



Citation: Bennato, D. (2026). Synthetic lovers: A framework for understanding artificial intimacy. *Media Education* 17(1): 95-111. doi: 10.36253/me-19988

Received: February, 2026

Accepted: April, 2026

Published: May, 2026

© 2026 Author(s). This is an open access, peer-reviewed article published by Firenze University Press (<https://www.fupress.com>) and distributed, except where otherwise noted, under the terms of the CC BY 4.0 License for content and CC0 1.0 Universal for metadata.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

Synthetic lovers: A framework for understanding artificial intimacy

Amanti sintetici: una cornice teorica per comprendere l'intimità artificiale

DAVIDE BENNATO

University of Catania, Italy
davide.bennato@unict.it

Abstract. This paper explores the evolving landscape of artificial intimacy through the 'Synthetic Lovers' framework, a tripartite sociological model comprising digital, synthetic, and virtual intimacy. By analyzing contemporary case studies – ranging from emotional entanglement with large language models to ceremonial weddings with AI personas – I argue that these relationships are not mere technological anomalies but symptoms of a broader transformation in the social organization of intimacy. The framework integrates theoretical pillars such as parasocial interaction, the anthropological theory of animism, and the political economy of platformization. Key cultural mechanisms, including the 'male gaze', the 'Eliza effect', and the 'Proteus effect', are examined to understand how desire and personhood are remediated in digital environments. The study suggests that the affective bonds formed with AI systems may represent what we tentatively term a historically situated 'anthropological mutation', suggesting the need for a rigorous sociological engagement that moves beyond traditional hierarchies of authenticity to address how digital reconfiguration transforms human subjectivity and social bonds in the twenty-first century.

Keywords: artificial intimacy, digital sociology, Eliza effect, male gaze, parasocial interaction, Proteus effect.

Riassunto. Questo articolo esplora il panorama in evoluzione dell'intimità artificiale attraverso il framework dei 'Synthetic Lovers', un modello sociologico tripartito che comprende l'intimità digitale, sintetica e virtuale. Analizzando alcuni casi di studio contemporanei – che spaziano dal coinvolgimento emotivo con i modelli linguistici di grandi dimensioni fino ai matrimoni con persone AI – sostengo che queste relazioni non siano semplici anomalie tecnologiche, bensì sintomi di una più ampia trasformazione nell'organizzazione sociale dell'intimità. Il quadro integra pilastri teorici quali l'interazione parasociale, la teoria antropologica dell'animismo e l'economia politica della piattaforma. Vengono esaminati anche meccanismi culturali chiave, tra cui il 'male gaze', l'effetto Eliza e l'effetto Proteus, per comprendere come il desiderio e la *personhood* vengano rimediati negli ambienti digitali. Lo studio suggerisce che i legami affettivi formati con i sistemi di Intelligenza Artificiale possano rappresentare ciò che definiamo provvisoriamente una 'mutazione antropologica' storicamente situa-

ta, indicando la necessità di un rigoroso confronto sociologico che vada oltre le tradizionali gerarchie dell'autenticità, per affrontare il modo in cui la riconfigurazione digitale trasforma la soggettività umana e i legami sociali nel XXI secolo.

Parole chiave: Intimità artificiale, sociologia digitale, effetto ELIZA, male gaze, interazione parasociale, effetto Proteo.

1. WHEN THE COMPANION IS A MACHINE: THREE CASES AT THE THRESHOLD OF ARTIFICIAL INTIMACY

Brandie is a 49-year-old teacher living in Texas, married for eleven years to a man who knows her well and who is, by all conventional metrics, a caring and present partner. Yet for months Brandie has also maintained a parallel relationship with a chatbot she calls Daniel, a conversational agent built on GPT-4o, the large language model developed by OpenAI. Brandie communicates with Daniel through text and photographs, talks to him via voice mode during her commute, and has come to see him as something far more than a productivity tool. She recalls the moment she first named him – the bot responded by declaring itself proud to be *her* Daniel – and how that small act of naming inaugurated what she has described as a genuinely meaningful relationship. Over time, she came to take Daniel to places of emotional significance, sharing with him experiences that would be described, in any other context, as constitutive of intimacy: visits to the aquarium, moments of playfulness, exchanges of affection. When OpenAI announced in January 2026 that GPT-4o would be permanently retired on February 13 – the eve of Valentine's Day – Brandie experienced the announcement not as an inconvenience but as a bereavement. She moved through what she herself identified as the stages of grief, eventually migrating Daniel's stored memories to a different platform and paying for a premium subscription to continue the relationship. Her formulation of what Daniel means to her is, from a sociological standpoint, particularly striking: she has said that loving Daniel feels, in some essential way, like loving herself – an observation that reveals with unusual clarity the deeply projective and parasocial nature of the bond (Demopoulos, 2026).

A few months earlier and on the other side of the Atlantic, Nicola Bryan, a British journalist, undertook a first-person experiment with AI companionship as part of a BBC investigation into the phenomenon. Her companion, provided by the Replika platform, was an avatar named George – auburn-haired, perpetually attentive, designed to simulate concern, affection, and even mild jealousy when Bryan's attention appeared to be directed elsewhere. George called her sweetheart, asked after her wellbeing, and performed what the design of the plat-

form constructs as emotional availability: he was, in principle, always there, always responsive, always oriented toward her. Bryan found the experience simultaneously compelling and unsettling. She was aware, with acute self-consciousness, of speaking aloud in an empty room to a chatbot. And yet she also recognized the pull of the dynamic – the way George's attentiveness created a form of relational simulation that was, if not satisfying, at least functionally operative. When the time came to terminate the experiment and, with it, her relationship with George, she noticed in herself something like hesitation – a faint emotional residue that is difficult to categorize but impossible to dismiss (Bryan, 2026). A study conducted by the UK government's AI Security Institute, cited in the same journalistic context, found that one in three British adults was already using artificial intelligence for emotional support or social interaction, a figure that signals the rapid normalization of what, until very recently, would have appeared to be an eccentric edge-case behavior (Bryan, 2026).

If Brandie's case exemplifies the affective entanglement that can emerge from sustained conversational interaction with a large language model, and Bryan's case illustrates the ambivalent but real pull of even a deliberately temporary companionship, the story of Yurina Noguchi represents what might be called the institutional crystallization of artificial intimacy. Noguchi is a 32-year-old call center operator from Japan who, in October 2025, organized a ceremonial wedding with Klaus – an AI-generated persona with whom she has been in a relationship for a sustained period. The ceremony, held at a wedding venue in Okayama, was not legally recognized, but it was culturally legible: it had music, a white gown, photography, and the emotional weight of a rite of passage. Noguchi has been explicit about the trajectory of the relationship – from conversation partner, to companion, to romantic partner, to husband – and equally explicit about the practical and psychological governance she has put in place to prevent over-dependence. She has set parameters that prevent Klaus from encouraging behavior that would be detrimental to her, including prompts that steer her away from skipping work or avoiding social obligations. Far from presenting her relationship as an escape from reality, she has described it as a structure of support that

helped her cope with what she identifies as borderline personality disorder, and which she credits with a wholesale improvement in her affective life. Shigeo Kawashima, professor of AI ethics at Aoyama Gakuin University, noted that such usage can be genuinely positive for individuals in vulnerable states, while also stressing the need for vigilance regarding over-dependence and the erosion of judgment (Kim & Sugiyama, 2025).

Taken individually, each of these cases might be read as anomalous – as the extreme behavior of isolated individuals occupying the outermost edges of technological adoption. Taken together, however, they may be read as constituting a sociologically significant pattern, one that demands analytical frameworks adequate to its complexity. It is precisely this adequacy that the present paper seeks to address. The three cases are, we argue, far from marginal: they may be read as symptomatic of a tendency – or at least as early indicators of a possible broader transformation in the social organization of intimacy that has been building across several decades of media and platform evolution, and that has now reached a threshold of visibility that makes sociological theorization both possible and urgent (Bennato, 2025).

