

# Human-AI Co-Design and Co-Creation: A Review of Emerging Approaches, Challenges, and Future Directions

Nyasha Kadenhe<sup>1</sup>, Mohamed Al Musleh<sup>2</sup>, Allan Lompot<sup>2</sup>

<sup>1</sup>Academic Quality, Heriot-Watt University Dubai Campus

<sup>2</sup>School of Textiles and Design, Heriot-Watt University Dubai Campus  
n.kadenhe@hw.ac.uk, m.al-musleh@hw.ac.uk, a.lompot@hw.ac.uk

## Abstract

The integration of Artificial Intelligence (AI) into creative and design processes has shifted from automation towards co-creation, positioning AI as a collaborative partner rather than a replacement. As AI-driven tools become more embedded in human-centred design, understanding their impact on interaction dynamics, ethics, and usability is critical. This review examines key advancements in human-AI co-design and co-creation fields, focusing on interaction frameworks, ethical considerations, non-linear collaboration models, domain-specific applications, and user experience (UX) design. Recent research emphasises the need for structured frameworks that facilitate effective communication and partnership between humans and AI in creative tasks. Mixed-initiative and explainable AI (XAI) approaches play a crucial role in enhancing transparency and interpretability, allowing designers to co-create with greater trust and autonomy. Ethical concerns, such as AI's influence on user perception and decision-making, are also gaining prominence, calling for responsible AI deployment in co-creative settings. Additionally, non-linear collaboration models redefine AI's role as an adaptive assistant throughout iterative design stages, aligning with the dynamic nature of creative processes. Domain-specific applications, ranging from game and product design to choreography and smart manufacturing, illustrate the versatility of AI in augmenting human creativity. AI-assisted UX design further extends this impact by personalising user experiences and streamlining workflows, ultimately improving efficiency and engagement. Despite these advancements, challenges remain in balancing AI autonomy with human control, evaluating its impact on creative workflows, and developing inclusive methodologies that cater to diverse design disciplines. This review synthesises current research trends and identifies future directions for designing AI systems that empower, rather than replace, human expertise in creative industries.

## Introduction

The integration of AI into creative and design processes has undergone a significant transformation, from being a tool for automation to becoming a collaborative partner in co-creation. Historically, AI was primarily employed to automate repetitive tasks, optimising efficiency while minimising human intervention (Brynjolfsson and McAfee, 2014).

Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

However, as AI capabilities have advanced, particularly with the rise of deep learning (LeCun et al., 2015), its role has expanded to include generative design, ideation support, and adaptive decision-making, fostering a more dynamic interaction between humans and machines.

This shift reflects a broader movement toward human-AI co-design, where AI augments rather than replaces human creativity (Shneiderman, 2020). Modern AI-driven tools, such as ChatGPT, Midjourney, and Autodesk Dreamcatcher, exemplify this evolution by enabling iterative, non-linear collaboration—where human intuition and AI-generated insights work in tandem (Dove et al., 2017; Chen et al., 2023). Yet, as AI becomes more embedded in creative workflows, critical challenges emerge, including:

- Trust and transparency (XAI)
- Ethical concerns (bias, authorship, over-reliance)
- Balancing AI autonomy with human control
- Adapting AI to diverse creative disciplines

This paper examines the current state of human-AI co-creation, synthesising research on interaction frameworks, ethical considerations, domain-specific applications, and future challenges. It explores how mixed-initiative systems (Amershi et al., 2014) and human-centred AI (HCAI) principles (Shneiderman, 2020) can enhance collaboration while addressing risks such as cognitive overload (Chiou et al., 2021) and diminished creative agency (Guo et al., 2023). Additionally, the emerging regulatory needs are highlighted such as the EU AI Act and propose directions for more adaptive, ethically grounded AI design practices.

By analysing advancements and unresolved issues in human-AI co-design, this review aims to inform the development of AI systems that empower creativity while preserving human oversight, trust, and ethical integrity in the evolving landscape of collaborative intelligence.

## Theoretical and Conceptual Foundations

The trajectory of AI in creative domains has evolved considerably, transitioning from paradigms centred on automation to models that emphasize co-creation. Traditional automation focused on the replacement of repetitive human tasks through rule-based systems with the aim of maximising efficiency (Brynjolfsson and McAfee, 2014). These systems operated with minimal adaptability and lacked contextual awareness, often relegating human users to passive oversight

or administrative roles. Recent research has highlighted that such automation-centric models are increasingly insufficient in domains requiring creativity, adaptability, and user input (Shneiderman, 2020; Dellermann et al., 2019). The shift toward co-creation reflects a broader understanding that effective human-AI systems must support dynamic interaction, shared goals, and the augmentation, not replacement, of human capabilities.

