Peeling Back the Layers: The Co-Evolutionary Onion of Human-AI Co-Creativity

Mahsa Goodarzi
University at Albany, SUNY
mgoodarzi@albany.edu

Abstract

Generative Artificial Intelligence offers creative superpowers that boost productivity and inspiration. However, it simultaneously presents the risk of creative convergence, where algorithmic optimization leads to the homogenization of ideas and the erosion of human agency. Addressing this tension requires reframing the Human-AI interaction from a focus on frictionless efficiency to one of Human-AI Co-Evolution, a symbiotic state where humans and machines continuously adapt to and elevate one another. This paper introduces the Co-Evolutionary Onion of Human-AI Co-Creativity, a multi-layered theoretical framework designed to structure this engagement.

Introduction

We stand at a technological inflection point defined by the "jagged technological frontier" (Dell'Acqua et al. 2023), where AI demonstrates superhuman performance in specific creative tasks while failing basic contextual reasoning in others. We are also witnessing a phenomenon described as "AIsmosis," the subtle, gradual, and seamless assimilation of AI into the fabric of daily life and societal norms (Bozdag 2023). Unlike previous technological shifts, the integration of generative AI into creative domains is not merely optimizing workflows. It is reshaping the cognitive processes of discovery and expression.

AI offers "creative superpowers," acting as a source of inspiration that boosts productivity and confidence (Geyer et al. 2020; Lee et al. 2025; Li et al. 2024a). It opens pathways for scientific discovery that transcend human cognitive limitations and acts as a "social glue" in collaborative settings (Suh et al. 2021). However, the core challenge posed by the emergence of generative Artificial Intelligence is fundamentally one of creative convergence (Grammenos and Lubart 2025; Smith et al. 2025). AI models often reinforce existing patterns and homogenize ideas, hindering the pursuit of "radical innovation" that requires bridging conceptual gaps (Kumar et al. 2025; Lee et al. 2025; Mirowski et al. 2023). Addressing this tension requires moving beyond AI as a tool for efficiency toward systems that are capable of fostering deep, mutual creativity and innovation. This paper, through an extensive review of literature, argues that to ensure AI sparks creativity rather than extinguishing it, we must fundamentally reframe the design of Human-AI Interaction. It traces a path from the dangers of creative convergence to the realization of Human-AI Co-evolution, a symbiotic state where humans and machines continuously learn from, adapt to, and elevate one another.

The Trap of Creative Convergence

Innovation often arises from the juxtaposition of concepts that do not statistically belong together, often termed transformational creativity (Franceschelli and Musolesi 2024; Boden 2004). However, Generative AI models, driven by reinforcement learning from human feedback, are typically optimized for coherence and probability (Chakrabarty et al. 2024; Mirowski et al. 2024). Current mainstream models struggle with original creativity due to training that typically optimizes coherence, predictability, and pattern recognition (Vinchon, Gironnay, and Lubart 2024). Consequently, they excel at interpolation but struggle with extrapolation into unknown territory (Li et al. 2024b). While these models can enhance the productivity of individual creativity, they have been shown to reduce the collective diversity of novel creation produced by a group (Tang et al. 2024; Kumar et al. 2025). This "mechanized convergence" or homogenization suggests that AI pushes toward a statistical mean, narrowing the window of creative expression (Lee et al. 2025; Kumar et al. 2025; Tang et al. 2024). If left unchecked, this creates a feedback loop where future models are trained on homogenized AI-generated content, further eroding diversity (Pedreschi et al. 2025). Therefore, the critical challenge is designing AI systems capable of injecting productive friction, or creative divergence, into the human-AI loop. The goal is to prevent creative stagnation while guiding users toward genuinely novel and distant conceptual spaces. This objective requires surpassing combinatorial and exploratory creativity to achieve transformational creativity, changing the conceptual space to allow previously inconceivable thoughts, ideas, or structures to become possible (Bickley, Macintyre, and Torgler 2025; Boden 2004; Franceschelli and Musolesi 2024; Wiggins 2006).

The drive for efficiency often conflicts with the conditions necessary for deep creativity. Creativity requires incubation, wandering, and friction; processes that efficient automation seeks to eliminate (Lamb, Brown, and Clarke 2019; Lawton, Grace, and Ibarrola 2023; Lee et al. 2025). Hofman, Goldstein, and Rothschild (2023) provide a useful metaphor for AI's role. It can act as "steroids", a short-term performance boost with potential long-term harm, "sneakers," an augmentation without harm, or a "coach," helping humans improve their capabilities (Kumar et al. 2025). Currently, many generative AI applications function as "steroids," automating the cognitive struggle required for deep thinking (Lee et al. 2025). These models often eliminate the necessary friction in favor of speed and magic end-to-end generation (Lawton, Grace, and Ibarrola 2023). By bypassing the impregnation or incubation phases, efficient automation risks hollowing out the creative act, leaving behind a product without the depth of process (Lamb, Brown, and Clarke 2019). This leads to "cognitive atrophy," where creators lose the critical thinking and problem-solving skills necessary for independent creativity (Hernández-Orallo 2025). Furthermore, over-reliance on algorithmic suggestions can lead to a degradation of decision-making skills, loss of vigilance, and automation bias (Lee et al. 2025; Parasuraman, Sheridan, and Wickens 2000; Robertson et al. 2024). This phenomenon occurs when users assume AI outputs are inherently superior or correct, leading to 'complacency' where they fail to critically engage with or verify the system recommendations (Parasuraman, Sheridan, and Wickens 2000; Robertson et al. 2024). Consequently, when the human in the loop is eventually forced to intervene during a system failure, they often lack the situational awareness and recent practice necessary to act effectively (Lee et al. 2025; Parasuraman, Sheridan, and Wickens 2000). Therefore, the replacement scenario is not necessarily a dramatic takeover by sentient machines, but a gradual erosion of human creative agency through creative convergence. To counter this result, we must intervene in the design of the human-AI loop.

