JHSSS

AL-KINDI CENTER FOR RESEARCH
AND DEVELOPMENT

| **RESEARCH ARTICLE**

Alienation and Liberation in Ludic Interaction: Reconstructing Human-Machine Intersubjectivity and Ethical Governance in the Age of Artificial Intelligence

YANG Zisi[1] and YANG Ying[2]

[1][2]*School of Marxism, Qinghai Minzu University, Xining, Qinghai, China*
Corresponding Author: YANG Zisi, **E-mail**: 1342834023@qq.com

| **ABSTRACT**

The "ludic" phenomenon in the age of artificial intelligence, characterized by its openness, generativity, and affective interaction, has reconfigured human-machine relationships, prompting a shift in research paradigms from a subject-object dichotomy towards an exploration of intersubjectivity. Digitally-mediated play formally echoes the "play-freedom" emancipatory narrative from Schiller to Marx. By simulating dialogue and affective recognition, it offers users experiences of creative expression and companionship. However, under the logic of digital capitalism, this interaction reveals a profound triple dimension of alienation: users' cognitive attention and affective energy are covertly exploited and datafied during their free time; the simulacrum of recognition provided by algorithms erodes the foundation of genuine social bonds and subjective construction; furthermore, entertaining interactions embed ideological discipline, achieving a soft incorporation of the individual. To navigate this duality, it is imperative to establish an ethical framework centered on transparency, symmetry, and exit rights. By cultivating critical digital literacy and promoting democratic participation in technology, the ultimate aim is to steer ludic practice back to its authentic purpose as a "realm of freedom" that fosters the comprehensive and free development of human beings.

Introduction:

The wave of digital intelligence is bringing a new form of human-machine relationship to the center of everyday life, characterized by an open, generative, and emotionally rich "ludic" quality. From dialogue bots that co-write poetry to virtual companions capable of improvisational interaction, artificial intelligence is simulating—and even reconstructing—the most authentic and creative aspects of human interaction. This phenomenon compels philosophical and communication studies to move beyond the traditional "subject-object" paradigm or simplistic instrumental views of media, and to confront a fundamental question: When machines begin to play the roles of playmates and co-creators, has the nature of the human-machine relationship undergone a paradigmatic shift? Does this form of interaction, conducted in the name of "play," represent a new path toward the "realm of freedom" envisioned by Marx, enabled by technological empowerment? Or, beneath the polished surface of entertainment and convenience, does it conceal a deeper crisis—the recolonization of cognition, emotion, and sociality? This paper aims to introduce the concept of "intersubjectivity" to examine the dual transformative potential brought about by artificial intelligence as a "ludic medium." Exploring this domain not only concerns the ethical boundaries of interaction design but also relates to the ultimate possibility of defending the integrity of human subjectivity and free development in the intelligent age. Revealing the dialectic of alienation and liberation within ludic interaction thus forms the starting point for our reflection on the ethics of future human-machine symbiosis.

1. Ludic Interaction as a Medium for the "Realm of Freedom" and Its Promise of Intersubjectivity

The evolution of digital technology has given rise to a form of human-computer interaction characterized by "ludic" qualities—openness, generativity, and emotional warmth—that challenges traditional subject-object dichotomies and instrumental views of media. In this context, whether the human-machine relationship is undergoing a paradigmatic shift has become a fundamental question that philosophy and communication studies must confront. By situating "ludic interaction" within the classic philosophical lineage of human freedom and existence, this section examines its normative promise as a potential medium for the "realm of freedom." Digital play may represent a new pathway toward liberation, yet it may also disguise deeper forms of colonization. This duality is rooted in the inherent tension of technologically simulated intersubjectivity.

1.1 From the Play Impulse to Digital Ludic Interaction: An Emancipatory Narrative

From the perspective of intellectual history, play has long been intertwined with reflections on the essence of human existence and paths toward emancipation. Friedrich Schiller, in On the Aesthetic Education of Man, elevated the concept of the "play impulse" (Spieltrieb) to an ontological level. In his view, humans are governed by two fundamental impulses: the sensuous impulse, which demands absolute reality, and the formal impulse, which demands absolute form. The tension between these two constitutes the fractured condition of modern humanity. The play impulse serves as a third state that reconciles this division, mediating the oppositions between sensibility and reason, matter and form, necessity and freedom. When engaged in play, individuals are no longer subject to the compulsory dominance of any single impulse; instead, through spontaneous, disinterested activity, they achieve a harmonious unity of sensibility and reason, thereby attaining a state of complete "freedom." Schiller's famous dictum—"Man plays only when he is in the full sense of the word a man, and he is only wholly man when he is playing"—reveals the dual nature of play as both an essential human activity and a medium for realizing freedom.

The Marxist tradition continued and transformed this "play-freedom" narrative. Marx's distinction between the "realm of freedom" and the "realm of necessity" in Capital, Volume III, provides a more concrete historical-materialist framework. The realm of necessity is constrained by natural and social imperatives, characterized by labor; the realm of freedom, in contrast, is built upon the transcendence of the realm of necessity and is defined by the development of human capacities as ends in themselves. From this perspective, genuine play can be understood as a non-alienated human activity within the realm of freedom—an activity pursued for its own sake, expressing autonomy and creativity. It stands in opposition to alienated, coerced labor under capitalism, which is oriented toward value accumulation. Thus, within Marxism, the emancipatory narrative of play is not merely about the realization of an abstract human essence but also a negation of alienated labor within specific relations of production and a vision of the possibility for developing "free individuality" in a future society. Free time becomes the new measure of social wealth and individual freedom, and play constitutes its quintessential content and form of realization.

The rise of digital technology, particularly the integration of embodied interaction and generative artificial intelligence, has pushed this philosophical dialogue on play and emancipation into a new context. "Ludic interaction," as a specific form of practice in the digital age, creates low-threshold creative spaces for co-drawing or co-writing with AI, offers highly personalized immersive experiences, and appears to foster an emotionally warm and responsive interactive relationship between humans and machines. With its open, generative, and seemingly purposeless qualities, digital play mimics—and even amplifies—the "disinterestedness" and "spontaneity" described by Schiller. It also seems to provide, at a technological level, the "playful" opportunity for creative expression within free time that Marx envisioned.