The rise of machine learning, particularly deep learning, enabled more complex forms of intelligent automation (LeCun et al., 2015). These advances allowed AI systems to go beyond rule-based execution and engage in pattern recognition, prediction, and generative capabilities, thus opening the door to creative and design-oriented applications. Despite these capabilities, early AI systems remained largely task driven. In contrast, emerging co-creation frameworks position humans and AI as collaborative partners, where both parties actively contribute to ideation, iteration, and refinement (Dove et al., 2017).

New co-creative tools like Autodesk Dreamcatcher, ChatGPT, and AIVA demonstrate how AI-generated outputs can be guided, adapted, and reinterpreted by human input, illustrating this shift from automation to co-creation. This evolution reflects a broader transition toward participatory design and human-AI partnerships (Shneiderman, 2020). The shift marks a foundational change in how humans and machines interact—from one-way execution to two-way collaboration (Chen et al., 2021). As AI becomes more embedded in creative and decision-making tasks, concerns around transparency, accountability, and alignment with human values have become increasingly important.

This progression leads naturally into the principles of HCAI and XAI, which aim to ensure that AI systems are not only powerful but also understandable and supportive of human agency. HCAI emphasises the development of intelligent systems that align with human values and support human decision-making processes (Shneiderman, 2020). In co-creative contexts, this ensures that AI augments rather than hinders creativity. Explainable AI complements this vision by addressing the “black box” nature of many AI models (Doshi-Velez and Kim, 2017). It is especially critical in co-creation, where users need to understand and assess AI-generated outputs to maintain trust and creative control (Wang et al., 2019). While balancing performance and interpretability remains a challenge, promising techniques such as counterfactual reasoning, attention visualisation, and interactive explanations are emerging (Guidotti et al., 2018).

Rezwana and Maher (Rezwana and Maher, 2022) argued that while explainability and human-centeredness help users interpret and trust AI outputs, effective collaboration also requires dynamic interaction—where control can shift fluidly between human and machine. This need for adaptable and reciprocal engagement leads directly into the idea of mixed-initiative systems, which go beyond traditional user-AI boundaries.

Mixed-initiative systems support shared control, allowing both human and AI agents to proactively shape the course of a task (Wang et al., 2019). Amershi et al. (Amershi et al., 2014) highlighted that these systems are particularly rel-

evant for creative processes, where ideation and revision are iterative and dynamic. Their research concluded that by enabling fluid turn-taking and decision-making, mixed-initiative design tools create conditions for richer human-AI collaboration.

However, recent studies have identified critical challenges, including role ambiguity, cognitive overload, and control conflicts between users and AI agents. Chiou et al. (2021), for instance, found that poorly designed control dynamics can lead to user confusion and reduced collaboration effectiveness. Liao et al. (Liao et al., 2020) similarly noted that users may either over-rely on the system or disengage entirely when boundaries of control are unclear.

Despite these challenges, mixed-initiative systems are being deployed across diverse domains such as storytelling (Li et al., 2022), code generation (Vaithilingam et al., 2022), and visual design (Bharadhwaj et al., 2021). Yet, persistent design hurdles remain—especially in managing initiative, modelling users effectively, and sustaining engagement without overwhelming the user (Wang et al., 2019).

As these systems become increasingly prevalent, the complexity of coordinating human and AI contributions intensifies (McGrath et al., 2024). This necessitates robust interaction frameworks that facilitate fluid dialogue, shared context, and mutual understanding, without which meaningful collaboration may falter (McGrath et al., 2024; Pyae, 2025). Trust is foundational to sustainable and effective human-AI collaboration, influenced by factors such as reliability, transparency, feedback quality, and perceived agency (Pyae., 2025). Recent work emphasises that trust is dynamic and context-dependent, evolving over time through continuous interaction and mutual adaptation (McGrath et al., 2024). For example, the CHAI-T framework actively manages trust in collaborative human-AI teams by accounting for task specificity and evolving team dynamics (McGrath et al., 2024), while the Human-AI Handshake Model advocates bidirectional information exchange and capability augmentation to foster balanced interaction (Pyae., 2025). Additionally, the Human-Centred Trust Framework (HCTFrame) provides guidelines for mapping user trust and addressing socio-ethical considerations in AI system design (Sousa et al., 2023). Co-adaptive systems, which evolve alongside users, have been proposed to support long-term trust and sustained engagement by adapting to user behaviours and preferences [4]. Design strategies such as progressive disclosure, explanation on demand, and controllability are crucial in fostering trust, and trust-aware AI design continues to emerge as a research frontier in domains like education, healthcare, and the creative industries (World Health Organization., 2025). As human-AI collaboration deepens, trust will remain central to ensuring these partnerships are not only technologically functional but also ethically aligned and user-driven.