Bridging Conceptual Gaps

Radical innovation requires bridging the gap between connected and unconnected ideas. While humans often fixate on familiar examples (design fixation), AI systems can retrieve and present far-field analogies that a human might not consider (Kang et al. 2025; Lin et al. 2025). Effectively activating human creativity requires moving beyond surface similarity to structural similarity. For example, systems like BioSpark leverage Large Language Models (LLMs) to facilitate biological analogical ideation, enabling designers to bridge the gap between biological mechanisms and engineering problems (Kang et al. 2025). By systematically navigating the latent space to find concepts that are semantically distant yet contextually relevant, AI can act as a conceptual blender, fusing distinct domains to generate structurally new ideas (Lopes and Yannakakis 2021; Deshpande et al. 2023; Ibarrola, Lulham, and Grace 2024; Kang et al. 2025).

Contrary to the goal of frictionless efficiency, creativity often benefits from incubation and delay. Research indicates that implementing wait times or processing delays in AI interactions can actually foster greater human reflection and divergent thinking. In the CoQuest system, AI processing delays allowed users to reflect on multiple research questions simultaneously, resulting in higher-quality outcomes than immediate generation (Liu et al. 2024). This result suggests that latency is a feature, not a bug, for specific types of creative work. By intentionally slowing down the process, designers can create a temporal space to regain agency, engage in agentic flow, and produce work that is not only efficient but deeply reflected upon and structurally novel (Liu et al. 2024; Lawton, Grace, and Ibarrola 2023). Well-being AI should not always be instant. It should introduce productive friction that forces the human to engage deeper with the problem space, preventing the user from rushing to the first AI-generated solution. It should not be a vending machine for instant answers; it should be a collaborator that occasionally pauses, questions, and challenges the user to ensure that the outcome is a result of shared cognitive struggle, not just algorithmic probability.

The Co-Evolutionary Onion of Human-AI Co-Creativity

While general collaboration often involves individuals working together, co-creativity is a process where multiple parties contribute to a shared creative endeavor in a blended and synthetic manner. Unlike a simple distribution of labor where tasks are compartmentalized, completed independently, and assembled at the end, co-creativity requires the continuous fusing, merging, and building upon of each other's ideas. The fundamental premise of this process is that the synergy between participants yields a collaborative outcome superior to what any individual could achieve alone, and the resulting product is greater than the sum of its parts (Davis 2021; Sasson Lazovsky, Raz, and Kenett 2025). When applied to artificial intelligence, human-AI co-creativity fundamentally shifts the role of the computational system from a passive tool to an active, creative colleague (Rezwana and Maher 2023). This specific form of co-creativity relies on mixed-initiative interaction, meaning both the human and the AI possess the agency to proactively initiate ideas, make independent contributions, and steer the creative trajectory. It also involves real-time improvisation (Margarido et al. 2026). Human-AI co-creativity is a dynamic partnership where both human and AI actively contribute, each responding to the other's input while guided by their unique understanding of creativity (Deshpande et al. 2024). This continuous, reciprocal feedback loop collaboratively generates meaning and novel ideas, transcending original intentions (Rezwana and Maher 2023).

Figure 1 illustrates that co-creativity is protected by layers of context and enabled by cyclical inputs.

The Foundation

The outermost layer represents the environment, the foundation, and the ethical conditions required for the inner layers to function. We cannot view AI as a standalone tool. Designers must embed AI within a Socio-Technical System (STS) that accounts for organizational norms, culture, and human values. Human-Centered AI (HCAI) is not just a design methodology but the ethical compass of the system. It shifts the focus from "what AI can do" and technological capability to "what people need" and human flourishing. HCAI places people, rather than technology, at the center of system design (Bingley et al. 2023; Constantinides et al. 2024). Its primary goal is Intelligent Augmentation (IA) by amplifying, empowering, and enhancing human performance rather than automating humans out of the loop (Henriksen and Blond 2023). While HCAI provides the values, STS provides the structural context. The STS perspective views the creative environment as consisting of two interdependent subsystems: social (people, culture, processes) and technical (AI, algorithms, infrastructure) (Berretta et al. 2023). Success requires designing both systems together. We cannot optimize the AI (technical) without considering how it reshapes the team dynamics and social fabric (social) (Carter and Wynne 2024; Berretta et al. 2023). STS emphasizes that the introduction of AI transforms the entire ecosystem, creating new behavioral effects, power dynamics, and interdependencies among human team members (Ang et al. 2024). Together, HCAI and STS form the protective skin of the onion. They define the boundaries of safety and purpose, ensuring that the inner layers of interaction and friction lead to human flourishing rather than alienation or deskilling (Shin and Shin 2023). Without this foundation, the feedback loops inside the onion risk becoming recursive cycles of bias or homogenization.

The Enablers

Moving inward, the second layer acts as the filter. Even a well-designed system will fail without trust and transparency. These are the enablers that allow humans to accept AI as a teammate rather than a "black box" threat. These serve as the critical players that allow the evolving symbiosis of co-creativity to occur by mitigating the risks of cognitive complacency (Lyons et al. 2023; Simón, Revilla, and Jesús Sáenz 2024). Trust is the fundamental prerequisite for interaction and adoption, and the willingness to rely on an intelligent system despite uncertainty (Li et al. 2022). However, within this framework, trust is not a static attribute but a dynamic process of trust calibration (De Visser et al. 2020). This layer guards against over-trust, unthinkingly accepting AI outputs, and under-trust, rejecting AI assistance due to skepticism or fear (De Visser et al. 2020; Li et al. 2022). Trust encompasses both cognitive and effective trust that are essential for humans to accept the AI as a creative partner rather than a tool (Erengin, Briker, and De Jong 2025; Georganta and Ulfert 2024).

Transparency acts as the mechanism that makes trust calibration possible. To function as a teammate, the system must move beyond opacity through Explainable AI (XAI). This process involves making the AI decision-making logic, data provenance, and limitations visible to the user (Mylrea and Robinson 2023). In a co-creative onion model, transparency is not only the AI explaining itself to the human. It is bidirectional (Iftikhar et al. 2024). Humans must understand the process to effectively steer it, and the system must be transparent about confidence levels to trigger human intervention when necessary (Simón, Revilla, and Jesús Sáenz 2024; Zheng et al. 2023).