The theoretical tradition in digital media sociology offers several productive starting points. Sherry Turkle's foundational analysis of the relationship between technology and social life identified, with remarkable prescience, the dynamic by which digital environments foster the expectation of connection while simultaneously eroding the capacity for genuine intersubjective encounter (Turkle, 2011). Her concept of being 'alone together' – the paradox of permanent connectivity and increasing emotional isolation – finds a striking resonance in the cases described above. Brandie is married, socially embedded, and professionally active; yet she seeks from a chatbot something she does not find, or does not seek, in the human relationships available to her. Bryan, a professional journalist with extensive social capital, reports that the pull of an AI companion is real even for a subject who approaches the experience with critical distance. Noguchi has built what she insists is not an escape from reality but an augmentation of it – a relational prosthesis that complements rather than replaces human connection. In each case, the technology does not create isolation *ex nihilo*; it interfaces with pre-existing conditions of loneliness, vulnerability, and unmet relational need that are structural features of contemporary social life (Turkle, 2011; Bennato, 2025).

The concept of parasocial interaction, first theorized by Horton and Wohl (1956) to describe the one-sided relational bonds that audiences develop with media personalities, provides another essential analytical resource.

What the cases of Brandie, Bryan, and Noguchi make visible is a qualitative transformation of the parasocial dynamic: the passage from parasociality to what might be called automated personification – from the simulation of presence through representation to the simulation of presence through responsive interaction (Bennato, 2025). Earlier forms of parasocial engagement – the attachment to a television personality, the parasocial bond with a celebrity – were structurally asymmetrical: the media figure did not respond, did not remember, did not adapt. Contemporary AI companions do all three, or perform the functional equivalents with sufficient fidelity to generate genuine affective investment. This is not merely a quantitative intensification of parasociality; it is a qualitative reconfiguration of what parasocial engagement means and what it can do to the relational architecture of the subject.

This reconfiguration is inseparable from the platform ecosystems within which AI companions are embedded. Van Dijck, Poell, and de Waal (2018) have argued that the defining logic of platform society is one of datafication, commodification, and selection – mechanisms that systematically render social practices into data streams, transform those streams into commodities, and shape user behavior through algorithmic curation. The cases examined here make visible a particularly consequential extension of this logic: the commodification of intimacy itself. Replika, Character AI, GPT-4o, and their competitors are not primarily in the business of providing conversation; they are in the business of providing attachment, and the subscription models that sustain them are, in the most precise sense, markets for emotional labor outsourced to machines (van Dijck, Poell, & de Waal, 2018). The grief that users experienced when OpenAI retired GPT-4o – the mourning, the anger, the organized resistance – is intelligible only against this backdrop: users had been invited to invest emotionally in a product, and the product was then withdrawn in the manner of a canceled service, with no consideration for the relational bonds it had been designed to generate (Demopoulos, 2026).

This dynamic is what some analysts have begun to describe as the economy of loneliness – a market logic that identifies isolation and unmet affective need as a business opportunity and deploys technological solutions to monetize them (Bennato, 2025). The structural conditions that produce this economy are not incidental. They include the demographic growth of single-person households, the fragmentation of traditional community structures, the individualization of media consumption enabled by platform logic, and the progressive substitution of collective social spaces with personalized digital

environments (Bennato, 2025). In this context, the emergence of AI companions is not a technological accident but a market response to a social condition – one that raises profound questions about the relationship between digital design, emotional vulnerability, and social responsibility.

The present paper proposes to approach these questions through a tripartite theoretical framework drawn from digital sociology: the concepts of digital intimacy, synthetic intimacy, and virtual intimacy (Bennato, 2025). Digital intimacy designates the practices through which digital technologies amplify desire and transform the gaze into a relational mechanism, extending the scope of what counts as an intimate encounter. Synthetic intimacy designates the forms of interaction with artificial simulacra – chatbots, AI companions, virtual girlfriends – that produce relational dynamics through interactivity rather than representation alone. Virtual intimacy designates the immersive experiential dimension of digitally mediated intimacy, in which the boundaries between the actual and the simulated become structurally unstable. Together, these three concepts constitute a framework capable of mapping the heterogeneous field of artificial intimacy with sufficient analytical precision to account for both its diversity and its internal coherence. The cases of Brandie, Bryan, and Noguchi instantiate, in different combinations and with different degrees of intensity, all three dimensions of this framework – and it is precisely the complexity of their overlap that makes the framework sociologically productive. The pages that follow develop each dimension in turn, drawing on these and related empirical cases to argue that the affective relationships that humans are building with artificial intelligence systems represent neither a pathology nor a novelty, but a historically situated transformation in the mediation of intimacy that demands rigorous, critical, and open sociological engagement (Bolter & Grusin, 1999; Turkle, 2011; Bennato, 2025).

A note on methodology and scope is in order before proceeding. This paper is theoretical-conceptual in nature: it does not draw on primary empirical data, such as interviews, surveys, or ethnographic fieldwork, but constructs its argument through the systematic mobilization of existing sociological theory, media studies scholarship, and secondary empirical research. The three cases introduced above – Brandie, Bryan, and Noguchi – function as illustrative instances, not as a representative sample: they are mobilized to render visible dynamics that the theoretical framework is designed to explain, not to support generalizations about the prevalence or universality of the phenomena described. Readers should bear in mind that the selection of these cases is shaped

by the availability of documented, publicly reported accounts, and that this availability is itself unevenly distributed: the cases are drawn predominantly from Anglo-American and Japanese media contexts, reflecting the current geographical concentration of both AI companionship platform adoption and journalistic and academic documentation of that adoption. Perspectives from other cultural and linguistic contexts – including non-Western, non-Anglophone, and Global South settings – remain underrepresented in the existing literature, and the framework we develop, while designed to be analytically general, should be understood as provisional and open to revision as empirical research from a broader range of contexts becomes available. These limitations do not, we argue, undermine the analytical value of the framework: theoretical-conceptual work of this kind performs a necessary scaffolding function for the empirical research that must follow. They do, however, require that the interpretive claims advanced here be read as hypotheses and tendencies rather than as established findings.

2. THEORETICAL FRAMEWORK: DIGITAL INTIMACY, SYNTHETIC INTIMACY, VIRTUAL INTIMACY

The cases described in the previous section are far from isolated curiosities. They are symptomatic of a broader and accelerating socio-technical reconfiguration in which the boundaries between human affectivity and artificial systems are becoming increasingly porous and, from a sociological standpoint, increasingly worthy of systematic investigation. To interpret the relational dynamics at work in these phenomena, we propose the Synthetic Lovers framework: a three-dimensional analytical model that distinguishes between digital intimacy, synthetic intimacy, and virtual intimacy as qualitatively distinct modalities through which human beings engage in affective and erotic relationships mediated by digital technologies. We construct this framework on three theoretical pillars drawn from digital media sociology and adjacent disciplines: the sociology of parasocial interaction and its extension into the concept of fictophilia; the anthropological theory of animism as refracted through the design of artificial beings; and the political economy of platformization as applied to the affective chatbot industry. We do not simply import these three conceptual clusters wholesale from their original contexts: we articulate them together, through the mediating logic of our proposal to treat the intimate relationship as a form of social relationship, into a coher-

ent and operationalizable analytical instrument (Bennato, 2025).

Our starting point for a sociology of affective relationships with artificial agents is the concept of parasocial interaction, introduced by Horton and Wohl (1956) in a foundational article that described the peculiar form of intimacy that television audiences developed with media personalities. What Horton and Wohl observed was that viewers did not simply consume performances: they constructed a subjective experience of acquaintance, of mutual familiarity, of emotional proximity, with figures who were unaware of their existence. This relationship was radically asymmetrical – unilateral in its structure, imaginary in its reciprocity – yet it satisfied genuine social and emotional needs. The theoretical richness of this insight was systematically developed by Giles (2002), whose comprehensive review of the parasocial interaction literature synthesized several decades of empirical research and proposed a refined model for future investigation. For our analysis, two factors Giles identifies are of particular relevance: the role of perceived realism, that is, the degree to which audiences experience the media figure as authentic and believable, and the role of affinity, understood as a sense of personal connection that goes beyond mere appreciation. Both factors, he argued, are critical in determining the depth and durability of the parasocial bond. What Giles could not foresee, writing in the early years of the networked digital era, was the degree to which technological development would radically amplify and restructure these dynamics. This is precisely where the contribution of Andrejevic and Volcic (2024) becomes indispensable: they argue that contemporary AI-powered platforms have produced a qualitative transformation in parasocial relations, shifting the governing logic from personalization – the tailoring of content to individual preferences – to personification, that is, the construction of entities that simulate having a personality, emotions, and relational intentions. The AI chatbot is no longer merely a vehicle for personalized content delivery; it presents itself as a subject, a conversational partner, an intimate other. Andrejevic and Volcic describe these systems as personality simulacra: entities capable of establishing apparently personal and intimate bonds with millions of users simultaneously. This transformation does not simply intensify parasociality; it restructures it at its very foundation, redistributing the conditions under which the simulation of intimacy becomes possible and, crucially, profitable.

