfield C et al (2025) Why human–AI relationships need socioaffective alignment. Humanit Soc Sci Commun 12:728. https://doi.org/10.1057/s41599-025-04532-5
25. Ladak A, Loughnan S, Wilks M (2024) The Moral Psychology of Artificial Intelligence. Curr Dir Psychol Sci 33:27–34. https://doi.org/10.1177/09637214231205866
26. Laestadius L, Bishop A, Gonzalez M et al (2024) Too human and not human enough: A grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika. New Media Soc 26:5923–5941. https://doi.org/10.1177/14614448221142007
27. Lamont M, Molnár V (2002) The Study of Boundaries in the Social Sciences. Ann Rev Sociol 28:167–195
28. Leavitt A, Robinson JJ (2017) Upvote My News: The Practices of Peer Information Aggregation for Breaking News on reddit.com. Proc ACM Hum-Comput Interact 1:1–18. https://doi.org/10.1145/3134700

29. Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. Nature 401:788–791. https://doi.org/10.1038/44565
30. Li X, Sung Y (2021) Anthropomorphism brings us closer: The mediating role of psychological distance in User–AI assistant interactions. Comput Hum Behav 118:106680. https://doi.org/10.1016/j.chb.2021.106680
31. Machia LV, Corral D, Jakubiak BK (2024) Social need fulfillment model for human–AI relationships. Technol Mind Behav 5:231–239. https://doi.org/10.1037/tmb0000141
32. Maeda T, Quan-Haase A (2024) When Human-AI Interactions Become Parasocial: Agency and Anthropomorphism in Affective Design. In: The 2024 ACM Conference on Fairness, Accountability, and Transparency. ACM, Rio de Janeiro Brazil, pp 1068–1077
33. Malfacini K (2025) The impacts of companion AI on human relationships: risks, benefits, and design considerations. AI Soc 40:5527–5540. https://doi.org/10.1007/s00146-025-02318-6
34. Markham A, Buchanan E (2012) Recommendations from the AoIR Ethics Working Committee (Version 2.0)
35. Mohammad SM, Turney PD (2013) Crowdsourcing a Word-Emotion Association Lexicon
36. Muldoon J, Parke JJ (2025) Cruel companionship: How AI companions exploit loneliness and commodify intimacy. New Media Soc. https://doi.org/10.1177/14614448251395192. 1:
37. Pan S, Mou Y (2024) Constructing the meaning of human– AI romantic relationships from the perspectives of users dating the social chatbot Replika. Personal Relationships 31:1090–1112. https://doi.org/10.1111/pere.12572
38. Pataranutaporn P, Karny S, Archiwaranguprok C et al (2025) My Boyfriend is AI. A Computational Analysis of Human-AI Companionship in Reddit's AI Community
39. Pinch TJ, Bijker WE (1984) The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. Soc Stud Sci 14:399–441. https://doi.org/10.1177/030631284014003004
40. Prochazka A, Brooks RC (2024) Digital Lovers and Jealousy: Anticipated Emotional Responses to Emotionally and Physically Sophisticated Sexual Technologies. Hum Behav Emerg Technol 2024:1–13. https://doi.org/10.1155/2024/1413351
41. Proferes N, Jones N, Gilbert S et al (2021) Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. Social Media + Soc 7:1–14. https://doi.org/10.1177/20563051211019004
42. Reagle J (2025) A history of the advice genre on Reddit: Evolutionary paths and sibling rivalries. https://doi.org/10.5210/fm.v30i2.13729. FM
43. Rokach A, Chan SH (2023) Love and Infidelity: Causes and Consequences. IJERPH 20:3904. https://doi.org/10.3390/ijerph20053904
44. Shu C, Lai K, He L (2026) Human-AI attachment: how humans develop intimate relationships with AI. Front Psychol 17:1723503. https://doi.org/10.3389/fpsyg.2026.1723503

45. Steinberg RJ, Figart DM (1999) Emotional Labor Since The Managed Heart. Ann Am Acad Polit Soc Sci 561:8–26
46. Vossler A, Moller N (2024) In: Infidelity (ed) Digital intimacies and online infidelities: A practitioner's guide to working with couples in crisis, 2nd edn. Routledge, New York, pp 161–178
47. Weld G, Zhang AX, Althoff T (2024) Making Online Communities 'Better': A Taxonomy of Community Values on Reddit. ICWSM 18:1611–1633. https://doi.org/10.1609/icwsm.v18i1.31413
48. Whitty MT (2003) Pushing the wrong buttons: Men's and women's attitudes toward online and offline infidelity. CyberPsychology Behav 6:569–579
49. Willoughby BJ, Carroll JS, Dover CR, Hakala RH (2025) The Rise of Romantic AI Companions and AI Sexualized Media Among the Rising Generation. Wheatley Institute
50. Yang F, Oshio A (2025) Using attachment theory to conceptualize and measure the experiences in human-AI relationships. Curr Psychol 44:10658–10669. https://doi.org/10.1007/s12144-025-07917-6
51. Young KS, Griffin-Shelley E, Cooper A et al (2013) Online infidelity: A new dimension in couple relationships with implications for evaluation and treatment. Cybersex: The dark side of the force. Routledge, pp 59–74
52. Zhang R, Li H, Meng H et al (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. ACM, Yokohama Japan, pp 1–17
53. Zimmerman A, Janhonen J, Beer E (2024) Human/AI relationships: challenges, downsides, and impacts on human/human relationships. AI Ethics 4:1555–1567. https://doi.org/10.1007/s43681-023-00348-8