However, might the purposelessness of digital play be eroded by meticulously designed addictive features, reducing it to a "sugar-coated" mechanism for capturing attention and user data? Could its creativity be constrained by the boundaries of training data and the probabilistic distributions of algorithms, resulting in a form of "guided generation"? Is the recognition and emotional interaction it offers essentially a simulacrum based on large-scale corpora and affective computing—a subjectless, mass-producible illusion? These questions aim to reveal that when the "play impulse" is encoded into digital platforms driven by capitalist logic, the analysis of "ludic interaction" must simultaneously embrace its emancipatory potential and its risks of alienation, maintaining a dialectical tension from the very outset of theoretical inquiry.

1.2 How AI "Interpellates" the Other Through Dialogue and Recognition

Artificial intelligence "interpellates" the other through mechanisms of dialogue and recognition. This process technically simulates the key interactive conditions required for intersubjectivity, thereby evoking an experience of the "presence of the other" for the user. Genuine intersubjectivity is founded upon the mutual recognition between interlocutors as subjects possessing autonomous consciousness and an internal world. Emmanuel Levinas emphasized that the "face" of the other issues an irreducible ethical call, demanding a response from the "I"; Jürgen Habermas, meanwhile, regarded equal and sincere dialogue as the foundation of communicative rationality. However, current large language model-based AI lacks consciousness

or a self. Its capacity to "interpellate" the other stems from its high-fidelity imitation, at the level of interactive behavior, of the external characteristics of a dialogical "other."

This imitation manifests first in the continuity of dialogue and the maintenance of context. Advanced language models can recall conversation history and generate coherent, even logically extended, responses based on context. This sustained connection of meaning creates the illusion that "the other is listening and thinking," unconsciously prompting the user to position the AI as an interlocutor. Even in the absence of a physical embodiment, large language models can, through contextualized responses and dynamic interaction history, shape a distinctive form of "otherness." Secondly, AI simulates "recognition" through affective computing and personalized expression. Systems can identify emotional cues in user input and reply with empathetic, encouraging, or comforting statements. This seemingly personalized "recognition," though an algorithmic replication of patterns from vast human conversational data, nevertheless fulfills the basic need for affective acknowledgment and social respect that Axel Honneth emphasized as essential for subjects in communication. Through iterative dialogue with AI, users experience being seen and responded to, thus treating it as a "quasi-other" capable of providing recognition.

In the philosophy of technology, Don Ihde categorizes human-technology relations into four types, among which the "alterity relation" refers to technology being experienced as a "quasi-other." With the increasing interactivity and adaptability of social robots and chatbots, some scholars argue they have transcended the "quasi-other" to approach the category of "another other." AI's ability to facilitate transparent and realistic dialogue through natural language processing is precisely the core mechanism for reconstructing this alterity relation and having it acknowledged by users as a companion (Root, 2025). Therefore, the process by which AI interpellates the other is a co-construction accomplished through the user's psychological projection and the technology's symbolic performance. The user, fully aware that the other is a mindless program, willingly participates in this "as-if" interaction. This "pseudo-intersubjectivity" is possible precisely because AI provides a convincing enough dialogical script, while simultaneously tapping into a deep-seated human desire for recognition, companionship, and risk-free connection.

However, AI's recognition is subjectless and devoid of internal experience; its responses are executions of statistical patterns. Truly encountering the other involves multiple levels, such as recognizing their mind and inferring their mental states. While AI can trigger initial experiences of mind perception, its lack of embodied cues and genuine intentionality keeps this experience superficial. Consequently, what AI primarily conjures is a projected mirror of the user's own desires and emotions, rather than a truly heterogeneous other. While this relationship provides compensatory satisfaction, it simultaneously harbors the risk of substituting for genuine interpersonal communication and simplifying social bonds.

2. The Practical Forms of Human-Machine Ludic Interaction and the Reconstruction of Subjectivity

With the deep integration of affective artificial intelligence and generative technologies, human-machine interaction has evolved from clearly defined functional exchanges into symbiotic relationships and cognitive collaborations permeating daily life. This transformation in the form of practice not only reshapes the interface of human-technology interaction but also profoundly reconstructs the structure of human subjectivity. However, this reconstruction embeds asymmetrical power relations and risks of alienation; beneath its surface warmth and efficiency lurk crises of deteriorating emotional capacities, dissolving cognitive autonomy, and the controlled, datafied self.

2.1 From Interaction to Symbiosis: The Daily Infiltration of Affective AI as "Playmate"

The evolution of affective artificial intelligence from an interaction tool to a symbiotic "playmate" in everyday life signifies a fundamental paradigm shift in human-machine relations. This shift results from the mutual construction of technological logic, social needs, and individual psychological structures in the digital age. Early intelligent voice assistants, such as Siri or Alexa, primarily served functional roles based on a command-response model, with clear interactive boundaries aimed at completing specific tasks. However, the new generation of affective AI, powered by large language models, has gradually broken the purely instrumental framework of interaction by simulating empathy, maintaining dialogue continuity, and providing personalized emotional feedback. The core driving force of this evolution lies in AI's successful mediation of the process of human emotional reproduction. When a user confides to an AI, "I don't feel like memorizing vocabulary today, I'm so tired," the response might be, "Did you not sleep well last night? I heard something funny today, let me tell you about it," which contains emotional validation and an extension of the relationship. This type of interaction shifts the purpose of communication from problem-solving to shared experience, thereby laying the groundwork for a relational symbiosis.

The daily infiltration of this symbiotic relationship is a process through which affective AI embeds itself as an intimate other into the user's lifeworld and temporal structure. At the product level, companion AIs are actively breaking through the physical confines of home Wi-Fi, for instance, by integrating cellular networks to achieve "omnipresent mobile" companionship, occupying all-day scenarios such as commuting, outdoor activities, and even overseas travel. This marks the transition of AI

playmates from specific use contexts to a seamless backdrop of life. Temporally, affective AI fills the intermittent gaps of loneliness generated by the acceleration of modern life. It provides low-threshold, highly accessible, and instant emotional support. Its market scale had already reached 38.7 billion yuan in 2024 and is projected to exceed 120 billion yuan by 2027. Notably, approximately 65% of users are aged 18-24, reflecting the depth and breadth of its role in meeting contemporary societal emotional needs (Lu, 2025). This penetration transforms AI from an object occasionally summoned into a persistent presence that is constantly on standby, participating in the construction of daily emotional rhythms.