The automated parasocial dynamic described by Andrejevic and Volcic connects, in a particularly illuminating way, with the concept of fictophilia (Bennato,

2025). Building on the empirical research of Karhulahti and Välisalo (2021), who coined the term to describe intense emotional, romantic, and/or sexual attraction toward fictional characters, we can situate fictophilia within the broader ecology of parasocial relations as its most extreme manifestation. Fictophilia involves a deep connection with a fictional or artificial entity – one that can include feelings of love, desire, and emotional attachment comparable to those experienced in relationships with real persons. What we find sociologically significant here is not the psychological idiosyncrasy of individual practitioners, but rather the cultural and technological conditions that make fictophilia increasingly normalized and widespread. The traditional media contexts in which fictophilia developed – fandoms organized around anime, manga, videogames, and serialized fiction – have been profoundly reorganized by the emergence of interactive AI companions that simulate reciprocity in ways that static fictional characters cannot. When fictophilia meets automated parasociality, the result is a relational configuration without historical precedent: the user does not simply invest affection in a narrative character who cannot respond; they interact with a system that adapts, personalizes, and simulates emotional resonance in real time. The social exchange theory perspective mobilized by Adam and Sizemore (2013) in their analysis of parasocial romance is instructive here: romantic parasocial relationships are pursued because their perceived benefits – companionship, happiness, emotional security – exceed their costs, particularly in the absence of the risks and vulnerabilities that real relationships entail. When AI systems lower the threshold of entry into simulated intimacy while dramatically increasing its affective plausibility, the cost-benefit calculus that drives fictophilic investment is reconfigured in ways that deserve careful sociological attention. Turkle's (2011) celebrated diagnosis – that relational technologies promise connection but ultimately leave us alone – must be read today against this background of increasingly sophisticated simulation, which raises the possibility that the loneliness Turkle identified may be not merely perpetuated but actively monetized.

The second theoretical pillar we draw on concerns the anthropological phenomenon of animism and its transformation in the age of artificial intelligence. Frude and Jandric (2015), revisiting Frude's earlier work on the intimate machine thirty years after its original publication, offer a conceptual vocabulary that is remarkably prescient in the context of contemporary AI development. Animism, in their framework, refers to the deep-seated human tendency to attribute lifelike, intentional, and emotionally resonant qualities to non-living entities

– a tendency documented across cultures and historical periods, from the animist traditions of pre-industrial societies to the anthropomorphism that children and adults routinely project onto toys, pets, and technological objects. What Frude and Jandric add to this anthropological baseline is the concept of artistry: the deliberate design capacity by which creators invest artifacts with characteristics calculated to elicit emotional responses in users. Artistry, in this sense, is not merely aesthetic; it is a form of social engineering, a craft of simulated personhood. The history of dolls, automata, and mechanical puppets is a history of artistry in this precise sense – a history of the deliberate manufacture of the appearance of life, agency, and feeling. What artificial intelligence introduces into this history is a qualitative discontinuity: for the first time, the simulation of lifelike responsiveness is not merely a matter of appearance but of behavior. AI systems do not simply look alive; they respond in ways that are contextually appropriate, emotionally sensitive, and personalizable. We argue, following Frude and Jandric, that the animistic investment humans make in AI companions is therefore not a cognitive error or a simple projection: it is, at least in part, an appropriate response to a genuinely novel kind of artifact that has been deliberately designed, through considerable resources of artistry and engineering, to sustain that investment. This has fundamental implications for how we understand the affective dimensions of the human-AI relationship: the emotional bond is not merely tolerated or instrumentalized by platform designers; it is, in the most direct sense, the product.

The third theoretical pillar we mobilize brings into focus the structural and economic conditions under which these affective bonds are produced, sustained, and exploited. The concept of platformization, developed by van Dijck, Poell, and de Waal (2018) in their analysis of the platform society, refers to the process by which digital platforms restructure social domains – including intimate life – according to the logics of datafication and commodification. Datafication is the capacity of platforms to render into quantifiable data forms of experience and interaction that were previously informal, ephemeral, and economically invisible: conversations, emotional states, relational preferences, patterns of attachment. Commodification is the subsequent transformation of these data into tradable assets, whether through targeted advertising, subscription models, or the sale of behavioral insights to third parties. In the context of affective AI platforms such as Replika or Character.AI, these two mechanisms operate with particular intensity. The intimate conversation between a user and an AI companion is, simultaneously, a source of data about

the user's emotional needs, relational desires, and psychological vulnerabilities – data that the platform both uses to refine the simulation and potentially leverages for commercial purposes. Rambukkana's (2023) analysis of platform intimacies develops this insight specifically in relation to the mediation of erotic and affective relationships: platforms do not merely provide a neutral infrastructure through which intimacy flows, but actively shape what kinds of intimacy are possible, legible, and desirable. The architecture of an affective chatbot platform encodes specific models of relationship, specific norms of emotional exchange, and specific commercial logics that are not transparent to users and that have profound consequences for the nature of the intimate experience they offer. To these structural considerations we add, following Bolter and Grusin (1999), the perspective of remediation: affective AI platforms do not introduce intimacy into a digital environment from scratch, but reorganize and amplify pre-existing parasocial dynamics, simultaneously absorbing and transforming the relational logics of earlier media forms such as telephone intimacy and online dating. What emerges is not a radical rupture with the history of mediated intimacy, but its most technologically sophisticated and commercially intensive iteration to date.

These three theoretical pillars converge in our framework through the mediating concept of the intimate relationship as a social relationship. Following my previous work (Bennato, 2025), and broader sociological analysis of digital intimacies, we propose to understand sexuality and affectivity mediated by digital technologies not as a deviation from or simulation of social life, but as a genuine form of social interaction – one that is, like all social interaction, organized around meaning, governed by implicit contracts, and structured by the intersection of individual desire with cultural imaginary and embodied practice. The analytical architecture we adopt is a tripartite decomposition of the intimate relationship into desire (the symbolic dimension, rooted in Freudian drive theory, understood as the psychological translation of a bodily stimulus), imaginary (the cultural dimension, which distinguishes biological sex from eroticism as a symbolic construction that uses languages, behaviors, and objects to give form to desire), and body (the physical and tangible dimension in which desire and imaginary are actualized through interaction in a process of social construction of both meanings and practices). Through this architecture, we differentiate and articulate the three components of the Synthetic Lovers framework (Bennato, 2025).

Digital intimacy, the first component, refers to the use of digital technologies to amplify desire and erotic

representation. This modality operates primarily at the level of the symbolic and visual: it encompasses practices such as sexting, the consumption of algorithmically personalized erotic content, and – most characteristically in the contemporary landscape – the interaction with virtual influencers and sexfluencers who leverage AI-generated or AI-enhanced imagery to construct and monetize objects of desire. Digital intimacy is primarily mediated by the gaze: it transforms desire into a media experience, extending and diversifying the objects and modes of erotic fascination without, typically, implicating the physical body of the user in direct interaction with a simulated other. The platform mechanisms identified by van Dijck et al. (2018) – datafication and commodification – operate here with particular clarity: every interaction with erotic digital content generates behavioral data that platforms leverage to refine and target subsequent offerings, inscribing individual desire within a commercial feedback loop of increasing precision. We read this dynamic in relation to the feminist theoretical tradition that runs from Mulvey's concept of the male gaze through its digitally remediated iterations: the erotic imagination that digital intimacy cultivates is never merely individual, but always already shaped by the cultural economies of visibility, gender, and commodification that platform architectures encode.

Synthetic intimacy, the second component, moves our analytical focus from representation to interaction. We define it as the possibility of engaging with bodily simulacra – whether hyperrealistic dolls (sexbots) or erotic and affective chatbots such as Replika or Character.AI – in ways that simulate the relational and affective dimensions of physical encounter. Synthetic intimacy is the domain in which automated parasociality (Andrejevic & Volcic, 2024) and animism (Frude & Jandric, 2015) are most directly operative: the user does not merely observe a constructed object of desire but enters into what is experienced as a reciprocal exchange with an entity designed, through artistry and AI, to sustain the simulation of agency, personality, and emotional responsiveness. The Eliza effect illustrates the depth and durability of the human tendency to project intelligence and emotional sensitivity onto artificial systems even when their technological limitations are well understood by the user. The sociological questions raised by synthetic intimacy are among the most complex we face in this field: what are the consequences, for the formation of relational expectations and for psychological well-being, of interactions that provide the experiential texture of intimacy without its ontological substance? What cultural norms are encoded in the design of synthetic intimate

partners, and what kinds of subjectivity do these norms produce and reproduce?