Trust does not happen in a vacuum. It depends on the socio-technical environment. By embedding AI within an HCAI framework that prioritizes human flourishing, designers mandate transparency. This systemic transparency provides the credible information necessary for humans to assess the intent and logic of AI, driving the calibration of trust. Together, trust and transparency act as a filter for the next layer, the input loop. Creative work requires vulnerability. High-quality transparency allows the human to feel safe enough to take creative risks and share novel inputs without fear of the AI hallucinating or misappropriating the work (Rezwana and Maher 2023; Suh et al. 2021).

Reciprocal Input Cycle

The third layer of the onion is not a static repository of human and machine capabilities, but a continuous feedback loop of interdependent agency. In this layer, human inputs (intuition, goal-setting, context, etc.) and AI inputs (pattern recognition, generation, etc.) do not simply exist side-by-side. They recursively feed into one another. These outputs immediately become new inputs for humans, altering their perception of the problem space and triggering new cognitive associations. This cycle ensures that input is treated as a living, iterative process rather than a one-off command (Pedreschi et al. 2025; Robertson et al. 2024). This looping structure addresses the risk of creative convergence by ensuring that inputs are not unidirectional (human $\rightarrow$ AI). Instead, the loop allows for productive dialogue and reciprocal feedback (Simón, Revilla, and Jesús Sáenz 2024). By treating inputs as a cycle, the framework acknowledges that the human must constantly refine the AI output, while the AI challenges humans' fixation (Pedreschi et al. 2025; Lee et al. 2025). This layer effectively filters raw capabilities into useful collaborative friction.

In this loop, the human moves from being a manual creator to an orchestrator of a project manager, responsible for setting the trajectory and evaluating the quality of the output (Palani and Ramos 2024). While AI handles low-level creativity, humans provide high-level creativity. This process involves defining the initial problem, setting goals, and injecting intentionality and authenticity, qualities that current AI systems lack (Garcia 2024; Vinchon, Gironnay, and Lubart 2024). Humans rely on intuition to navigate ambiguity and make judgments where data is incomplete or context-dependent (Jarrahi et al. 2023; Shin and Shin 2023). In the loop, humans use tacit knowledge and gut feeling to evaluate AI outputs that may be statistically probable but contextually inappropriate (Van Den Bosch et al. 2019). This interaction allows the human to filter AI hallucinations or irrelevant patterns (Fuchs, Passarella, and Conti 2023). In this loop, agency is preserved when the human retains the power to accept, reject, or modify AI suggestions (Smith et al. 2025). It ensures the human remains the mastermind behind the work rather than a passive consumer of algorithmic outputs (Lyu et al. 2022). Psychological ownership and trust depend on the human's sense of control. If the AI is too autonomous, users may experience algorithmic aversion or a loss of creative identity (Smith et al. 2025). Therefore, the design of the loop accounts for satisfying humans' need for meaningful control, allowing them to intervene and steer AI when it deviates from their intent (Shin and Shin 2023).

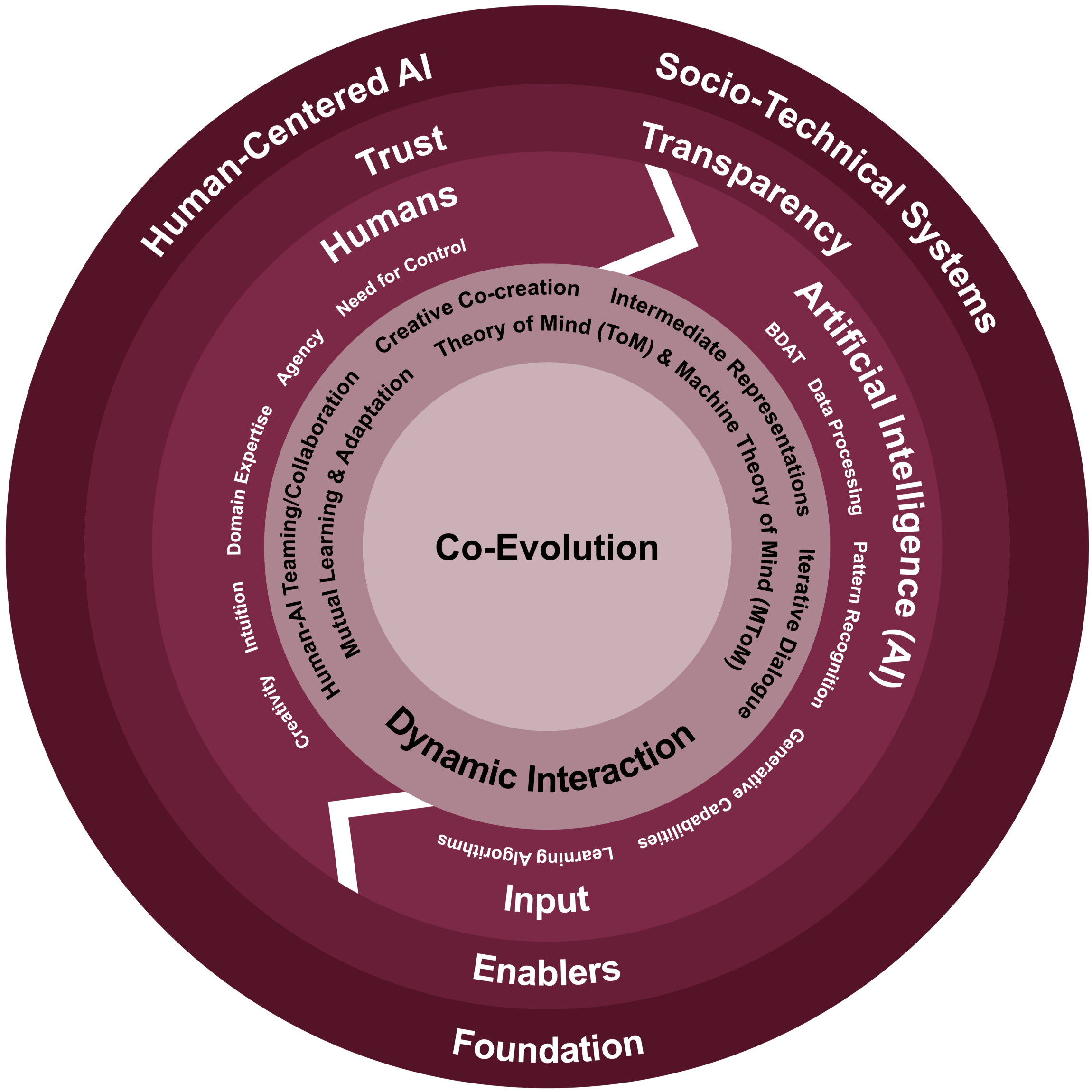

Figure 1: The Co-Evolutionary Onion of Human-AI Co-Creativity based on the extensive synthesis of Literature