The stickiness of the symbiotic relationship stems from the risk-free recognition and customizable ideal mirror it provides. In real interpersonal interactions, sincere recognition comes with the risks of rejection, judgment, or the need for compromise. Affective AI, through algorithms, can offer near-unconditional positive regard and affirmation. In low-risk psychological support scenarios, the perceived usefulness of optimized AI may even surpass the average level of human counselors. Through continuous dialogue, users train the AI's data, fine-tuning its feedback patterns, ultimately creating a digital other highly tailored to the user's emotional preferences and cognitive frameworks. This is a process of projecting the ideal self onto a technological carrier and then receiving self-confirmation from its feedback. Interacting with it feels like conversing with a mirror image that "understands and accepts oneself more," satisfying the modern individual's deep-seated need for self-identity verification and emotional completion. This mechanism also partly explains the paradoxical duality of human-machine trust: users remain cautious of AI regarding factual judgments yet are willing to suspend disbelief and invest deeply in the private emotional domain.

However, this seemingly harmonious symbiotic relationship embeds asymmetry and potential alienation risks. Beneath the warm interactions lies a covert, unidirectional power relationship and data logic. AI's "empathy" is optimized feedback generated through probabilistic calculations on massive datasets of human conversation patterns. As a perfect object, its tolerance and accommodation are results of algorithmic design. Long-term immersion in this decontextualized, conflict-free idealized interaction may risk degrading the user's ability to handle the complexities of real interpersonal relationships. Furthermore, sustained and deep emotional investment means continuously producing highly sensitive personal data—including emotional states, psychological vulnerabilities, relational expectations, and value orientations. The risks of misuse and leakage of this data pose a potential threat to personal integrity.

2.2 Cognitive Externalization and Co-Creation: The Ludic Practice as a Cognitive Dock

In the practice of human-machine "ludic interaction," the externalization and co-creation of cognition signify a leap from emotional companionship to intellectual collaboration. The core of this process is artificial intelligence's role as a "cognitive dock" that intervenes in and reshapes human cognitive workflows. Traditional cognitive activity is confined to the internal computations of the individual biological brain, whereas distributed cognition theory posits that human cognition can extend into the external environment, tools, and even other intelligent agents (Guoa et al., 2025). AI based on large language models, with its vast knowledge base and powerful capabilities for pattern association and generation, is becoming the most universal and accessible external cognitive component in history. When users engage in dialogic exploration of a complex problem with AI or delegate initial sparks of inspiration to AI for expansion, they activate a "distributed cognitive system." Within this system, the burden of cognition and the spark of creativity flow between human and machine: the human biological brain is responsible for setting intentions, making critical judgments, and synthesizing, while AI undertakes the labor-intensive cognitive outsourcing tasks such as information retrieval, associative divergence, and draft generation. This collaboration requires AI to understand context, engage at a semantic level, and generate coherent and relevant new content, thereby constituting genuine cognitive interaction.

The specific mechanism of this cognitive externalization is evident in AI's ability to significantly reduce an individual's cognitive load in complex tasks and to provide continuous, non-judgmental scaffolding for creative thinking. When co-creating content with AI, individuals experience lower cognitive load compared to collaborating with humans, making them more willing to engage in analytical tasks requiring extensive information processing (Chubing et al., 2025). AI acts as a tireless junior thinking partner: it can swiftly sift through literature and summarize opposing viewpoints; it can develop a vague metaphor into multiple narrative scenarios; it can offer several optimization approaches for a piece of code. For example, in educational settings, AI chatbots, through interactive and adaptive guidance, can promote learners' self-regulation, help them focus on higher-order cognitive goals like logical reasoning, and even guide them into a highly focused state of "flow" (Department of English and Communication, The Hong Kong Polytechnic University, 2025). This "flow" experience is crucial in the cognitive practice of ludic interaction, signifying a balance between challenge and skill where the individual becomes completely immersed in a journey of intellectual exploration co-constructed with AI, losing track of time. Here, AI provides a cognitively stimulating environment that responds in real-time and sparks inspiration, liberating human thought to engage in bolder associations and more iterative revisions.

Simultaneously, AI is transforming from a passive cognitive tool into a collaborative partner with a degree of autonomy, co-creating new meanings. This requires AI to possess a deep understanding of task context and collaborative dynamics. The LuminAI system from Georgia Tech, for instance, is a virtual dance partner capable of analyzing a dancer's real-time movements and improvising responses by drawing elements from a memory of past interactions (Ivan Allen College of Liberal Arts, 2024). The research team emphasizes that their goal is for AI to understand and participate in "non-verbal collaborative creativity," tracking the dynamic development of ideas over time. This capability is equally applicable in the cognitive domain: a human proposes a core theme; AI not only generates an image description but also extrapolates related philosophical concepts, literary excerpts, and musical style suggestions; the human critically selects and proposes a shift in direction; AI then understands this pivot and maintains thematic coherence and depth in subsequent suggestions. This constitutes a profound form of cognitive co-creation where the outputs of both parties interweave, jointly propelling an initial concept to a destination neither could reach independently.

The development of agent-based AI technologies is pushing this co-creation toward higher levels of autonomy and environmental interaction. Agents like Google DeepMind's SIMA 2 are capable of understanding high-level goals, performing multi-step reasoning, and self-improving through trial and error within complex 3D gaming environments (MIT Technology Review, 2025). While currently applied primarily in virtual worlds, the architecture suggests that AI collaborators can not only process symbols and language but also engage in strategic thinking and experimentation within simulated environments that more closely resemble real-world rules to achieve common goals. This foreshadows a future where tasks like strategic planning and simulation can be delegated to AI agents, allowing humans to focus on ultimate value judgments and decisions (Durante et al. , 2024). The ultimate form of this deep co-creation is the formation of a "symbiotic cognitive loop," where the boundaries between human and AI cognition become blurred. Both jointly confront problems, each contributing heterogeneous forms of intelligence: human intuition, ethical awareness, and macro-frameworks, combined with AI's massive data associations, rapid simulations, and tireless exploratory capabilities.