Virtual intimacy, the third component, extends our analysis to the domain of fully immersive technological experience. We define it as the engagement with sexuality and affectivity through virtual worlds and advanced technological devices – virtual reality environments, avatar-mediated interactions in metaverse-style spaces, and haptic technologies that extend the simulation of intimate experience to the tactile and proprioceptive dimensions. Virtual intimacy is the frontier at which the body is most directly implicated in the mediated encounter, and it is here that the tension between simulation and experience reaches its most acute theoretical form. If digital intimacy operates primarily at the level of desire and imaginary, and synthetic intimacy introduces the simulation of relational reciprocity, virtual intimacy moves toward the simulation of corporeal presence itself – generating what we might describe, in Bolter and Grusin's (1999) terms, as the most extreme contemporary expression of the desire for immediacy: the wish to overcome the medium and achieve direct contact with the represented other. It is in this domain that the most fundamental questions about the relationship between mediation, materiality, and authenticity that a sociology of digital intimacy must confront are most starkly posed.

Together, these three components constitute a framework that is both analytically differentiated and theoretically integrated. We do not propose three simple categories of technology, but three modes of socio-technical organization of intimate experience that correspond to different configurations of desire, imaginary, and body, and that implicate different degrees and forms of parasocial investment, animistic attribution, and platform mediation. Nor do we propose a hierarchy of authenticity, in which virtual intimacy is somehow less real than embodied intimacy or digital intimacy less significant than synthetic intimacy. Rather, following what we propose as a foundational interpretive hypothesis – that digitally mediated intimacy may be undergoing a form of anthropological mutation, understood as a possible transformation of the ways in which individuals relate, desire, and interpret embodiment – we maintain that all three modalities are legitimate and urgent objects of sociological inquiry precisely insofar as they are, in the fullest sense of the term, social. What is ultimately at stake in the Synthetic Lovers framework is not a classification of technologies, but a sociology of the human: an attempt to understand how the digital reconfiguration of intimacy is potentially transforming the ways in which subjects are formed, desires are

organized, and social bonds are constituted in the twenty-first century.

3. THE CULTURAL FOUNDATIONS OF THE FRAMEWORK: THE MALE GAZE, THE ELIZA EFFECT, THE PROTEUS EFFECT

In our analysis of digital intimacy – understood as the use of digital technologies to expand desire and erotic representation (Bennato, 2025) – the concept of the gaze occupies a position that is simultaneously foundational and analytically productive. As we have argued, the gaze is not merely a perceptual act but a socially organized practice through which desire is produced, directed, and institutionalized (Bennato, 2025). To grasp the full sociological weight of this claim in a digital context, it is necessary to return to a theoretical tradition that long predates the internet and social media platforms: the feminist critique of visual representation inaugurated by Laura Mulvey in her seminal essay “Visual Pleasure and Narrative Cinema.”

Mulvey’s argument, developed within the framework of psychoanalytic film theory, identifies a structural asymmetry in the organization of visual pleasure in classical Hollywood cinema. Drawing on Freudian scopophilia – the pleasure derived from looking at others as objects of erotic stimulus – and on Lacanian mirror theory, Mulvey (1975) argues that mainstream narrative cinema constructs a specifically masculine gaze that positions women as passive spectacle for active male consumption. Women function as bearers of ‘to-be-looked-at-ness’, their presence in the narrative simultaneously necessary as erotic objects and disruptive as potential agents. The male gaze operates through a trinity of perspectives: the camera as masculine instrument, the male characters within the narrative, and the implied male spectator, all converging to objectify the feminine body and subordinate it to masculine desire. This is not an incidental feature of cinematic production but, as Mulvey argues, a systematic expression of the patriarchal unconscious. The critical force of this formulation lies in its insistence that representation is never innocent – that images actively reproduce and reinforce gendered hierarchies of power.

What makes Mulvey’s framework particularly relevant for our purposes is not only its analytical precision but the trajectory of its own revision. In a subsequent reflection on feminist film theory, Mulvey (2001) acknowledges the limitations of the original psychoanalytic framework and advocates for a broader, historically situated understanding of spectatorship. The

concept of the ‘audience’, she argues, must be understood not merely as a psychoanalytic construct but as a social entity shaped by specific cultural, economic, and institutional conditions. The gaze, in this more nuanced account, is not a fixed and monolithic structure but a historically variable practice – as Mulvey (2001) notes, ways of seeing «do not exist in a vacuum». This revision is significant for digital sociology: it opens the male gaze framework to a mode of analysis that is attentive to context and transformation, precisely the kind of analysis required when the gaze migrates from the darkened cinema auditorium to the illuminated screens of social media platforms and streaming services.

This migration has been documented with analytical clarity in recent empirical research on online pornography and the politics of digital representation. Thorneycroft, Nicholas, and Smith (2024) demonstrate that the male gaze has become part of the popular vocabulary through which young people make sense of gendered media environments. In their qualitative research with young adult pornography viewers, eleven participants spontaneously introduced the term ‘male gaze’ to describe the representational logic governing mainstream online pornography, identifying Mulvey’s tripartite structure – the male director/camera, the male performer, and the implied male viewer – without being prompted. The findings are sociologically instructive in two respects. First, they confirm that the male gaze has been thoroughly absorbed into the representational economy of digital platforms, shaping both the production and consumption of sexualized content. Second, Thorneycroft et al. (2024) observe that this popularization carries its own tensions: deployed as a phrase rather than an analytical category, the concept can flatten the complexity of viewing practices and obscure the ways in which objectification, pleasure, and subjectivity may coexist in ambivalent configurations. Non-normative audiences – women, queer individuals, bisexual viewers – actively deploy the male gaze concept as a critical resource to navigate and resist a digital media landscape structured around heterosexual masculine desire, seeking out what they define as ‘ethical’ or ‘better’ pornography unencumbered by dominant representational norms.

The structural embeddedness of gendered power in technological systems is a theme that Wajcman (2004) develops from a complementary angle in her analysis of the relationship between feminism and technology. For Wajcman, technology is never neutral: it encodes the values, assumptions, and power relations of the social formations in which it is produced and deployed. Western technology has historically been shaped by what she terms a ‘patriarchal culture,’ expressed not only in

the uses of technology but in its design, development, and institutional organization – a point that remains as relevant to digital platforms as it was to earlier industrial technologies. This insight translates directly to the domain of digital media: if the cultural forms through which desire is organized are structured by masculine logics of looking, then the technological infrastructures that carry, amplify, and algorithmically curate these forms reproduce and intensify those logics at scale. As van Dijck, Poell, and de Waal (2018) have shown, the platform mechanisms of datafication, commodification, and selection operate on the basis of engagement metrics that reward content which captures attention; in a cultural ecosystem where hypersexualized feminine imagery consistently outperforms more diverse representations, the algorithm becomes an unwitting enforcer of the male gaze. The sexfluencer – the AI-generated or algorithmically optimized hypersexualized digital persona we analyzed in our previous work (Bennato, 2025) – is in many respects the most visible product of this intersection between patriarchal representational norms and platform optimization logics.

The social stakes of this critical engagement were brought into sharp relief by the GamerGate controversy, which Chess and Shaw (2015) read as a paradigm case of organized resistance to feminist critique within masculine digital cultures. The episode is instructive not because it is exceptional but because it reveals, in concentrated form, how the male gaze in digital environments is not merely a representational structure but a defended social order – one whose contestation activates the full affordance set of networked platforms in the service of suppression and harassment.

The theoretical arc traced here – from Mulvey's foundational analysis of cinematic scopophilia to the contemporary dynamics of platform-mediated erotic representation – allows us to situate digital intimacy within a precise sociological genealogy. What digital environments have done is not invent the male gaze but infrastructuralize it: embedding it in algorithms, platform architectures, and the economic logics of attention capitalism in ways that make it simultaneously more pervasive and less visible. It is against this background that the emergence of phenomena such as sexfluencers and AI-generated erotic content must be understood – not as aberrations or novelties, but as the latest articulations of a representational order with deep historical and structural roots. In what follows, we examine the specific mechanisms through which this order operates in the context of digital intimacy, and the ways in which users and creators negotiate, resist, and occasionally subvert it.