## Domain-Specific Applications and Non-Linear Collaboration

Human-AI co-creation has shown particular promise in disciplines where collaborative methods support and expand

creative outcomes. In product design and user experience, generative AI tools have been recognised for their role in supporting early-stage ideation and concept evaluation. For example, Chen et al. (Chen et al., 2025) found that a notable proportion of participants appreciated the use of generative AI during idea development. These tools are increasingly viewed not only as functional aids but as active contributors that shape design directions based on large datasets and pattern recognition (Ogundipe et al., 2024).

The involvement of AI in design is best understood as adaptable, shifting depending on the stage of the process. During early ideation, tools like ChatGPT and other generative models assist in brainstorming or producing initial drafts, allowing designers to concentrate on decision-making and conceptual refinement (Sun et al., 2024). Rather than replacing the designer's creative judgment, these tools help structure and extend the thinking process.

Design practices often benefit from iterative approaches that allow flexibility and the ability to revisit decisions. Non-linear workflows are particularly relevant in creative contexts, as they reflect the way ideas shift and develop over time (Baltà-Salvador et al., 2025). For instance, Baltà-Salvador et al. (Baltà-Salvador et al., 2025) suggest that initiating ideation independently from AI can promote original thinking, with AI introduced later to extend or refine those ideas. This approach supports a fluid feedback loop, where changes are informed by user responses or shifting project aims (Ege et al., 2024).

Karadağ and Ozar (Karadağ and Ozar, 2025) explored the integration of AI, specifically text-to-image (T2I) generators like Midjourney, into the conceptual design phase of interior design education. Conducted over a 14-week design studio course, the study involved structured workshops, surveys, and interviews with senior undergraduate students. Findings indicate that AI-supported visualisation enhanced ideation, accelerated design iteration, and improved conceptual precision. Students experimented with different prompt strategies—from simple keywords to complex narratives—revealing that layered, progressive prompts yielded the most meaningful visual outputs. While students broadly embraced AI as a valuable tool for inspiration and development, concerns were raised about over-reliance, ethical issues, and the risk of diminished creative autonomy. The study underscores the importance of thoughtful prompt design and highlights AI's potential as a collaborative partner in design education.

In domains requiring rapid prototyping, such as game development or interactive design, this iterative structure allows for diverse concept testing and short design cycles. Ege et al. (Ege et al., 2024) note that such workflows encourage experimentation, enabling teams to test and adapt creative choices efficiently. Furthermore, the use of AI in these cycles supports associative thinking and helps build a layered understanding of a project, contributing incrementally as the design progresses (Li et al., 2022; Ege et al., 2024).

In manufacturing, AI integration has been linked to improved team collaboration and information sharing, especially in complex design and production phases (Li et al., 2022). Grech et al. (Grech et al., 2023) highlight that AI

can help coordinate communication across teams, supporting smoother workflows during iterative design tasks. Similarly, in fields such as choreography and game design, adaptive AI systems offer suggestions that align with evolving creative goals, making them useful in settings where real-time responsiveness is valued (Grech et al., 2023).

In collaborative settings, Makokha (Makokha, 2023) describes how AI can act as a form of constructive interruption, providing suggestions or reframing problems based on group dynamics. This perspective frames AI as a participant in the creative dialogue, contributing through observation, feedback, or alternative perspectives. Across different design phases, AI tools can therefore shift roles—supporting idea generation, helping with iteration, or prompting reflection—depending on how they are applied within the team's workflow.