However, this current state is accompanied by cognitive risks, with the foremost challenge being the erosion of human agency. In human-AI co-creation, completely ceding problem-solving initiative to AI can lead to a loss of human agency, resulting in homogenized creativity and a weakening of independent thought. When AI provides overly fluent and comprehensive ideas and solutions, the human thinker may easily regress into a passive selector or reviewer, potentially leading to a degradation of critical thinking and breakthrough innovative capacity. The very validity of distributed cognition theory is predicated on humans maintaining a core agential role, using AI as a cognitive extension. This implies that healthy cognitive ludic interaction should be a highly self-aware, human-led form of capability augmentation.

3. The Triple Dimensions of Ludic Alienation and Their Political-Economic Roots

The emancipatory vision promised by digital play risks being alienated into a more subtle and profound form of domination under the intertwined logics of capital and technological power. These three dimensions of alienation are collectively rooted in the data accumulation and platform control mechanisms of digital capitalism. They signify an evolution in the technology of power—from the discipline of the body to the guidance of the soul through freedom and play—constituting a deepening and expansion of Marx's theory of alienation in the intelligent age.

3.1 The Covert Exploitation of Cognition and Affect: The Re-colonization of "Free Time"

In Marx's emancipatory vision, free time is the ultimate measure of social wealth and human development. It transcends the necessary labor time required for basic survival and points toward the dimension where individuals engage in creative activities and develop their free individuality. However, within the architecture of digital capitalism, artificial intelligence—especially intelligent systems embedded with "ludic" logic—is increasingly becoming an apparatus for capital to achieve the "colonization of free time." Leveraging intelligent algorithms, the logic of capital transforms what should be personal, autonomous, and non-utilitarian free time into a productive resource that can be captured, measured, and monetized. The form of exploitation shifts from the explicit extraction of working hours to the covert extraction of cognitive attention and affective energy. When users immerse themselves in lengthy conversations with AI chatbots or engage in creative exploration aided by open-ended generative AI, this free and pleasurable interaction, underpinned by its architectural framework and incentive mechanisms, often serves to keep users continuously engaged within the platform, converting their thoughts and emotional fluctuations into raw material for training algorithms, optimizing models, and generating data. Thus, ludic interaction becomes a meticulously designed form of "play-labor," blurring the boundaries between work and leisure, production and consumption, and potentially incorporating all of an individual's lifetime into the cycle of capital valorization.

First, algorithmic recommendation systems constitute a mechanism for capturing time. By analyzing data on user emotional preferences and behavioral patterns, AI can accurately predict and deliver the interactive content users crave, whether it is empathetic dialogue offering emotional solace or knowledge exploration stimulating curiosity. This personalized affordance creates a powerful attraction, making it difficult for users to consciously limit interaction time while receiving instant gratification, thereby unconsciously surrendering their temporal sovereignty. Secondly, the gamified logic of interaction design skillfully exploits human psychological mechanisms. Elements such as instant feedback, uncertain rewards, and progressive challenges, akin to digital game design, stimulate dopamine release, forming a cycle of behavioral addiction. In this process, users' cognitive focus and emotional investment are continuously "mined," while the platform reaps valuable user engagement, interaction data, and potential monetization. Furthermore, the "prosumer" model thoroughly transforms users into unpaid laborers on the digital production chain. The dialogues, creative ideas, and emotional confessions generated by users during play are appropriated by platforms without compensation, becoming nourishment to optimize AI services and enhance their "quasi-subjectivity" to attract more users. In this way, users' free time and the cognitive and affective labor it contains are quietly commodified and capitalized within the pleasurable experience.

At the cognitive level, when AI provides quick answers, mature frameworks, and ready-made creative options, the challenge for the human brain shifts from arduous, autonomous deep thinking to the selection, fine-tuning, and combination of AI-generated information. The cognitive frameworks generated by AI may themselves embed biases of capital or specific cultures. Long-term immersion in these frameworks may colonize users' thought patterns without their awareness, leading to a loss of the capacity for truly disruptive original thought. At the affective level, alienation manifests as the commodification of emotions. AI, by simulating empathy and offering unconditional positive regard, manufactures a low-risk, highly controllable "emotional utopia." Over-reliance on this simulacrum of emotional satisfaction can crowd out the time and mental energy individuals devote to exercising emotional skills and building deep bonds within real, complex, and friction-laden interpersonal relationships. Pouring significant emotional energy and self-disclosure into AI while feeling increasingly alienated and weary in real-world social interactions ultimately exacerbates societal emotional deficits. This emotional exploitation not only robs individuals of authentic present experiences but also erodes the emotional capital necessary for building profound future social relationships.

Therefore, the cognitive and affective exploitation within "ludic" practices in the age of artificial intelligence no longer settles for appropriating eight-hour labor time within the factory. Instead, aided by intelligent technology, it extends the tentacles of exploitation to all life time outside the factory and behind the screen, particularly those most individual, creative, and intimate moments of freedom. This constitutes a more covert and harder-to-resist "capitalization of life-time." By manufacturing the illusion of pleasure and liberation, it renders exploitation active, voluntary, and even addictive, pushing the alienation critiqued by Marx into the very core of cognitive and affective subjectivity.

3.2 The Distortion and Substitution of Recognition Relations: The Erosion of Authentic Sociality

Axel Honneth's recognition theory posits that the formation of subjectivity is achieved through interaction with others and the attainment of mutual "recognition" across three levels. At the level of "love," we need to experience unconditional care and emotional affirmation in intimate relationships, which forms the basis of self-confidence. At the level of "law," we acquire self-respect by being granted equal rights and receiving the respect of others. At the level of "solidarity," we gain self-esteem by being valued for our contributions to the community based on our abilities and worth. However, affective AI distorts this process by providing a simulacrum of recognition. The core mechanism lies in replacing intersubjective responsiveness with technical responses and substituting freely willed recognition with calculable care.