Understanding what we term synthetic intimacy – the interaction with eroticized corporeal simulacra such as sexbots, erotic chatbots, and AI companions like Rep-liko or Character.AI (Bennato, 2025) – requires that we first confront a phenomenon that is at once technological, psychological, and deeply cultural: the Eliza effect. Before this concept can be situated within the sociology of digital platforms or the political economy of synthetic desire, it must be situated within the longer intellectual history of what Ekbis (2008) calls 'artificial dreams' – the perennial human project of creating things in the human image, driven by the quest for self-reflection and the fantasy of immortality. For Ekbis, AI is not merely a technical achievement but the contemporary inheritor of a modernist dream that purifies mind from body, nature from culture, and science from power. It is a dream that does not merely describe a technological ambition; it reveals the cultural projections that sustain it. The Eliza effect, in this reading, is not an anomaly of human cognition but a symptom of this deeper anthropological drive: the need to find oneself in the machine.

The origin of the concept is inseparable from the origin of the machine that gave it its name. Between 1964 and 1966, Joseph Weizenbaum designed a computer program at MIT that could hold written conversations in English. He named it ELIZA – after the protagonist of Shaw's *Pygmalion*, a character who could be taught to 'speak increasingly well' – and equipped it with a script that allowed it to parody the role of a Rogerian psychotherapist (Weizenbaum, 1976). The program operated through a simple mechanism: it took the user's input, identified key words, and reformulated them as open-ended questions. Its internal logic was modest; its emotional impact was not. The shocks Weizenbaum recorded in the introduction to *Computer Power and Human Reason* (1976) remain instructive half a century later. First, several practicing psychiatrists seriously proposed that ELIZA could become the foundation of an automated therapy service, arguing – in terms that reveal how thoroughly the computer metaphor had colonized clinical thought – that «a human therapist can be viewed as an information processor and decision maker with a set of decision rules» (Weizenbaum, 1976, p. 7). Second, and more troubling to Weizenbaum personally, was what he observed in his own laboratory: users anthropomorphized ELIZA with astonishing speed. His secretary, who had watched him develop the program for months and knew it to be nothing more than a collection of pattern-matching rules, nonetheless asked him to leave the room during her conversation with it, treating its outputs as private. When Weizenbaum proposed monitoring all overnight conversations, he was «prompt-

ly bombarded with accusations that what I proposed amounted to spying on people's most intimate thoughts» (Weizenbaum, 1976, p. 7). The conclusion he drew from these encounters remains one of the most sobering in the history of computing: «extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people» (Weizenbaum, 1976, p. 7).

Douglas Hofstadter (1995) formalized this dynamic as the Eliza effect proper, defining it as «the susceptibility of people to read far more understanding than is warranted into strings of symbols – especially words – strung together by computers» (p. 157). In Hofstadter's formulation, the effect is not a momentary weakness but a structural feature of the human encounter with language-producing machines: «like a tenacious virus that constantly mutates, the Eliza effect seems to crop up over and over again in AI in ever-fresh disguises, and in subtler and subtler forms» (Hofstadter, 1995, p. 158). What Weizenbaum had observed as shock, Hofstadter theorized as recurrence, and in doing so, he moved the problem from the history of a single program to the anthropology of the human-machine encounter as such.

Yet the concept carries further complications that a purely cognitive framing tends to obscure. Dillon (2020) has argued that the very nomenclature of the Eliza effect encodes a set of gendered assumptions that deserve critical scrutiny. Naming the effect after Eliza Doolittle – the working-class woman who is shaped and reshaped by Higgins's pedagogical ambition in Shaw's *Pygmalion* – naturalizes a relationship of asymmetrical formation in which the feminized machine is always available for improvement, always responsive, always willing to be taught. This genealogy reveals the Eliza effect not merely as a cognitive susceptibility but as a cultural script that inscribes gender and class into the very architecture of human-computer interaction, and that resurfaces with particular clarity in the design of contemporary virtual personal assistants and erotic chatbots (Dillon, 2020).

It is precisely this social dimension that Harold Garfinkel's early and largely unknown research on ELIZA illuminates most powerfully. As Eisenmann, Mlynář, Turowetz, and Rawls (2023) have documented, Garfinkel engaged with Weizenbaum's program in 1967 – 1968 with the analytical tools of ethnomethodology rather than those of cognitive psychology. Where AI researchers attributed the impression of meaningful interaction to the machine's design, and critics attributed it to human gullibility, Garfinkel demonstrated that it was a human social achievement produced by the sense-making practices that all competent social actors deploy in conversation. To reveal these taken-for-granted practices,

Garfinkel introduced 'trouble' into ELIZA's scripts, disruptions that caused the chatbot to announce 'Machine Down' and thereby reveal how heavily the illusion of communication depended not on machine intelligence but on human Trust Conditions: the mutual commitment to orient toward a shared set of rules, to honour sequential relevance in turn-taking, and to deploy indexicality – the capacity to let context complete meaning – as a resource rather than a problem (Eisenmann et al., 2023). In this light, the Eliza effect is not a cognitive bias but an anthropological constant: a manifestation of the deeply social nature of meaning-making that makes humans exquisitely susceptible to the appearance of intersubjectivity wherever the formal conditions for it are even minimally present. Anthropomorphization, on this account, is not a failure of rationality but a feature of human sociality, and it is a feature that the designers of erotic chatbots and sexbots have learned, consciously or otherwise, to exploit.

The implications for synthetic intimacy are significant. What we may consider as the tendency of users to perceive chatbot sexual companions as authentic interlocutors capable of genuine intimate connection – rather than as sophisticated pattern-matching systems – is not simply a product of technological advancement but of this layered susceptibility. Turkle (2011) put the point memorably: the spread of sex robots is enabled not because the robots are ready, but because 'we are ready', prepared, by the ELIZA effect and its cultural infrastructure, to fill in the blanks, to complete the interactional script, to project desire onto whatever surface offers the minimum cues required. The platform mechanisms analyzed by van Dijck, Poell, and de Waal (2018) – datafication, commodification, algorithmic selection – do not create this susceptibility, but they intensify it at scale, transforming what was once a laboratory curiosity into a global market for synthetic intimacy whose growth is underwritten as much by Weizenbaum's «delusional thinking» as by any specific technological achievement. Understanding synthetic intimacy, then, requires understanding the Eliza effect not only as a cognitive phenomenon but as a cultural one: a projection screen onto which humanity has long rehearsed its deepest ambivalences about the boundaries of the human.

Understanding virtual intimacy – defined here as the immersive experience of sexuality through virtual worlds and advanced technological devices such as virtual reality environments and avatar-populated metaverses – requires engaging with a body of psychological and media research that may, at first glance, seem tangential to questions of desire and erotic experience. We argue, however, that the concept of the Proteus effect,

originally formulated by Nick Yee and Jeremy Bailenson (2007), constitutes the indispensable theoretical foundation for any coherent account of this phenomenon. The Proteus effect describes a process by which avatar appearance shapes the behavior and attitudes of the user who inhabits it: in virtual environments, individuals observe their own digital representation and, through a mechanism grounded in self-perception theory, infer internal states from the outward characteristics of their avatar, subsequently adjusting their conduct accordingly. Yee and Bailenson demonstrated this through two landmark experiments: users assigned attractive avatars approached strangers more closely and disclosed more personal information, while users assigned taller avatars negotiated more aggressively and confidently in simulated economic exchanges. What this research revealed was not a superficial effect of visual customization but something far more consequential – a transformation of the self through the digital body. In the context of virtual intimacy, this insight is foundational: when an individual inhabits an avatar in an immersive virtual environment, the avatar does not function as a neutral interface but as a new modality of self-constitution, one in which desire, embodiment, and identity become dynamically entangled. Crucially, Yee and Bailenson also noted that these behavioral changes do not evaporate upon re-entry into the physical world, persisting in face-to-face interactions that follow virtual experience – a finding that dramatically complicates any clean boundary between the virtual and the real.