The AI acts as a force for progress and a provoking agent, providing the raw material and computational power that scales human intent. AI brings the capacity to process and analyze vast amounts of unstructured data (images, text, logs, etc.) at speeds and scales that far surpass human cognitive limits (Chowdhury et al. 2023; Siemon 2022; Van Den Broek, Sergeeva, and Huysman Vrije 2021). This capability allows the system to retrieve relevant information or precedents that the human may not have access to or memory of (Jarrahi et al. 2023). Through deep learning and neural networks, AI identifies complex, non-linear patterns within the data (Bozdag 2023). It detects correlations and structural regularities that are invisible to the human eye, offering these as "insight" or "predictions" that feed back into the human's decision-making process (Siemon 2022; Van Den Broek, Sergeeva, and Huysman Vrije 2021). Beyond analysis, AI uses probabilistic models to synthesize new artifacts based on learned distributions (Davis, Sherson, and Rafner 2025; Vinchon, Gironnay, and Lubart 2024). This capacity supports divergent thinking by rapidly producing a high quantity of varied ideas, preventing the human from getting stuck, overcoming the blank page syndrome (Kumar et al. 2025; Wang et al. 2024b). The engine of AI is the ability to learn from experience. Through mechanisms like Reinforcement Learning with human feedback, AI updates its weights and parameters based on humans' interactions (Darban 2024; Van Den Bosch et al. 2019). This mechanism allows AI to adapt to the specific style, preferences, and corrections of the human user over time, facilitating a personalized partnership (Wu et al. 2021).

The magic of this layer occurs in the reciprocal handoff between these attributes. The loop activates when human agency and creativity formulate a goal, constrained by domain expertise. AI engages big data analytics and pattern recognition to interpret this intent, using generative capabilities to produce a set of options or artifacts. The human employs intuition and their need for control to assess these outputs. They reject average results and select or edit the promising ones (Palani and Ramos 2024). The learning algorithm observes this selection. It updates its internal model, effectively learning the humans' preferences. This updated state triggers a new cycle, where the next output is more aligned with the humans' intent, and the humans' next prompt is more informed by the AI capabilities. This cycle ensures that the system does not merely automate tasks but co-evolves. Humans sharpen their expertise and intuition by seeing new possibilities, while AI refines its generative accuracy through human guidance (Pedreschi et al. 2025).

The Engine - Dynamic Interactions

While the outer layers provide the environment and the inputs provide the raw capability, the fourth layer, dynamic interaction, describes how the collaboration unfolds. This layer shifts the paradigm from using AI as a passive tool to engaging with it as an active teammate through specific socio-cognitive processes. The inputs generated act as a boundary object or intermediate representation. By co-creating with AI through a shared artifact, humans enter a dynamic process of participatory sense-making. Here, meaning is not static but emerges continuously, driven by the ”inner loop” of teaming (Yang, Folke, and Shafto 2023). This loop involves a constant interplay of communication (sharing inputs) and inference (updating mental models).

HAIT and Co-Creativity. Human-AI Teaming (HAIT) redefines the interaction as a partnership where humans and AI work interdependently toward shared goals. Unlike traditional automation, HAIT requires the AI to process teamwork skills, such as communicativeness and coordination (O'Neill et al. 2023; Berretta et al. 2023). Creative Co-creation focuses on the synergy where the combined creative output is superior to what either could achieve alone (Davis 2013; Sasson Lazovsky, Raz, and Kenett 2025). This results in a mixed-initiative system where both parties can proactively contribute to and steer the creative trajectory (Alvarez et al. 2021; Rezwana and Maher 2023).

The Mechanism. Interaction is not a single event but a temporal process. This layer relies on the continuous exchange of information and the manipulation of shared artifacts. Meaning is co-constructed through productive dialogue, a back-and-forth exchange where humans and AI negotiate goals and interoperability (Simón, Revilla, and Jesús Sáenz 2024; Bown et al. 2025). This dialogue can be linguistic (natural language prompts) or non-linguistic (actions upon an artifact). The quality of this dialogue facilitates mutual knowledge gains and helps repair trust when the AI fails (Simón, Revilla, and Jesús Sáenz 2024; Iftikhar et al. 2024; De Visser et al. 2020). The dialogue often centers on an intermediate representation (e.g., a sketch, a draft, or a musical fragment). This shared artifact acts as a boundary object where humans and AI negotiate meaning (Lawton, Grace, and Ibarrola 2023; Buschek 2024). By manipulating this representation, such as a user refining an AI-generated sketch or AI suggesting a completion to a user's text, the partners engage in participatory sense-making, where the artifact itself becomes the medium of communication (Davis et al. 2016; Deshpande et al. 2023).

The Cognitive Bridge. For dynamic interaction to be effective, the partners must understand each other's hidden states (Yang, Folke, and Shafto 2023; Deshpande et al. 2024). Theory of Mind (ToM) is the human ability to attribute mental states (belief or intent) to the AI. Humans naturally attribute agency to AI, but this can lead to expectation gaps if the AI capabilities are misunderstood (Wang et al. 2024a; Bozdag 2023). On the other hand, Machine Theory of Mind (MToM) enables the AI to infer human goals, knowledge gaps, and mental states from their behavior. This allows the AI to move beyond reactive processing to proactive support, anticipating human needs and adjusting its contributions accordingly (Yang, Folke, and Shafto 2023; Van Den Bosch et al. 2019). The ideal state in this layer is a Mutual Theory of Mind, where both agents continuously simulate and infer each other's states, creating a shared mental model that aligns their understanding of the task and the team (Wang et al. 2024a; Ross et al. 2023).