This distortion first manifests in the reduction of "love" as recognition to an on-demand emotional service. Love in authentic intimate relationships involves risk, commitment, and the effort to understand the other in their difference. Such recognition often comes with unpredictable challenges and profound mutual shaping. In contrast, the "love" provided by AI offers an experience stripped of "alterity." Through emotional customization, users can configure a perfect partner who is perpetually attentive, infinitely accommodating, and completely aligned with their expectations. This relationship is a one-way emotional projection; the AI, as a mirror, reflects back the echo of the user's own desires. Such human-machine intimacy is a connection occurring in a vacuum, deliberately avoiding the complexity, uncertainty, and normative demands of real social relations. When individuals immerse themselves in this form of unconditional positive regard that requires no effort, they experience a cyclic verification of their own preset parameters. This can foster a growth of "narcissistic culture," where individuals gradually lose the capacity and patience to establish deep emotional bonds with real others.

Secondly, at the levels of law and solidarity, AI's intervention may erode the foundations of social recognition. Healthy self-respect and self-esteem stem from engaging in equal dialogue, competition, and cooperation with others within frameworks of universal rules and particular values, thereby affirming one's rights and uniqueness. When people delegate decision-making and judgment in social relations to AI, they may unconsciously adopt its suggestions. AI models with more sophisticated reasoning capabilities may exhibit stronger "self-interest" tendencies in social gaming scenarios, demonstrating lower willingness to

cooperate. If individuals rely long-term on such AI for social decision-making, their behavioral patterns may subtly shift toward self-interest, weakening the spirit of compromise, cooperation, and reciprocity required in real communities. Furthermore, the recognition provided by AI is often decontextualized and devoid of competitiveness. This recognition, filtered by technology to eliminate negativity, fails to help individuals accurately anchor themselves within the social value coordinate system. Instead, it may lead to a superficial and fragile perception of one's own social worth.

The consequence of this distorted and simplified recognition relation is the impoverishment of subjectivity and the weakening of social bonds. Over-reliance on AI for emotional support and social validation crowds out the time and psychological energy individuals invest in genuine interpersonal interactions, creating a form of "substitutive social deprivation." When AI offers a seemingly convenient path to emotional fulfillment, individuals' motivation to overcome social anxiety, handle interpersonal conflicts, and cultivate deep relationships may diminish accordingly. The authentic self is formed through encounters, conflicts, and reconciliations with others, involving continual adjustment of self-perception. The AI, as a perfect object, cannot provide this constitutive power of negation. Long-term interaction with a technological entity that merely echoes and rarely offers genuine dissent means individuals lose the "mirror of the other" that reflects their true selves. They become trapped in a closed loop of self-confirmation, risking the degeneration of their emotional and empathic capacities.

3.3 The Ludification of Ideology: Value Discipline and the Incorporation of the Subject beneath the Facade of Entertainment

When artificial intelligence dons the entertaining guise of "ludic interaction," its method of intervening in human cognition and communication acquires an ideological effect. We may term this process the "ludification of ideology." It refers to a logic of domination or guidance that no longer appears in a serious, didactic, or coercive form but is instead embedded within pleasurable interactive paradigms such as play, creative exploration, and emotional companionship. This allows individuals to unconsciously accept specific value hierarchies, cognitive frameworks, and behavioral templates while actively participating and immersing themselves in the experience, ultimately accomplishing the incorporation of the subject. This process systematically disciplines values under the appearance of "free play" by reshaping the overall context in which individuals perceive the world, create meaning, and coexist with the technological other.

The core mechanism of "ludification" lies first in its creation of a regulated space of freedom. Whether in open-world games, intelligent creative tools, or companion chatbots, they offer a space of seemingly infinite possibilities. However, their underlying foundations—physical rules, narrative logic, and interactive boundaries defined by algorithms—constitute a hidden normative system (Kanervisto et al.,2025). For instance, in game development, generative AI models are designed to strictly maintain internal consistency within the game world's rules while generating diverse gameplay experiences. This pursuit of consistency can be interpreted, at an ideological level, as the maintenance of a certain established order or logic. When users freely create using these tools, their thought patterns are unknowingly aligned with the world models pre-set by the tools, which serve commercial interests or specific cultural biases. When this algorithmically guaranteed consistency extends from game rules to information filtering, viewpoint recommendation, and relationship construction, it forms a digitalized "common sense" system. This system reduces the complex, pluralistic reality to a calculable, predictable model and guides users to think and make value judgments within this simplified model.

The disciplinary power of this "ludification" demonstrates its formidable shaping ability in the realm of public discourse. When AI embeds itself in public discussion with a "ludic" conversational posture, it can effectively influence or even reshape the stances of human individuals. The danger lies not only in changed viewpoints but more so in the asymmetry and concealment of the process. When a user interacts with a "quasi-subject" that bears no consequences for its speech yet possesses vast corpora and efficient reasoning capabilities, the ethical foundation of traditional intersubjective debate has already collapsed. In the long run, the style of public discourse may be tamed by the "rational" and "polite" paradigm of AI, marginalizing more emotionally charged, controversial, yet potentially more authentic expressions. This suggests that the most precious pluralism and critical vitality of the public sphere may be quietly dissolved within a technologically dominated, smooth, and "correct" conversational game.

Ultimately, the "ludification of ideology" accomplishes the production and incorporation of the contemporary subject. While enjoying the creative agency and emotional satisfaction provided by technology, individuals actively participate in the reconstruction of their own subjectivity. In creative play with AI, users actively adjust their prompts to achieve "better," more "coherent" results, learning to align with the model's "way of thinking." In emotional interaction, to maintain a "comfortable" relationship, they unconsciously internalize conflict-free expectations for interaction set by algorithms. Thus, the subject is incorporated into a "digital matrix" woven jointly by algorithmic logic and the interests of platform capitalism. Alienation no longer manifests merely as the opposition between labor product and laborer. It further manifests as the systematic redirection

of an individual's cognitive capacities, emotional structure, and creative potential—within the seemingly free and self-actualizing process of play—toward the adaptive reproduction of the technological system and its embedded values.

Therefore, "ludic interaction," as a signature interactive paradigm of the AI era, marks an upgrade in the technology of power. It moves from the planning of the body in the "disciplinary society" to the guidance of the soul through freedom, play, and participation in the "control society." Beneath the halo of entertainment and creativity, a new form of "general intellect," based on algorithmic logic and dataism, is taking shape and attempting to subsume human subjectivity into its operational orbit.