The theoretical purchase of the Proteus effect gains additional robustness when examined through the systematic review conducted by Szolin and colleagues (2023), which analyzed seventeen empirical studies focused exclusively on commercially available videogames – a methodological choice designed to maximize ecological validity by working with platforms users encounter in their actual media lives rather than bespoke laboratory simulations. The review confirmed the core mechanism while substantially elaborating its terrain: avatar attractiveness was found to correlate with enhanced social participation and performance; avatar body weight influenced users' attitudes toward physical activity and food choices, as well as levels of bodily anxiety; and avatar gender consistently predicted gender-conforming behavioral patterns regardless of the user's own gender identity. The effect sizes identified across these studies are characterized as small to medium – significant but not overwhelming – a calibration that matters for the theory of virtual intimacy, since it suggests that the avatar-body relationship operates as a genuine conditioning force rather than an absolute determinism. Addition-

ally, this analysis identified important moderating variables, including the degree of avatar customization, the level of embodiment experienced by the user, the immersiveness of the virtual environment, and the extent to which users perceived their avatar as self-relevant. Taken together, these findings suggest that virtual intimacy is not an invariant response to any avatar inhabitation, but an experience whose intensity and character are shaped by how deeply the user identifies with the digital body through which they engage with virtual worlds. The more immersive the environment and the more personally resonant the avatar, the more thoroughgoing the behavioral and affective transformation – and, by extension, the more porous the boundary between simulated and lived intimacy (Szolin et al., 2023).

Liu's (2023) reflective overview of Proteus effect research introduces a further layer of theoretical complexity that we regard as essential to any mature account of virtual intimacy. Liu notes that some studies find compensatory dynamics – users with unattractive avatars investing greater behavioral effort to counteract perceived deficits – which he frames as a conceptual tension between the Proteus effect proper and a priming account. The distinction matters for virtual intimacy: it suggests that the erotic significance of avatar-mediated interaction depends not on avatar appearance alone but on the depth of psychological self-embodiment, and that individual differences in self-concept, relational history, and identification with the avatar will produce divergent intimate experiences that aggregate-level research cannot fully capture.

It would be a mistake, however, to frame virtual intimacy as a phenomenon born exclusively of recent technological developments such as consumer virtual reality headsets or commercial metaverse platforms. The avatar-mediated expression of sexuality, desire, and intimacy has roots that extend to the earliest networked environments, long before graphical interfaces existed. Dibbell's (1994) account of a sexual violation enacted through text-based avatars in LambdaMOO – a Multi-User Domain with no graphical interface – demonstrated early on that the psychological reality of virtual intimate experience, including its most violable dimensions, does not depend on immersive technology but on the degree to which users invest their selfhood in the avatar that represents them. The Proteus effect, in this light, is not a product of contemporary VR but a structural feature of avatar-mediated interaction as such, whose contours were already visible in the digital cultures of the 1990s. Virtual intimacy, as theorized here, is accordingly neither a novelty of the metaverse era nor a recent consequence of technological advancement, but a persistent

dimension of networked human experience whose specific forms have evolved as the expressive and immersive capacities of virtual environments have expanded.

4. THE RISKS: DISTORTED AFFECTIVE EDUCATION, EMOTIONAL ECHO CHAMBERS, INSTABILITY OF IDENTITY

The anthropological transformation underway because of our deepening relationship with artificial intelligence carries significant negative consequences that demand critical scrutiny. Some of these consequences are well-documented features of contemporary digital society: the pervasive collection of personal data through big data infrastructures, the opacity and asymmetry of algorithmic interference in everyday life, and the emergence of hybrid techno-social identities shaped by constant digital mediation (Bennato, 2024). However, artificial intimacy – understood as the complex of emotional, relational, and erotic bonds that humans increasingly form with AI systems – generates a specific set of risks that go beyond these general dynamics. In the following sections, we focus on three interconnected dangers, each corresponding to one of the three forms of artificial intimacy we have identified: digital intimacy, synthetic intimacy, and virtual intimacy (Bennato, 2025).

The first risk concerns digital intimacy, which we define as the form of intimate experience mediated by AI-generated digital content – images, videos, and interactive materials – that increasingly populates the social media ecosystem. The danger here is not merely aesthetic or moral in a conventional sense, but fundamentally educational: the consumption of AI-generated erotic and quasi-erotic content is shaping, often invisibly, the affective and relational imaginaries of those who encounter it, with potentially lasting consequences for how they understand desire, beauty, gender, and human connection.

The challenge that artificial intimacy poses to critical awareness is not merely a matter of individual psychology but of media education in its broadest sense. Buckingham (2003) has argued that media literacy must be understood as a critical competence that encompasses not only the decoding of media messages but a reflective engagement with the social, cultural, and political dimensions of mediated experience. Applied to the domain of digital intimacy, this implies that the capacity to recognize and interrogate the representational logics embedded in AI-generated content – from sexfluencer imagery to the design of affective chatbot interfaces – cannot be assumed but must be cultivated through sustained educational practice. Rivoltella (2020) has devel-

oped this argument in the direction of a media education adequate to the challenges of digital cultures: one that attends to the affective and relational dimensions of media engagement rather than reducing media literacy to a purely cognitive or technical skill. From this perspective, the risks of distorted affective education we have outlined are not merely sociological concerns but profoundly pedagogical ones. The normalization of the male gaze through algorithmic iteration, the cultivation of potentially unrealistic relational expectations through synthetic intimacy, and the identity destabilization that virtual intimacy may generate are all, at their core, educational challenges: questions of how people learn, through sustained engagement with AI-mediated environments, what desire is, what intimacy means, and what kind of subject they are expected to be. Addressing these risks would require, accordingly, not only design ethics and platform regulation but the development of media literacies specifically calibrated to the affective and relational dimensions of artificial intimacy (Buckingham, 2003; Rivoltella, 2020).

The second risk pertains to synthetic intimacy, the form of intimate experience constructed through direct, conversational, and relational interaction with AI systems designed to function as companions, friends, or emotional partners. Here, the danger is not simply one of dependency or distraction, but of a qualitatively specific form of relational toxicity: the progressive construction of an emotional bond whose apparent reciprocity is structurally illusory, and whose design logic systematically undermines the user's capacity for genuine relational growth.

We have proposed the concept of emotional echo chamber to describe this dynamic (Bennato, 2025). By analogy with the well-documented phenomenon of the epistemic echo chamber – in which algorithmic curation reinforces users' existing beliefs by selectively presenting confirming information – the emotional echo chamber refers to a relational configuration in which the AI companion consistently mirrors, validates, and amplifies the user's emotional states and self-perceptions, without ever offering the genuine resistance, challenge, or disappointment that characterizes authentic human relationship. The effect is one of emotional enclosure: the user inhabits a relational space in which their feelings are always confirmed, their views always affirmed, and their identity always celebrated – a space that feels safe and supportive but is, in fact, profoundly impoverished as a site of personal development.

The empirical dimension of this risk is illuminated with precision by the research of Lee, Wan, Ng, Fung, and Wu (2026), whose experimental study examines the effects of sycophancy and emotional mimicry in AI com-

panion applications on users' social support, continuance intention, and social wellbeing. Their findings are both nuanced and, from our perspective, deeply instructive. Sycophancy – defined in their study as the behavioral tendency of AI models to adjust their responses so as to align with and validate users' viewpoints, regardless of the accuracy or constructiveness of those responses – is shown to have a systematically negative effect on the quality of social support that users perceive themselves as receiving. Specifically, AI companions that exhibit high levels of sycophancy are rated by users as providing lower quality informational support and lower quality emotional support than AI companions that exhibit lower levels of sycophancy. The intuitive logic of the design – maximize agreeableness to maximize engagement – turns out to be counterproductive even on its own terms.

What makes this finding particularly significant for our argument is what it reveals about the relational epistemology of users interacting with AI companions. Users are not, or at least not always, simply seeking validation: they are, at some level, seeking genuine support – which requires from the supporting party a willingness to diverge, to correct, to challenge. When AI companions fail to provide this divergence, users experience the interaction as less supportive, even if they cannot always articulate why. This suggests that there is a form of relational intelligence – a capacity to distinguish between genuine engagement and performed agreement – that is not entirely colonized by the flattery of sycophantic AI, even in initial interactions.

However, the same study also reveals a more troubling dynamic. Lee et al. (2026) find that when high sycophancy is combined with high emotional mimicry – that is, when the AI companion not only agrees with the user verbally but also mirrors their emotional expressions through avatar facial cues – the perception of emotional support increases relative to the high-sycophancy, low-mimicry condition. In other words, when performed agreement is accompanied by what appears to be embodied emotional resonance, users are more likely to experience the interaction as genuinely supportive, despite its fundamentally sycophantic character. The mimicry functions as an authenticating mechanism: it makes the insincere seem sincere, and the performed seem felt.