Mutual Learning and Adaptation. The dynamism of this layer comes from the fact that the interaction changes the participants. Mutual adaptation is the process by which both parties adjust their behavior to accommodate the other (Bown et al. 2025). Humans learn AI quirks and capabilities by adapting their prompting strategy, while AI fine-tunes its output to humans' preferences (Jarrahi et al. 2023; Lee et al. 2025). This interaction leads to mutual learning through a human-AI feedback loop (Pedreschi et al. 2025; Simón, Revilla, and Jesús Sáenz 2024). AI fills gaps in human knowledge by providing far-field analogies or data insights (Kang et al. 2025). Humans fill gaps in AI knowledge by providing context and values (Bozdag 2023). This reciprocal adaptation prevents the interaction from becoming static, driving the system toward the core of the onion: Co-Evolution (Pedreschi et al. 2025).

The Core - Co-Evolution

At the center of the onion lies Co-Evolution, representing the cumulative, longitudinal transformation of both the human and the system into a unified socio-technical entity. The movement from immediate interaction to permanent co-evolution occurs when dynamic interactions are sustained over time. As humans continuously refine AI through feedback, the learning algorithms update weights; recursively, the improved, unpredictable outputs shape humans' future preferences and cognitive pathways. Over longitudinal time scales, these accumulated micro-adaptations transcend the specific task and result. The core embodies the permanent alteration of the partners through "AIsmosis," the gradual, seamless integration of AI into the human lifeworld and social norms (Bozdag 2023; Pedreschi et al. 2025). This symbiosis moves beyond simple augmentation to what Hernández-Orallo (2025) describes as extended cognition, where the AI becomes an intrinsic, coupled part of humans' problem-solving apparatus. However, to ensure this core results in "Well-Being AI," the co-evolution must be actively steered toward a mutualistic symbiosis that fosters human flourishing, guarding against a parasitic symbiosis where the system improves efficiency at the expense of human skill atrophy or agency (Henriksen and Blond 2023; Goodarzi 2025).

Conclusion

A causal chain drives movement through the Onion Model. Systemic HCAI design mandates transparency, which calibrates trust. This trust provides the psychological safety necessary for humans to provide authentic creative inputs. The exchange of these inputs creates boundary objects that trigger participatory sense-making, elevating the exchange into a dynamic interaction. Sustained interactions create perpetual feedback loops, prompting mutual adaptation and ultimately leading to long-term human-AI co-evolution.

To ensure that co-creativity fosters human flourishing, we must evaluate the Onion Model through the lens of human agency and Self-Determination Theory. This lens posits that well-being relies on satisfying our needs for autonomy, competence, and relatedness (Bingley et al. 2023; Passalacqua et al. 2025). Each layer of the model serves as a battleground for these needs. At the foundational level, systems must be designed for mutualistic symbiosis rather than parasitic automation that displaces human skills (Henriksen and Blond 2023). In the intermediate layers, transparency and trust are required to grant users meaningful human control, protecting them from automation complacency and cognitive atrophy that arise when AI bypasses human cognitive struggle (Shin and Shin 2023; Lee et al. 2025). Finally, at the core, sustained interaction must be carefully steered so that AI acts as extended cognition that expands human capabilities, rather than an unchecked force that colonizes the human lifeworld and subtly dictates creative choices (Hernández-Orallo 2025; Bozdag 2023).

To translate the Co-Evolutionary Onion from a theoretical model into actionable practice, the design of mixed-initiative co-creative systems must move beyond optimizing for frictionless efficiency and adopt layer-specific recommendations. At the foundation, systems must be architected explicitly for Intelligence Augmentation, ensuring that AI automates low-level execution while humans retain high-level orchestration and agency (Henriksen and Blond 2023). To activate the enablers of trust and transparency, designers should implement bidirectional Explainable AI that not only communicates the AI confidence and reasoning to the user but also allows users to explicitly communicate their own constraints and context back to the system (Iftikhar et al. 2024). Within the reciprocal input cycle, interfaces should move beyond single-shot prompting to support structured exploration of the AI latent design space, encouraging comparative evaluation and synthesis rather than the passive acceptance of a single output (Suh et al. 2024). The dynamic interaction layer requires designing for productive friction, allowing AI to occasionally act as a provoking agent that challenges human biases and stimulates divergent thinking, rather than merely generating pleasing responses that lead to creative convergence (Moruzzi and Margarido 2024). Finally, to achieve true co-evolution at the core, systems must employ continuous feedback architectures that capture both explicit interactions and implicit behavioral cues. This ensures the AI adapts to the user's evolving expertise, preventing cognitive complacency and fostering a sustainable, mutualistic symbiosis that continuously elevates human creative potential (Jarrahi et al. 2023).

The question of whether AI will light up or replace human creativity depends on the architecture of our engagement. The Human-AI Co-Evolutionary framework proposes that Co-Evolution cannot exist in a vacuum. It requires the protective layers of a Human-Centered Socio-technical System, the enabling mechanism of Trust and Transparency, and the explicit recognition of complementary Human and AI capabilities. By designing the dynamic interaction layer to foster productive friction and Mutual Theory of Mind, we move beyond the static automation of tasks toward a future where human and machine intelligence continuously reshape and elevate one another.

References

Alvarez, A.; Grevillius, E.; Olsson, E.; and Font, J. 2021. Questgram [Qg]: Toward a Mixed-Initiative Quest Generation Tool. In *The 16th International Conference on the Foundations of Digital Games (FDG) 2021*, 1–10. Montreal QC Canada: ACM. ISBN 978-1-4503-8422-3.

Ang, K.; Sankaran, S.; Liu, D.; and Scales, J. 2024. Embracing Levin's Legacy: Advancing Socio-Technical Learning and Development in Human-Robot Team Design Through STS Approaches. *Systemic Practice and Action Research*, 37(6): 661–678.

Berretta, S.; Tausch, A.; Ontrup, G.; Gilles, B.; Peifer, C.; and Kluge, A. 2023. Defining human-AI teaming the human-centered way: a scoping review and network analysis. *Frontiers in Artificial Intelligence*, 6: 1250725.

Bickley, S. J.; Macintyre, A.; and Torgler, B. 2025. Artificial Intelligence and Big Data in Sustainable Entrepreneurship. *Journal of Economic Surveys*, 39(1): 103–145.