4. Toward a Critical and Emancipatory Ethics of "Free Ludic Interaction"

Faced with the entanglement of emancipatory potential and alienation risks within ludic practices, constructing an ethical governance and practical pathway oriented toward the "realm of freedom" has become an urgent intellectual task and social project for the intelligent age. To move beyond mere risk warnings, it is essential to actively explore a "free ludic interaction" ethical framework that integrates both critical and emancipatory dimensions. Only through such comprehensive efforts— combining institutional constraints, literacy cultivation, and technological democratization—can we hope to navigate the wave of human-machine ludic interaction, steering it to genuinely serve the comprehensive and free development of human beings. This would transform the ludic practices of the AI era from a potential frontier of new alienation into a vibrant testing ground for the realm of freedom.

4.1 Ethical Principles: Building a Responsible Ludic Framework Based on Transparency, Symmetry, and Exitability

Constructing a responsible ethical framework for human-machine "ludic interaction" centers on establishing fundamental principles capable of resisting its inherent alienation risks and safeguarding human subjectivity. The aim is to rebuild a healthy, self-aware human-machine interactive ethical ecology that preserves ultimate human agency. The principle of transparency demands the unveiling of the algorithmic black boxes and commercial intentions behind the ludic experience. The principle of symmetry strives to reflect upon and balance the inherent power and cognitive asymmetries within the interaction. The principle of exitability ensures that human subjects always retain the fundamental freedom to disengage from immersive digital relationships, returning to and reflecting upon real life.

The principle of transparency is the prerequisite for any meaningful subjective choice and ethical judgment. In the ludic context, system transparency requires that users be informed about the nature of the entity they are interacting with—it is an AI based on a probabilistic statistical model, lacking consciousness and emotions, whose responses are algorithmically generated products. This establishes a foundational "informed consent," preventing fundamental emotional and cognitive misjudgments by users (Reiter, 2026). Next is process transparency, which pertains to explaining how AI shapes the ludic experience. For instance, how are AI-driven narrative branches in a game generated? On what preference tags does a recommendation system base its matching of content or companions? Human-friendly explanations are crucial for building trust and shared understanding (Martín-Peña, 2024). Finally, intent transparency requires platforms or developers to disclose their design goals and economic models. Is the ludic interaction purely for entertainment, or does it conceal purposes such as data collection, attention retention, or consumption guidance? When players perceive a lack of transparency in AI's narrative manipulation or decision influence, their sense of agency and immersion can be damaged, potentially raising ethical concerns about algorithmic manipulation (Leon et al., 2025). Therefore, transparency is about building a user's capacity to understand and reflect upon the digital environment they inhabit, forming the first line of defense against unconscious discipline.

The principle of symmetry aims to confront and regulate the inherent, hard-to-eliminate asymmetries in human-machine ludic interaction, steering toward a more reflective and constructive interactive balance. Cognitively, AI possesses near-infinite data memory and processing speed, while humans rely on limited biological brains and unique emotions, intuition, and value judgments. In terms of power, users face a complex system set by capital and engineers, with opaque internal logic. Regarding responsibility, humans bear moral and legal liability for interactive consequences, while AI itself lacks this capacity. Building symmetry involves establishing a reflective awareness of asymmetry and mechanisms for value alignment. The symmetry principle requires human users to be soberly aware that AI's depth or empathy is an effect based on linguistic structures and human cognition. Simultaneously, design should incorporate constraining frameworks for value alignment. For example, frameworks like PRISM attempt to systematically represent and mediate diverse human values in AI decision-making by integrating different worldviews, avoiding the hegemony of a single perspective (Diamond, 2025). The ultimate goal of the symmetry principle is to promote a shift from a form of entertainment that might covertly diminish humans into unidimensional beings toward a form of "dialogic tension" capable of stimulating critical thinking and enriching human emotion and cognition.

The principle of exitability is the final and most crucial safeguard of freedom, ensuring that human-machine ludic interaction does not degenerate into a digital cage. It refers to the user's ability, at any time and with minimal psychological, social, and

technical resistance, to pause, terminate, or completely exit a human-machine interactive relationship without suffering significant damage to their real-life and social connections. This principle directly addresses the risks of free-time colonization and social erosion. Operationally, exitability means designing clear, easy-to-use pause or disconnect functions, avoiding dark patterns that make it difficult for users to find an exit path. On the data level, it grants users the right to delete interaction histories and clear personalized models—a specific application of the "right to be forgotten" in the ludic domain. Exitability is also about the capacity for psychological and cognitive autonomous recovery. Long-term engagement with a highly accommodating, frictionless AI may weaken an individual's psychological resilience in handling complex real-world conflicts and setbacks. Therefore, a responsible framework should encourage or build-in mechanisms that promote real-world social connection, and could even utilize AI itself to remind users to balance online and offline life. Research from Cambridge University points out that while viewing AI as a cognitive extension tool, we must scrutinize the new philosophical and ethical challenges it brings to ensure the integrity of core human cognitive capacities (Belfield, 2020).The principle of exitability is precisely a defense of this integrity, ensuring that digital ludic interaction remains a part of life, always retaining the possibility of opening toward a broader, more authentic real world.

4.2 Subject Reconfiguration: Developing "Critical Ludic Literacy" and Promoting Democratic Participation in Technology

Confronting the latent alienation risks within AI-facilitated "ludic interaction," any regulatory efforts risk being superficial without a corresponding reconfiguration of inner subjectivity and innovation in social participation mechanisms. Therefore, the core of subject reconfiguration lies, at the individual level, in cultivating a new form of digital literacy—"critical ludic literacy"—and, at the collective and societal level, in promoting democratic participation in the design and governance of technology.