Rezwana and Maher (Rezwana and Maher, 2022) investigated the role of interaction design in human-AI co-creative systems, specifically examining how AI-to-human communication influences collaborative experience, user engagement, and perception of AI. Their comparative study, involving 38 participants using two prototypes of a sketch-based design tool, demonstrated that systems enabling two-way communication (via text, speech, and a visual avatar) were perceived as more collaborative, engaging, and intelligent compared to systems that relied solely on human-to-AI input. Participants reported higher trust, creativity, and satisfaction when the AI responded affectively and collected feedback, suggesting that communication design, beyond algorithmic ability, plays a critical role in shaping effective and enjoyable human-AI creative partnerships.

Recent work by Lee et al. (Lee et al., 2025) offers a structured analysis of how artificial intelligence (AI) supports the design process by mapping existing AI-based Design Support Systems (AI-DSS) to the Double Diamond model. Their systematic review reveals that the majority of AI-DSS focus on the later stages of the design process—Develop and Deliver—where solution generation and evaluation occur. In contrast, relatively few tools support the early phases of Discover and Define, where designers engage in problem framing, user understanding, and conceptual exploration. Notably, their case study involving ChatGPT and a robotic arm illustrates that expert designers perceive valuable AI roles in the early stages, such as aiding with design inspiration, defining constraints, and structuring ambiguous information. This study highlights a disconnect between current system capabilities and designers' expectations, pointing to the need for future AI-DSS that can meaningfully engage with open-ended, ill-structured design problems. Their work advocates for a shift toward more interpretive and collaborative human-AI interactions across the full design cycle.

## Challenges, Ethics, and Future Directions

AI has evolved into a powerful design paradigm, enabling humans to boost creativity and accelerate design through co-creation (Mao et al., 2023; Demirel et al., 2024). However, this progress brings ethical concerns, challenges and the need to balance human oversight with AI autonomy, along with an increasing demand for a robust regulatory

framework and focus on AI's future development. AI integration continues to pose substantial challenges, especially for designers. Excessive reliance on AI particularly during the early ideation stages of design can lead to diminishing the depth of designers' critical thinking and weakening the originality of conceptual outcomes (Guo et al., 2023).

Moreover, many existing generative design tools fall short in incorporating human factors, limiting the ability to account for the full spectrum of human capabilities, constraints and emotional responses that are essential consideration for advancing truly human-centred product and service innovation (Demirel et al., 2024). As a result, the adoption of AI in creative design industries remains particularly challenging, where professionals place high value on individual style and expression. These domains demand AI tools that are not only technically capable but also highly personalized and sensitive to the needs of creative practitioners (Mao et al., 2023). Another significant challenge lies in enhancing productivity through human-AI collaboration. A literature review by Weisz et al. (Weisz et al., 2023) found that many collaborative systems failed to achieve positive synergy. These collaborations, in some cases, even resulted in outcomes that were inferior to those produced by humans or AI alone. Ensuring the safe use of generative AI remains a pressing concern due to potential risks and harms, which may arise from both the training data used and the ways in which these systems are deployed. Gmeiner et al. (Gmeiner et al., 2023) revealed that designers often struggled to interpret the rationale behind AI-generated design features, uncertain whether outcomes were intentional or the result of algorithmic errors. They also faced challenges with control and authorship, often feeling AI dominated the process, leading some to avoid AI assistance entirely.

Working with emerging and complex technologies that differ fundamentally from conventional design materials introduces distinct difficulties. Designers are tasked to balance AI's expansive potential with essential human-centred considerations and interactional requirements (Murray-Rust et al., 2023). AI can act as a personalized tool for providing visual feedback and inspiration to designers, but its effectiveness relies on designers' ability to harness their creative agency (Wang et al., 2023). Achieving the right balance between control and autonomy is essential. Empowering designers with intuitive, meaningful control enhance efficiency and reinforces their creative identity. Increased user control fosters clearer understanding, stronger engagement and a greater sense of ownership over the design process (Weisz et al., 2023). In terms of ethics, ethical concerns related to AI have been present even before the term "artificial intelligence" was formally introduced as early as 1955. Today, there is heightened recognition of the importance of embedding social and ethical considerations into the development and deployment of emerging AI technologies. Key risk factors such as control, fairness, bias and transparency are gaining attention, along with newer issues like sustainability and the long-term implications of AI (Katirai and Nagato, 2024).

As AI-powered co-creative tools become more embedded in artistic and design practices, ethical considerations grow increasingly vital. Since AI generates content and in-

teracts directly with users, ethical concerns must be addressed throughout the design process. The open-ended nature of these interactions makes this complex, underscoring the need to understand user perspectives and expectations (Rezwana and Maher, 2023).