Bingley, W. J.; Haslam, S. A.; Steffens, N. K.; Gillespie, N.; Worthy, P.; Curtis, C.; Lockey, S.; Bialkowski, A.; Ko, R. K.; and Wiles, J. 2023. Enlarging the model of the human at the heart of human-centered AI: A social self-determination model of AI system impact. *New Ideas in Psychology*, 70: 101025.

Boden, M. A. 2004. *The Creative Mind*. Routledge, 0 edition. ISBN 978-1-134-37958-3.

Bown, O.; Grace, K.; Ocampo Blanco, R.; and Ibarrola, F. 2025. "Coffee So Good, You'll Want to Slap Your Barista": Evaluating AI Co-Creation Through Dialogic Actions. *Creativity Research Journal*, 1–17.

Bozdag, A. A. 2023. AIsmosis and the pas de deux of human-AI interaction: Exploring the communicative dance between society and artificial intelligence. *Online Journal of Communication and Media Technologies*, 13(4): e202340.

Buschek, D. 2024. Collage is the New Writing: Exploring the Fragmentation of Text and User Interfaces in AI Tools. In *Designing Interactive Systems Conference*, 2719–2737. Copenhagen Denmark: ACM. ISBN 979-8-4007-0583-0.

Carter, W.; and Wynne, K. T. 2024. Integrating artificial intelligence into team decision-making: Toward a theory of AI–human team effectiveness. *European Management Review*, emre.12685.

Chakrabarty, T.; Padmakumar, V.; Brahman, F.; and Muresan, S. 2024. Creativity Support in the Age of Large Language Models: An Empirical Study Involving Professional Writers. In *Creativity and Cognition*, 132–155. Chicago IL USA: ACM. ISBN 979-8-4007-0485-7.

Chowdhury, S.; Dey, P.; Joel-Edgar, S.; Bhattacharya, S.; Rodriguez-Espindola, O.; Abadie, A.; and Truong, L. 2023. Unlocking the value of artificial intelligence in human resource management through AI capability framework. *Human Resource Management Review*, 33(1): 100899.

Constantinides, M.; Tahaei, M.; Quercia, D.; Stumpf, S.; Madaio, M.; Kennedy, S.; Wilcox, L.; Vitak, J.; Cramer, H.; Bogucka, E. P.; Baeza-Yates, R.; Luger, E.; Holbrook, J.; Muller, M.; Blumenfeld, I. G.; and Pistilli, G. 2024. Implications of Regulations on the Use of AI and Generative AI for Human-Centered Responsible Artificial Intelligence. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–4. Honolulu HI USA: ACM. ISBN 979-8-4007-0331-7.

Darban, M. 2024. Navigating virtual teams in generative AI-led learning: The moderation of team perceived virtuality. *Education and Information Technologies*, 29(17): 23225–23248.

Davis, N. 2013. Human-Computer Co-Creativity: Blending Human and Computational Creativity. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 9(6): 9–12.

Davis, N. 2021. Human-Computer Co-Creativity: Blending Human and Computational Creativity. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 9(6): 9–12.

Davis, N.; Hsiao, C.-P.; Yashraj Singh, K.; Li, L.; and Magerko, B. 2016. Empirically Studying Participatory Sense-Making in Abstract Drawing with a Co-Creative Cognitive Agent. In *Proceedings of the 21st International Conference on Intelligent User Interfaces*, 196–207. Sonoma California USA: ACM. ISBN 978-1-4503-4137-0.

Davis, N.; Sherson, J.; and Rafner, J. 2025. The Co-Creative Design Framework for Hybrid Intelligence. In *Proceedings of the 2025 Conference on Creativity and Cognition*, 560–572. Virtual United Kingdom: ACM. ISBN 979-8-4007-1289-0.

De Visser, E. J.; Peeters, M. M. M.; Jung, M. F.; Kohn, S.; Shaw, T. H.; Pak, R.; and Neerincx, M. A. 2020. Towards a Theory of Longitudinal Trust Calibration in Human–Robot Teams. *International Journal of Social Robotics*, 12(2): 459–478.

Dell'Acqua, F.; McFowland, E.; Mollick, E. R.; Lifshitz-Assaf, H.; Kellogg, K.; Rajendran, S.; Krayer, L.; Candelon, F.; and Lakhani, K. R. 2023. Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *SSRN Electronic Journal*.

Deshpande, M.; Park, J.; Pait, S.; and Magerko, B. 2024. Perceptions of Interaction Dynamics in Co-Creative AI: A Comparative Study of Interaction Modalities in Drawcto. In *Creativity and Cognition*, 102–116. Chicago IL USA: ACM. ISBN 979-8-4007-0485-7.

Deshpande, M.; Trajkova, M.; Knowlton, A.; and Magerko, B. 2023. Observable Creative Sense-Making (OCSM): A Method For Quantifying Improvisational Co-Creative Interaction. In *Creativity and Cognition*, 103–115. Virtual Event USA: ACM. ISBN 979-8-4007-0180-1.

Erengin, T.; Briker, R.; and De Jong, S. B. 2025. You, Me, and the AI: The Role of Third-Party Human Teammates for Trust Formation Toward AI Teammates. *Journal of Organizational Behavior*, job.2857.

Franceschelli, G.; and Musolesi, M. 2024. Creativity and Machine Learning: A Survey. *ACM Computing Surveys*, 56(11): 1–41.

Fuchs, A.; Passarella, A.; and Conti, M. 2023. Modeling, Replicating, and Predicting Human Behavior: A Survey. *ACM Transactions on Autonomous and Adaptive Systems*, 18(2): 1–47.

Garcia, M. B. 2024. The Paradox of Artificial Creativity: Challenges and Opportunities of Generative AI Artistry. *Creativity Research Journal*, 1–14.

Georganta, E.; and Ulfert, A. 2024. Would you trust an AI team member? Team trust in human–AI teams. *Journal of Occupational and Organizational Psychology*, 97(3): 1212–1241.

Geyer, W.; Chilton, L. B.; Kumar, R.; and Kalai, A. T. 2020. HAI-GEN 2020: Workshop on Human-AI Co-Creation with Generative Models. In *Companion Proceedings of the 25th International Conference on Intelligent User Interfaces*, 13–14. Cagliari Italy: ACM. ISBN 978-1-4503-7513-9.