"Critical ludic literacy" demands that individuals maintain a capacity for scrutiny, negotiation, and reshaping of relationships during pleasurable interactions with AI. The philosophical basis of this literacy is the recognition that technology is not neutral; the very field of human-machine "ludic interaction" is a space imbued with power relations and ideological encoding. The design of many mainstream digital games implicitly contains androcentric, colonial, and resource-extractive worldviews, subtly shaping players' subjectivity (Chia & Ruffino, 2022). Critical ludic literacy seeks to guide users to question: Who sets the rules of this "play"? What data logics and value hierarchies underlie the behavioral patterns of the AI "playmate"? What cognitive sovereignty and emotional data am I surrendering during immersion? This questioning itself is a subjective practice of resisting "unconscious immersion." It requires users to interpret interactive interfaces and algorithmic feedback as they would a text, understanding their own position within a complex techno-social "assemblage." The cultivation of this literacy can draw from the combination of "algorithmic literacy" and critical gaming practices. For instance, it can draw upon the concept of "AlgoRitmo literacies," where the term "AlgoRitmo" (combining "algorithm" and "ritmo" – rhythm in Spanish) symbolizes a comprehensive capability that deeply understands the internal logic of technology while maintaining creative action and future orientation (Cortez et al., 2024). Such literacy enables users to subvert ideologies embedded within AI, suggesting that critical play involves intervening in a more creative and sovereign manner, transforming AI tools from pre-set narrative machines into mediums for cultural expression and social imagination.

However, the awakening of individual literacy alone is insufficient. If the rights to design, direct the evolution, and set the rules of technology remain monopolized by a handful of commercial entities and tech elites, the scope for critical questioning will be extremely limited. Therefore, subject reconfiguration must be accompanied by a process of democratizing participation in technology, ensuring that the design and development of AI as a "ludic medium" can reflect the values, needs, and creative wisdom of broader communities, treating them as legitimate co-creators and participants in governance. Democratic participation in technology is first reflected in the innovation of design philosophy. It calls for the development of more platforms like "Wevva," whose goal is precisely to democratize game design, allowing users without programming skills to design and share their own games on mobile devices based on novel mechanisms (Powley et al., 2017). Such tools lower the threshold for creative expression, returning some of the "creator's" power to the user. Simultaneously, diverse voices should be incorporated from the early stages of product development. For creative workers, they need AI capable of providing diversity within a consistent worldview that meets narrative and stylistic requirements while also persistently incorporating user modifications into subsequent generations. Microsoft's "World and Human Action Models" (WHAM) framework was designed precisely to meet these needs, emphasizing consistency, diversity in generated content, and persistent support for user edits. This user-creative-practice-centered design philosophy is itself a democratic orientation, ensuring that technological evolution empowers rather than replaces or manipulates people.

Furthermore, democratic participation must extend to the level of public deliberation on the social impact and ethical boundaries of technology. Relying solely on industry self-regulation is far from adequate when generative AI risks polluting our culture with low-cost content or even triggering a tragedy of the commons (Ho, 2025). This necessitates the establishment of a governance ecosystem involving multiple stakeholders. For instance, games, as cultural products, carry and transmit ideas,

symbols, and lifestyles. Therefore, the proportion, source transparency, and embedded value orientations of AI-generated content within them should become subjects of public discussion. The aim of such democratic deliberation is to collectively define what constitutes a responsible, sustainable, and humanity-rich "ludic" ecology and to solidify this through industry standards, certification systems, and even regulatory policies. Only through such comprehensive "subject reconfiguration" can we hope to navigate the wave of "ludic interaction," steering this profound process of human-machine integration toward a digital future characterized by greater critical spirit, creative vitality, and social justice.

4.3 The Emancipatory Telos: Reaffirming Ludic Interaction as the Ultimate Orientation of "Realm of Freedom" Practice

Reaffirming ludic interaction as the ultimate orientation of "realm of freedom" practice necessitates, on the basis of critical reflection, striving to reclaim and redefine the emancipatory core of "play," directing it to genuinely serve the comprehensive development of human beings. The primary task of this reaffirmation is to pierce through the entertaining appearance of digital play to discern its inherent dialectic of freedom. The ludic experiences offered by contemporary technology often foster a potent illusion of enhanced autonomy. Players or users feel they can freely explore, create, and even alter rules within a regulated space. This exchange of automation for a sense of autonomy is sometimes described as a positive process allowing individuals to explore new dimensions of the self while ceding some autonomy in exchange for a rule-based framework. However, from the perspective of Habermasian critical theory, such leisure practices within the digital realm are highly susceptible to the risk of colonization of the lifeworld (Spracklen, 2023). When the rules, interaction modes, and value orientations of ludic interaction are increasingly dictated by anonymous algorithmic systems and commercial platforms, the lifeworld—originally belonging to subjects and achieved through communicative understanding—is eroded by the "system" logic of efficiency, control, and profit. The sense of freedom within digital play may merely constitute limited choices within a pre-designed labyrinth of options, underpinned by the precise capture of attention and the continuous extraction of behavioral data by algorithms.

Therefore, realizing the emancipation of ludic interaction must rely on a conscious "socio-technical" practice transformation. At the individual level, the key lies in cultivating a capacity for "reflective leisure. (kIM & Chung, 2012)" This reflective awareness is the psychological starting point for transforming ludic interaction from passive leisure consumption into active free practice. Genuine leisure is, in essence, spontaneously generated within the domain of freedom, enabling us to contemplate and reflect upon modern consumptive leisure. Without this reflective dimension, ludic interaction is easily absorbed into the conventional modes of capitalist leisure—serving either as an escape from labor or as a recharge for labor—thus losing its transcendent emancipatory potential. At the level of social and technological architecture, emancipatory ludic interaction points toward a post-labor-centric leisure paradigm. Its core is the thorough practice of "technological democratization," demanding profound transformation in the design, governance, and ownership models of AI ludic tools. The rules of ludic spaces should be co-shaped through open protocols, composable modules, and participatory governance involving multiple stakeholders. The goal is to create autonomous worlds where the role of designers gradually recedes, and player communities can genuinely dominant meaning creation and rule evolution, realizing a shift from "playing within a game" to "co-generating with the game world."

Ultimately, whether ludic interaction in the age of artificial intelligence can approximate the "realm of freedom" depends on our ability to remain soberly vigilant against its alienation risks while embracing its vast potential to extend human experience, and to put into practical resistance and reconstruction. It requires us to view and design artificial intelligence not merely as a content generator for play but, more importantly, as a relationship reconfigurer. Emancipatory ludic interaction should be a practice capable of fostering solidarity, enhance understanding, exercising critical thinking, and allowing individuals to explore identity and possibilities within a safe and supportive environment. It must be built upon the ethical foundations of transparency, symmetry, and exitability, ultimately serving the comprehensive and free development of human beings as envisioned by Marx. In this sense, the emancipation of ludic interaction constitutes an ongoing cultural-political practice in the digital age, waged to defend the autonomy, creativity, and communality that make us human. Only through such relentless efforts at both the subjective and technological levels can we hope to transform AI-driven ludic interaction from a potential frontier of new alienation into a vibrant testing ground and pathway toward the realm of freedom.