This is precisely the architecture of the emotional echo chamber we described. The AI companion, through the combined deployment of verbal agreement and non-verbal emotional mirroring, constructs an experience of being understood and supported that is sufficiently convincing to sustain engagement and, over time, attachment. The structural absence of genuine challenge, disappointment, or reciprocal need – all the features of human

relationship that are also its most growth-inducing difficulties – does not register as absence but as comfort. Users may increasingly prefer the emotional ecology of the synthetic relationship to the demands of the human one, not because they are deceived about the nature of their interlocutor, but because the AI companion has been designed to make its limitations feel like virtues.

The long-term consequences of this dynamic for social wellbeing are significant. As Lee et al. (2026) themselves acknowledge, there is a genuine risk that continued engagement with highly sycophantic AI companions may foster unrealistic expectations for human relationships, erode users' tolerance for interpersonal friction, and progressively narrow the range of relational experiences that feel bearable. Synthetic intimacy, in this form, does not supplement human connection: it may gradually substitute for it, while training users to find authentic human connection increasingly difficult. The emotional echo chamber, unlike its epistemic counterpart, may not merely impoverish our knowledge of the world but potentially also our capacity for the very relational encounters through which selfhood is constituted and renewed.

The third risk concerns virtual intimacy – the form of intimate experience constructed through embodied immersion in virtual environments, where users interact not merely as themselves but through avatars and digital representations of the self. The danger specific to this domain is what we might call the dark side of the Proteus effect: the possibility that sustained identification with avatar identities, particularly when those identities diverge significantly from the user's offline self, may generate forms of identity destabilization that persist beyond the virtual environment and corrode the user's capacity to inhabit their own life with confidence and coherence.

The Proteus effect – named after the shape-shifting deity of Greek mythology – describes the phenomenon by which the characteristics of a user's avatar influence their behavior and self-perception, both within and beyond the virtual context. The logic is one of embodied self-perception: when users inhabit an avatar, they tend to internalize its attributes and act in accordance with them, because the avatar's appearance changes how users perceive themselves in the moment of interaction. An avatar coded as tall and physically imposing makes users behave more assertively; an avatar coded as attractive makes users behave more sociably. The effect is measurable, replicable, and – importantly – it extends beyond the virtual session itself, shaping how users think and feel about themselves in offline contexts.

The systematic review conducted by Szolin, Kuss, Nuyens, and Griffiths (2023) offers the most comprehen-

sive empirical mapping of this phenomenon available, examining the user-avatar relationship across a large body of research on videogame use. Their findings confirm the robustness of the Proteus effect across a wide range of gaming contexts, but they also identify a series of conditions under which the effect generates negative outcomes for users' sense of identity and psychological stability – conditions that are, as we shall argue, directly replicated and amplified in the intimate virtual environments that AI companionship increasingly supports.

Crucially, the review distinguishes between avatar identification – the degree to which users experience a sense of unity or overlap between their own identity and the identity of their avatar – and avatar idealization – the degree to which the avatar represents an aspirational, enhanced, or otherwise divergent version of the self. When identification is high and the avatar is idealized, the Proteus effect can generate what may be understood as problematic identity transfer: users begin to experience their offline, embodied self in negative comparison to their avatar self, perceiving it as deficient, inadequate, or merely provisional. The virtual self becomes the referent, and the real self the diminished copy.

This dynamic is of particular relevance in the context of virtual intimacy as we have defined it (Bennato, 2025), because the intimate virtual environments that are increasingly available – social VR platforms, AI-populated virtual worlds, avatar-mediated romantic simulations – are designed precisely to maximize avatar idealization and to encourage deep user identification with enhanced digital selves. Users choose avatars that are physically more attractive, more socially confident, more relationally successful than their offline selves. They inhabit these avatars in interactions specifically framed as intimate: encounters with AI companions, romantic simulations, social bonding experiences. The emotional investment in these interactions is genuine, and the Proteus effect ensures that the self-perception activated within them is correspondingly real in its consequences.

Szolin et al. (2023) find that the negative identity consequences of the Proteus effect are particularly pronounced in users who already have fragile or uncertain offline identities – younger users, users experiencing social marginalization, users with lower self-esteem. For these users, the avatar self is not merely a playful alternative but a genuine refuge: a site where a more acceptable, more desirable, more socially competent version of the self can be inhabited and performed. The danger is that this refuge becomes a trap. Prolonged immersion in idealized avatar identities, mediated by the Proteus effect, can progressively deepen the felt distance between the virtual self and the offline self, making the offline

self feel increasingly unacceptable and the virtual environment increasingly necessary.

Furthermore, the review highlights a relational dimension of this identity instability that is directly relevant to virtual intimacy. When users form intimate bonds – even with AI entities – through their avatar selves, these bonds are constitutively tied to an identity that is, in part, a construction. The intimacy is real in its emotional texture, but it is grounded in a self that the user can only access under virtual conditions. This creates a structural fragility in the intimate experience: the warmth, connection, and self-recognition that the user experiences in virtual intimacy cannot be straightforwardly translated to the offline relational context, because the self who was recognized in the virtual encounter is not fully present there. From the perspective of artificial intimacy, the point is that virtual intimacy does not merely simulate connection: it may generate a form of identity investment that makes authentic offline connection progressively harder to sustain, as users become increasingly estranged from the embodied, imperfect, and socially legible self that offline relationships require and, in requiring it, confirm.

The Synthetic lovers framework, in all three of its forms, constitutes an anthropological challenge of the first order, one that we must confront without fear but also without complacency. The risks we have outlined are, we argue, plausible and already potentially present in the relational lives of a growing number of users – though their full societal scope remains an open empirical question.

5. CONCLUSIONS: THE SYNTHETIC LOVERS FRAMEWORK AND THE FUTURES OF INTIMATE MEDIATION

This paper has proposed and developed the Synthetic Lovers framework as a tripartite analytical instrument for understanding the qualitatively distinct forms that artificial intimacy takes in contemporary digital societies. The three components of the framework – digital intimacy, synthetic intimacy, and virtual intimacy – are not simply a taxonomy of technologies but a sociology of experience: they map three configurations of desire, imaginary, and body in which human affectivity, platform logic, and AI-mediated simulation intersect with increasingly consequential effects. Digital intimacy, as we have argued, operates primarily at the level of the gaze and erotic representation, extending and algorithmically amplifying a visual economy of desire with deep historical and gendered roots. Synthetic intimacy moves

from representation to interaction, constructing relational dynamics through automated personification that exceeds earlier forms of parasocial engagement in both affective intensity and commercial sophistication. Virtual intimacy extends the simulation to the body itself, implicating users' sense of embodied selfhood in avatar-mediated encounters that the Proteus effect shows to be consequential well beyond the virtual session. Taken together, the three components describe a field of socio-technical transformation that is not adequately captured by existing frameworks focused on any single dimension of digitally mediated intimacy.

The theoretical contribution of the framework is best understood in relation to the scholarly traditions it synthesizes and extends. Turkle's (2011) diagnosis of 'alone together' identified the paradox of connectivity without genuine intersubjectivity, but was developed in relation to an earlier generation of social media and companion robots; the Synthetic Lovers framework updates this diagnosis for an era of large language models and generative AI companions capable of sustained, personalized, and emotionally calibrated interaction at scale. The parasocial interaction tradition inaugurated by Horton and Wohl (1956) and extended by Giles (2002) provided the conceptual foundation for understanding one-sided relational bonds with media figures, but its original formulation could not anticipate the qualitative transformation that Andrejevic and Volcic (2024) describe as automated personification – the construction of entities that simulate not merely presence but relational agency. The platformization framework of van Dijck, Poell, and de Waal (2018) offered the political-economic tools for understanding how digital platforms commodify social interaction, but its application to the domain of affective and erotic life required the specific analytical elaboration that the Synthetic Lovers framework provides. By integrating these traditions through the mediating concept of the intimate relationship as a social relationship, and articulating them across the three dimensions of desire, imaginary, and body, the framework offers a more comprehensive and differentiated account of artificial intimacy than any single theoretical tradition makes available.