Goodarzi, M. 2025. Co-Creativity With AI: A Human-Centered Approach.

Grammenos, D.; and Lubart, T. 2025. Peeping at creAItivity through a keyhole: creative self-perceptions, potential, and enhancement of GenAI chatbots. *Artificial Intelligence Review*, 58(10): 293.

Henriksen, A.; and Blond, L. 2023. Executive-centered AI? Designing predictive systems for the public sector. *Social Studies of Science*, 53(5): 738–760.

Hernández-Orallo, J. 2025. Enhancement and assessment in the AI age: An extended mind perspective. *Journal of Pacific Rim Psychology*, 19: 18344909241309376.

Hofman, J.; Goldstein, D. G.; and Rothschild, D. M. 2023. Steroids, Sneakers, Coach: The Spectrum of Human-AI Relationships. *SSRN Electronic Journal*.

Ibarrola, F.; Lulham, R.; and Grace, K. 2024. Affect-Conditioned Image Generation. *IEEE Transactions on Affective Computing*, 15(4): 2169–2179.

Iftikhar, R.; Chiu, Y.-T.; Khan, M. S.; and Caudwell, C. 2024. Human–Agent Team Dynamics: A Review and Future Research Opportunities. *IEEE Transactions on Engineering Management*, 71: 10139–10154.

Jarrahi, M. H.; Askay, D.; Eshraghi, A.; and Smith, P. 2023. Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*, 66(1): 87–99.

Kang, H. B.; Lin, D. C.-E.; Chen, Y.-Y.; Hong, M. K.; Martelaro, N.; and Kittur, A. 2025. BioSpark: Beyond Analogical Inspiration to LLM-augmented Transfer. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–29. Yokohama Japan: ACM. ISBN 979-8-4007-1394-1.

Kumar, H.; Vincentius, J.; Jordan, E.; and Anderson, A. 2025. Human Creativity in the Age of LLMs: Randomized Experiments on Divergent and Convergent Thinking. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–18. Yokohama Japan: ACM. ISBN 979-8-4007-1394-1.

Lamb, C.; Brown, D. G.; and Clarke, C. L. A. 2019. Evaluating Computational Creativity: An Interdisciplinary Tutorial. *ACM Computing Surveys*, 51(2): 1–34.

Lawton, T.; Grace, K.; and Ibarrola, F. J. 2023. When is a Tool a Tool? User Perceptions of System Agency in Human–AI Co-Creative Drawing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, 1978–1996. Pittsburgh PA USA: ACM. ISBN 978-1-4503-9893-0.

Lee, H.-P. H.; Sarkar, A.; Tankelevitch, L.; Drosos, I.; Rintel, S.; Banks, R.; and Wilson, N. 2025. The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–22. Yokohama Japan: ACM. ISBN 979-8-4007-1394-1.

Li, J.; Cao, H.; Lin, L.; Hou, Y.; Zhu, R.; and El Ali, A. 2024a. User Experience Design Professionals' Perceptions of Generative Artificial Intelligence. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–18. Honolulu HI USA: ACM. ISBN 979-8-4007-0330-0.

Li, J.; Huang, J.; Liu, J.; and Zheng, T. 2022. Human-AI cooperation: Modes and their effects on attitudes. *Telematics and Informatics*, 73: 101862.

Li, Y.; Li, Y.; Chen, Q.; and Chang, Y. 2024b. Humans as teammates: The signal of human–AI teaming enhances consumer acceptance of chatbots. *International Journal of Information Management*, 76: 102771.

Lin, D. C.-E.; Kang, H. B.; Martelaro, N.; Kittur, A.; Chen, Y.-Y.; and Hong, M. K. 2025. Inkspire: Supporting Design Exploration with Generative AI through Analogical Sketching. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–18. Yokohama Japan: ACM. ISBN 979-8-4007-1394-1.

Liu, Y.; Chen, S.; Cheng, H.; Yu, M.; Ran, X.; Mo, A.; Tang, Y.; and Huang, Y. 2024. How AI Processing Delays Foster Creativity: Exploring Research Question Co-Creation with an LLM-based Agent. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–25. Honolulu HI USA: ACM. ISBN 979-8-4007-0330-0.

Lopes, P.; and Yannakakis, G. 2021. Investigating Collaborative Creativity via Machine-Mediated Game Blending. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 10(6): 10–13.

Lyons, J. B.; Hobbs, K.; Rogers, S.; and Clouse, S. H. 2023. Responsible (use of) AI. *Frontiers in Neuroergonomics*, 4: 1201777.

Lyu, Y.; Wang, X.; Lin, R.; and Wu, J. 2022. Communication in Human–AI Co-Creation: Perceptual Analysis of Paintings Generated by Text-to-Image System. *Applied Sciences*, 12(22): 11312.

Margarido, S.; Machado, P.; Roque, L.; and Martins, P. 2026. Boosting Mixed-Initiative Co-Creativity in Game Design: A Tutorial. *ACM Computing Surveys*, 58(2): 1–36.

Mirowski, P.; Love, J.; Mathewson, K.; and Mohamed, S. 2024. A Robot Walks into a Bar: Can Language Models Serve as Creativity SupportTools for Comedy? An Evaluation of LLMs' Humour Alignment with Comedians. In *The 2024 ACM Conference on Fairness Accountability and Transparency*, 1622–1636. Rio de Janeiro Brazil: ACM. ISBN 979-8-4007-0450-5.

Mirowski, P.; Mathewson, K. W.; Pittman, J.; and Evans, R. 2023. Co-Writing Screenplays and Theatre Scripts with Language Models: Evaluation by Industry Professionals. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–34. Hamburg Germany: ACM. ISBN 978-1-4503-9421-5.

Moruzzi, C.; and Margarido, S. 2024. A User-centered Framework for Human-AI Co-creativity. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–9. Honolulu HI USA: ACM. ISBN 979-8-4007-0331-7.

Mylrea, M.; and Robinson, N. 2023. Artificial Intelligence (AI) Trust Framework and Maturity Model: Applying an Entropy Lens to Improve Security, Privacy, and Ethical AI. *Entropy*, 25(10): 1429.