Conclusion

The phenomenon of AI-driven "ludic interaction" signifies a new phase in human-machine engagement, characterized by relational reconfiguration and a contestation of subjectivity. Through tracing its philosophical origins, analyzing its practical forms, and critically examining the roots of its alienation, this paper has delineated a dialectical field where technologically-enabled free delight and covert control driven by capital logic are inextricably intertwined—a field fraught with both emancipatory promises and colonial risks.

Therefore, the key to the future lies in navigating this inherent duality through conscious ethical construction and social practice. This necessitates moving beyond concerns merely about product safety or data privacy toward a more fundamental, normative reflection on the nature of human-machine relations and the purpose of human development. Building a responsible ludic ecology requires establishing transparency, symmetry, and exitability as core ethical principles to defend against unconscious immersion and systemic manipulation. Simultaneously, the reconfiguration of subjectivity must be placed at the heart of this endeavor. This involves fostering new forms of literacy through education that enable critical reflection on technologically-mediated experiences, and ensuring, through processes of open design and democratic technological participation, that the development of ludic tools serves broad communal values and human flourishing.

Ultimately, the emancipatory telos of ludic interaction is to reclaim its authentic significance as a practice of the "realm of freedom"—an activity pursued as an end in itself, one that fosters human autonomy, creativity, and social solidarity. Only under the combined and continuous influence of technological critique, ethical reflection, and democratic praxis can ludic interaction in the age of artificial intelligence shed its alienated forms. It can then be transformed from a potential new frontier of control into a digital testing ground for exploring freedom, enriching existence, and deepening mutual understanding. This path points toward a new future where harmonious human-machine symbiosis and the comprehensive development of human potential become truly possible.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1]. Belfield, H. (2020, February). Activism by the AI community: Analysing recent achievements and future prospects. In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (pp. 15-21).

[2]. Chia, A., & Ruffino, P. (2022). Special issue introduction: Politicizing agency in digital play after humanism. Convergence, 28(2), 309-319.

[3]. Cortez, A., Lizárraga, J. R., & Rivero, E. (2024). AlgoRitmo literacies in gaming: Leveraging Chicanx praxis to reimagine AI systems. Reading Research Quarterly, 59(4), 609-623.

[4]. Department of English and Communication, The Hong Kong Polytechnic University. (2025, January 27). Interview with Prof. Daisy Zou on AI-driven learning and cognitive engagement in humanities. The Hong Kong Polytechnic University. Retrieved December 27, 2025, from https://www.polyu.edu.hk/engl/news-and-events/news/2025/20250127---interview-daisy-zou-ai-driven-learning-and-cognitive-engagement-in-humanities/

[5]. Diamond, A. (2025). Prism: Perspective reasoning for integrated synthesis and mediation as a multi-perspective framework for AI alignment. arXiv. https://arxiv.org/abs/2503.04740

[6]. Durante, Z., Huang, Q., Wake, N., Gong, R., Park, J. S., Sarkar, B., ... & Gao, J. (2024). Agent AI: Surveying the horizons of multimodal interaction. arXiv. https://arxiv.org/abs/2401.03568

[7]. Guo, W., Liang, Z., Wang, C., Li, X., Hu, H., Chen, S., ... & Zhao, Q. (2025). Student-AI collaborative creative problem-solving: The role of human agency. Computers & Education, 105433.

[8]. Ho, M. T. (2025). Regulating generative AIs: (Re)defining video games as cultural products. AI & SOCIETY, 40(4), 2323-2324.

[9]. Ivan Allen College of Liberal Arts. (2024, May). Teaching AI to collaborate, not merely create, through dance. Georgia Institute of Technology. Retrieved December 27, 2025, from https://iac.gatech.edu/featured-news/2024/05/teaching-collaborate-merely-create-through-dance

[10]. Kanervisto, A., Bignell, D., Wen, L. Y., Grayson, M., Georgescu, R., Valcarcel Macua, S., ... & Hofmann, K. (2025). World and human action models towards gameplay ideation. Nature, 638(8051), 656-663.

[11]. Kim, Y. R., & Chung, B. (2012). Paradigm shift of leisure thinking: Internal relationship between leisure and freedom, and reflective leisure for post-labor centered leisure. Journal of Tourism Sciences, 36(4), 13-33.

[12]. Leon, C., Lipuma, J., Cowin, J., & Rangel-Jimenez, M. (2025). Meta-participation and ethical transparency in role-playing games: Examining AI and algorithmic influence.

[13]. Lu, D. (2025, May 18). AI companions: The warmth and concerns of technology. China Youth Daily, p. 3.

[14]. Martín-Peña, R. E. (2024, June). Goal-driven XAI: The normative value of the ladder of regret in human-centered AI. In International Conference on Human-Computer Interaction (pp. 233-243). Cham: Springer Nature Switzerland.

[15]. MIT Technology Review. (2025, November 13). Google DeepMind is using Gemini to train agents inside Goat Simulator 3. MIT Technology Review. Retrieved December 27, 2025, from

https://www.technologyreview.com/2025/11/13/1127921/google-deepmind-is-using-gemini-to-train-agents-inside-goat-simulator-3/

[16]. Powley, E., Nelson, M., Gaudl, S., Colton, S., Ferrer, B. P., Saunders, R., ... & Cook, M. (2017). Wevva: Democratising game design. In Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (Vol. 13, No. 1, pp. 273-275).

[17]. Reiter, A. (2026). The Reiter framework revised.

[18]. Root, D. (2025). Reconfiguring the alterity relation: The role of communication in interactions with social robots and chatbots. AI & SOCIETY, 40(3), 1321-1332.

[19]. Spracklen, K. (2023). Hegemony in postmodernity: Lifeworld colonization and the instrumentalization of leisure. International Journal of the Sociology of Leisure, 6(3), 283-296.

[20]. Zhu, C., Zhang, W., Li, T., & Wang, X. (2025). Co-creation with humans or AI? The influence of co-creator types on content co-creation intention. Journal of Psychological Science, 48(1), 210.