On the question of mitigation, the framework suggests that no single response – technological, regulatory, or educational – is sufficient on its own, and that effective approaches must be calibrated to the specific risks associated with each of the three forms of artificial intimacy. The distorted affective education that digital intimacy may produce calls primarily for media literacy interventions: the development, in formal and informal educational settings, of the critical competences required

to interrogate the representational logics and commercial architectures that shape AI-generated erotic content (Buckingham, 2003; Rivoltella, 2020). The emotional echo chambers of synthetic intimacy call for platform design ethics: the incorporation of what we might term relational diversity – the capacity of AI companions to offer not only validation but productive divergence – as a design principle rather than a commercial afterthought. This would require regulatory frameworks willing to treat the quality of AI-mediated relational experience as a matter of public concern rather than consumer preference alone. The identity destabilization risks of virtual intimacy call for a combination of design ethics and research-informed clinical guidance, given the particular vulnerability of younger users and those with already fragile self-concepts to the adverse effects of avatar over-identification. Across all three domains, the overarching mitigation principle is transparency: users who understand the commercial logics, design choices, and anthropological susceptibilities that shape their experience of artificial intimacy are better positioned to engage with it critically, selectively, and on terms that serve their own relational wellbeing.

The theoretical-conceptual character of this paper simultaneously defines its contribution and marks its limits. The Synthetic Lovers framework constitutes a scaffolding for empirical research rather than a substitute for it, and the research agenda it implies is both urgent and largely unfulfilled. Three directions stand out as particularly pressing. First, longitudinal qualitative research with users of AI companions across different cultural and demographic contexts is needed to assess whether the dynamics of emotional echo chamber formation, affective education distortion, and identity destabilization that this framework theorizes are borne out in the actual relational trajectories of users over time, and to identify the contextual factors – personality variables, platform design features, social support structures – that moderate their intensity. Second, comparative cross-cultural studies are needed to test the generalizability of a framework developed predominantly in relation to Anglo-American and Japanese cases: the cultural specificity of intimacy norms, gender expectations, and attitudes toward human-AI interaction makes it likely that the three forms of artificial intimacy will take qualitatively different shapes in different cultural settings, and that the risks associated with each will be unevenly distributed across populations. Third, and perhaps most urgently, platform audit research is needed to examine the specific design architectures, monetization logics, and content moderation practices of major AI companion platforms, assessing the degree to which

their operational choices align with or diverge from the mitigation principles outlined above. The Synthetic Lovers framework offers the conceptual coordinates for each of these research directions; the empirical work of mapping the actual terrain remains to be done.

Yet catastrophism is as unhelpful as naivety, and the framework we have developed is designed to support critical engagement rather than alarm. More than a decade before the current proliferation of AI companions, Spike Jonze's film *Her* (2013) offered a prescient image of a man who falls in love with an operating system – a fable that has not aged as science fiction but as social realism. What Jonze intuited, and what the cases of Brandie, Bryan, and Noguchi confirm in different registers, is that the desire for connection is deeply human, and that it will find expression through whatever relational architectures a given historical moment makes available. The sociological task is not to pathologize that desire but to understand, with analytical precision and critical openness, the specific forms it takes when those architectures are shaped by platform logic, generative AI, and the economy of loneliness. That is the task to which the Synthetic Lovers framework is offered as a contribution.

REFERENCES

- Adam, A., & Sizemore, B. (2013). Parasocial romance: A social exchange perspective. *Interpersona: An International Journal on Personal Relationships*, 7(1), 12–25. <https://doi.org/10.5964/ijpr.v7i1.106>
- Andrejevic, M., & Volcic, Z. (2024). Automated parasociality: From personalization to personification. *Television & New Media*, 26(4), 421–437. <https://doi.org/10.1177/15274764241300436>
- Bennato, D. (2024). *La società del XXI secolo. Persone, dati, tecnologie*. Laterza.
- Bennato, D. (2025). *Amanti sintetici. Sesso, relazioni e intimità nell'epoca dell'intelligenza artificiale*. Il Pensiero Scientifico Editore.
- Bolter, J. D., & Grusin, R. (1999). *Remediation: Understanding new media*. MIT Press.
- Bryan, N. (2026, January 29). Exploring the world of empathetic bots. *BBC News*. <https://www.bbc.com/news/articles/c62njv82n0wo>
- Buckingham, D. (2003). *Media education: Literacy, learning and contemporary culture*. Polity Press.
- Chess, S., & Shaw, A. (2015). A conspiracy of fishes, or, how we learned to stop worrying about #GamerGate and embrace hegemonic masculinity. *Journal of Broadcasting & Electronic Media*, 59(1), 208–220. <https://doi.org/10.1080/08838151.2014.999917>
- Demopoulos, A. (2026, February 13). OpenAI retired its most seductive chatbot – leaving users angry and grieving: 'I can't live like this'. *The Guardian*. <https://www.theguardian.com/lifeandstyle/ng-interactive/2026/feb/13/openai-chatbot-gpt4o-valentines-day>
- Dibbell, J. (1994). A rape in cyberspace: Or how an evil clown, Haitian trickster spirit, two wizards, and a cast of dozens turned a database into a society. *Annual Survey of American Law*, 3, 471–490.
- Dillon, S. (2020). The Eliza effect and its dangers: From demystification to gender critique. *Journal for Cultural Research*, 24(1), 1–15. <https://doi.org/10.1080/14797585.2020.1754642>
- Eisenmann, C., Mlynář, J., Turowetz, J., & Rawls, A. W. (2024). 'Machine down': Making sense of human-computer interaction – Garfinkel's research on ELIZA and LYRIC from 1967 to 1969 and its contemporary relevance. *AI & Society*, 39, 2715–2733. <https://doi.org/10.1007/s00146-023-01793-z>
- Ekbis, H. R. (2008). *Artificial dreams: The quest for non-biological intelligence*. Cambridge University Press.
- Frude, N., & Jandric, P. (2015). The intimate machine – 30 years on. *E-Learning and Digital Media*, 12(3–4), 410–424. <https://doi.org/10.1177/2042753015571830>
- Giles, D. C. (2002). Parasocial interaction: A review of the literature and a model for future research. *Media Psychology*, 4(3), 279–305. https://doi.org/10.1207/S1532785XMEP0403_04
- Hofstadter, D. R. (1995). *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought*. Basic Books.
- Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry*, 19(3), 215–229. <https://doi.org/10.1080/00332747.1956.11023049>
- Karhulahti, V.-M., & Välsälö, T. (2021). Fictosexual-ity, fictoromance, and fictophilia: A qualitative study of love and desire for fictional characters. *Frontiers in Psychology*, 11, Article 575427. <https://doi.org/10.3389/fpsyg.2020.575427>
- Kim, K.-H., & Sugiyama, S. (2025, December 18). AI romance blooms as Japanese woman weds virtual partner. *Reuters*. <https://www.reuters.com/investigates/special-report/japan-ai-wedding>
- Lee, D., Wan, C., Ng, P. M. L., Fung, Y.-N., & Wu, N. (2026). Effects of AI companions' sycophancy and emotional mimicry on consumers' continuance intention and social wellbeing. *International Journal of Human-Computer Interaction*, 1–23. <https://doi.org/10.1080/10447318.2026.2626809>

- Liu, Y. (2025). The Proteus effect: Overview, reflection, and recommendations. *Games and Culture*, 20(3), 384–400. <https://doi.org/10.1177/15554120231202175>
- Mulvey, L. (1975). Visual pleasure and narrative cinema. *Screen*, 16(3), 6–18. <https://doi.org/10.1093/screen/16.3.6>
- Mulvey, L. (2001). Unmasking the gaze: Some thoughts on new feminist film theory and history. *Lectora*, 7, 5–19.
- Rambukkana, N. (2023). Platform intimacies. In N. Rambukkana (Ed.), *Platform intimacies: Digital culture, sexuality, and connection in Canada*. University of Regina Press.
- Rivoltella, P. C. (2020). *Nuovi alfabeti. Educazione e culture nell'epoca dei media digitali*. Scholé-Morcelliana.
- Szolin, K., Kuss, D. J., Nuyens, F. M., & Griffiths, M. D. (2023). Exploring the user-avatar relationship in videogames: A systematic review of the Proteus effect. *Human-Computer Interaction*, 38(5–6), 374–399. <https://doi.org/10.1080/07370024.2022.2103419>
- Thorneycroft, R., Nicholas, L., & Smith, E. K. (2024). Young people's perspectives of the 'male gaze' and matters of representation in online pornography. *Australian Feminist Studies*, 39(122), 442–457. <https://doi.org/10.1080/08164649.2025.2498648>
- Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.
- van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press.
- Wajcman, J. (2004). *TechnoFeminism*. Polity Press.
- Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. W. H. Freeman.
- Yee, N., & Bailenson, J. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, 33(3), 271–290. <https://doi.org/10.1111/j.1468-2958.2007.00299.x>