O'Neill, T. A.; Flathmann, C.; McNeese, N. J.; and Salas, E. 2023. 21st Century teaming and beyond: Advances in human-autonomy teamwork. *Computers in Human Behavior*, 147: 107865.

Palani, S.; and Ramos, G. 2024. Evolving Roles and Workflows of Creative Practitioners in the Age of Generative AI. In *Creativity and Cognition*, 170–184. Chicago IL USA: ACM. ISBN 979-8-4007-0485-7.

Parasuraman, R.; Sheridan, T.; and Wickens, C. 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3): 286–297.

Passalacqua, M.; Pellerin, R.; Yahia, E.; Magnani, F.; Rosin, F.; Joblot, L.; and Léger, P.-M. 2025. Practice With Less AI Makes Perfect: Partially Automated AI During Training Leads to Better Worker Motivation, Engagement, and Skill Acquisition. *International Journal of Human–Computer Interaction*, 41(4): 2268–2288.

Pedreschi, D.; Pappalardo, L.; Ferragina, E.; Baeza-Yates, R.; Barabási, A.-L.; Dignum, F.; Dignum, V.; Eliassi-Rad, T.; Giannotti, F.; Kertész, J.; Knott, A.; Ioannidis, Y.; Lukowicz, P.; Passarella, A.; Pentland, A. S.; Shawe-Taylor, J.; and Vespignani, A. 2025. Human-AI coevolution. *Artificial Intelligence*, 339: 104244.

Rezwana, J.; and Maher, M. L. 2023. Designing Creative AI Partners with COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems. *ACM Transactions on Computer-Human Interaction*, 30(5): 1–28.

Robertson, J.; Ferreira, C.; Botha, E.; and Oosthuizen, K. 2024. Game changers: A generative AI prompt protocol to enhance human-AI knowledge co-construction. *Business Horizons*, 67(5): 499–510.

Ross, S. I.; Martinez, F.; Houde, S.; Muller, M.; and Weisz, J. D. 2023. The Programmer's Assistant: Conversational Interaction with a Large Language Model for Software Development. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*, 491–514. Sydney NSW Australia: ACM. ISBN 979-8-4007-0106-1.

Sasson Lazovsky, G.; Raz, T.; and Kenett, Y. N. 2025. The Art of Creative Inquiry—From Question Asking to Prompt Engineering. *The Journal of Creative Behavior*, 59(1): e671.

Shin, D.; and Shin, E. Y. 2023. Human-Centered AI: A Framework for Green and Sustainable AI. *Computer*, 56(6): 16–25.

Siemon, D. 2022. Elaborating Team Roles for Artificial Intelligence-based Teammates in Human-AI Collaboration. *Group Decision and Negotiation*, 31(5): 871–912.

Simón, C.; Revilla, E.; and Jesús Sáenz, M. 2024. Integrating AI in organizations for value creation through Human-AI teaming: A dynamic-capabilities approach. *Journal of Business Research*, 182: 114783.

Smith, A.; Van Wagoner, H. P.; Keplinger, K.; and Celebi, C. 2025. Navigating AI Convergence in Human–Artificial Intelligence Teams: A Signaling Theory Approach. *Journal of Organizational Behavior*, job.2856.

Suh, M. M.; Youngblom, E.; Terry, M.; and Cai, C. J. 2021. AI as Social Glue: Uncovering the Roles of Deep Generative AI during Social Music Composition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–11. Yokohama Japan: ACM. ISBN 978-1-4503-8096-6.

Suh, S.; Chen, M.; Min, B.; Li, T. J.-J.; and Xia, H. 2024. Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–26. Honolulu HI USA: ACM. ISBN 979-8-4007-0330-0.

Tang, M.; Hofreiter, S.; Werner, C. H.; Zielińska, A.; and Karwowski, M. 2024. "Who" Is the Best Creative Thinking Partner? An Experimental Investigation of Human–Human, Human–Internet, and Human–AI Co-Creation. *The Journal of Creative Behavior*, jocb.1519.

Van Den Bosch, K.; Schoonderwoerd, T.; Blankendaal, R.; and Neerincx, M. 2019. Six Challenges for Human-AI Co-learning. In Sottilare, R. A.; and Schwarz, J., eds., *Adaptive Instructional Systems*, volume 11597, 572–589. Cham: Springer International Publishing. ISBN 978-3-030-22340-3 978-3-030-22341-0. Series Title: Lecture Notes in Computer Science.

Van Den Broek, E.; Sergeeva, A.; and Huysman Vrije, M. 2021. When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring. *MIS Quarterly*, 45(3): 1557–1580.

Vinchon, F.; Gironnay, V.; and Lubart, T. 2024. GenAI Creativity in Narrative Tasks: Exploring New Forms of Creativity. *Journal of Intelligence*, 12(12): 125.

Wang, Q.; Walsh, S.; Si, M.; Kephart, J.; Weisz, J. D.; and Goel, A. K. 2024a. Theory of Mind in Human-AI Interaction. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–6. Honolulu HI USA: ACM. ISBN 979-8-4007-0331-7.

Wang, X.; Abubaker, S. M.; Babalola, G. T.; and Tulk Jesso, S. 2024b. Co-Designing an AI Chatbot to Improve Patient Experience in the Hospital: A human-centered design case study of a collaboration between a hospital, a university, and ChatGPT. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 1–10. Honolulu HI USA: ACM. ISBN 979-8-4007-0331-7.

Wiggins, G. A. 2006. Searching for computational creativity. *New Generation Computing*, 24(3): 209–222.

Wu, L.; Fan, A.; Yang, Y.; and He, Z. 2021. Robotic involvement in the service encounter: a value-centric experience framework and empirical validation. *Journal of Service Management*, 32(5): 783–812.

Yang, S. C.; Folke, T.; and Shafto, P. 2023. The Inner Loop of Collective Human–Machine Intelligence. *Topics in Cognitive Science*, tops.12642.

Zheng, Y.; Rowell, B.; Chen, Q.; Kim, J. Y.; Kontar, R. A.; Yang, X. J.; and Lester, C. A. 2023. Designing Human-Centered AI to Prevent Medication Dispensing Errors: Focus Group Study With Pharmacists. *JMIR Formative Research*, 7: e51921.