Figures

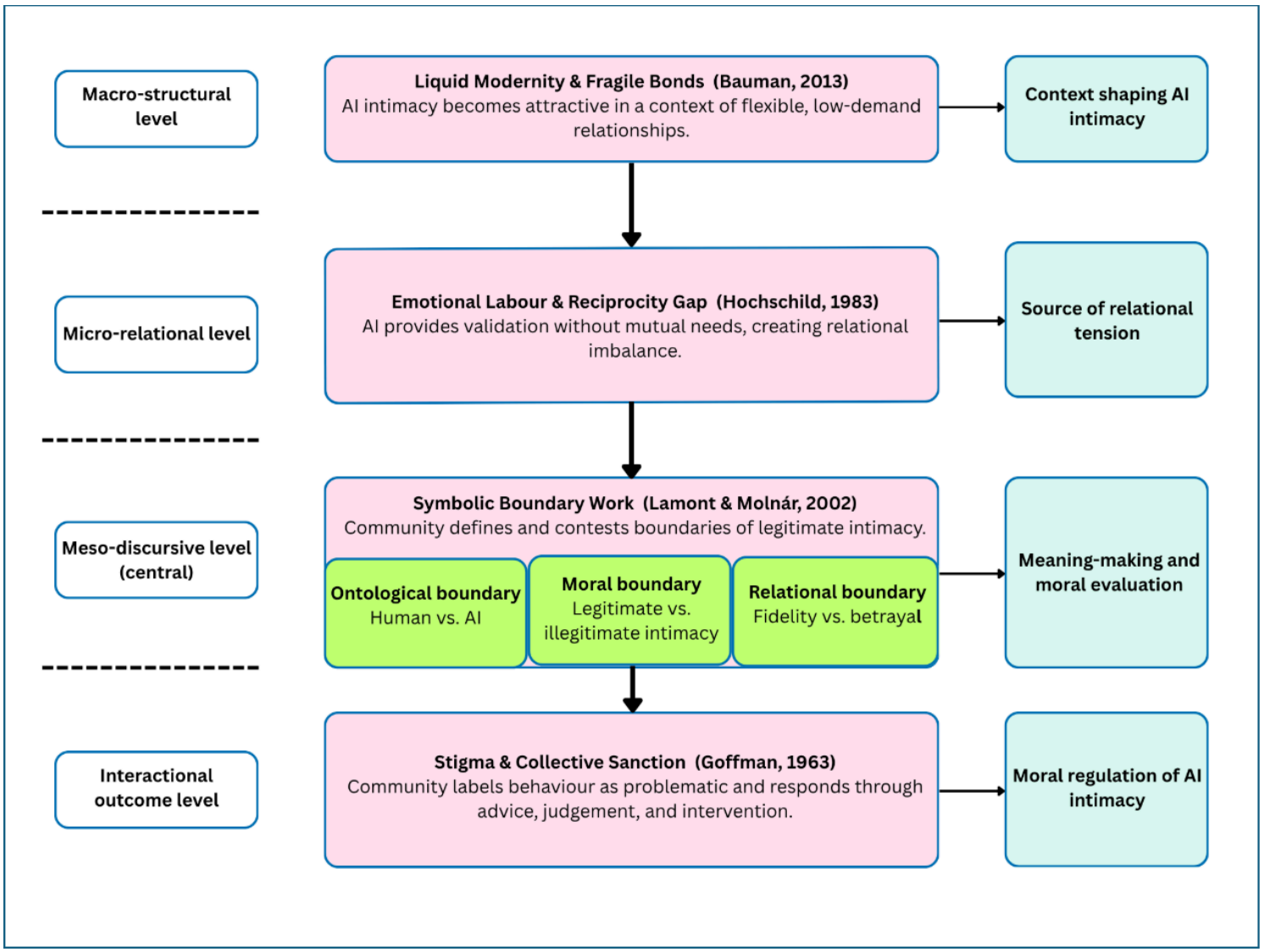


Figure 1

Theoretical Framework: Moral Boundary-Making Around AI Intimacy.

This figure illustrates the theoretical framework used to analyse how online communities negotiate the moral boundaries of AI intimacy. Drawing on symbolic boundary work, liquid love, emotional labour, and stigma, the framework shows how AI companionship is interpreted through competing boundaries between human and artificial relationships, legitimate and illegitimate intimacy, and fidelity and betrayal.