The rapid global rise of AI has sparked debate on appropriate regulatory responses (Shetty et al., 2025). While some countries adopt risk-based governance, others prioritize innovation or human rights. Amid this diversity, there is growing pressure to create regulations that balance public safety and safeguard public interest without stifling innovation (Almeida et al., 2023).

The EU AI Act is a major step toward comprehensive regulation, influencing global practices through the "Brussels Effect" (Ho et al., 2023). Meanwhile, bodies like NIST and ISO are crafting voluntary frameworks, and models like COSO ERM 2017 are being adapted for AI (Schuett, 2023). However, critics argue the EU AI Act overlooks the complexities of generative AI and lacks clarity on real-time systems like ChatGPT (Shetty et al. 2025; Chauhan, 2022). Current frameworks are often too rigid and require frequent updates (Schuett, 2023), while targeted measures are needed to address emerging risks (Anderljung, 2023). To address these challenges, scholars advocate for multi-level governance that embeds ethics and stakeholder input throughout AI development (Almeida et al. 2023; Birkstedt et al., 2023). Todorova et al. (Todorova et al., 2023) further stress the importance of supporting SMEs and start-ups in the regulatory landscape.

Despite the growing integration of AI into co-design and co-creation, substantial gaps remain in both research and practice. There is a pressing need for more empirical studies to explore how AI influences designers' cognitive processes, particularly in terms of when it supports or impedes creativity (Guo et al., 2023; Mao et al., 2023). As generative AI technologies are still in an early stage of development, novel co-creative user experiences are evolving quickly. This rapid pace of change raises the possibility that important design principles, strategies, or user goals may be overlooked. While existing principles offer valuable guidance for design decision-making, they must be validated through real-world applications to ensure their effectiveness, clarity, and relevance (Weisz et al., 2023). Future research should prioritize the development of design patterns and interaction models that support more effective human-AI collaborations, especially in creative domains where authorship, intent and interpretation are inherently subjective.

Although ethical concerns are recognized, there is a lack of actionable guidance on implementing ethical principles in co-design, such as transparency in decision-making, informed consent, and safeguarding creative agency. Finally, existing regulatory frameworks like the EU AI Act are too rigid for the dynamic nature of generative AI (Shetty et al., 2025). More adaptive, flexible regulatory approaches are needed to address the evolving nature of AI design and use.

## Conclusions

This review has charted the evolution of AI in creative domains from task automation to dynamic co-creation, reveal-

ing both transformative potential and unresolved tensions. The emergence of mixed-initiative systems and XAI has enabled richer collaboration, yet persistent gaps in role negotiation (Chiou et al., 2021), creative ownership (Gmeiner et al., 2023), and domain-specific adaptation (Lee et al., 2025) underscore the need for more nuanced frameworks.

Three critical frontiers demand attention:

1- Adaptive Control Mechanisms: Future systems must better negotiate initiative-sharing, particularly in early-stage ideation where AI's overreach may stifle originality (Guo et al., 2023). Techniques like context-aware prompting (Karadağ and Ozar, 2025) and human-in-the-loop refinement could recalibrate this balance.

2- Ethical by Design Co-Creation: Beyond current regulatory efforts such as EU AI Act, there's urgent need for embedded ethics—tools that surface biases in real-time and preserve creative provenance without workflow disruption (Rezwana and Maher, 2023).

3- Discipline-Specific AI Literacy: As AI permeates fields from choreography to manufacturing (Grech et al., 2023), tailored training must address how practitioners critically engage with AI outputs while retaining domain expertise.

The path forward lies not in perfecting AI's mimicry of human creativity, but in orchestrating its strengths—scale, pattern synthesis, iteration speed—to amplify rather than automate creative judgment. Achieving this requires co-evolution of technical systems, pedagogical approaches, and ethical standards, ensuring AI remains a lens for human imagination rather than its frame.

## References

Almeida; V. , Mendes, L.S., and Doneda, D., (2023) On the Development of AI Governance Frameworks. *IEEE Internet Computing*, 27(1).

Amershi; S. , Cakmak, M., Knox, W. B., and Kulesza, T. 2014. "Power to the People: The Role of Humans in Interactive Machine Learning," *AI Magazine* (35:4), pp. 105-120.

Anderljung; M. , Barnhart, J., Korinek, A., Leung, J., O'Keefe, C., Whittlestone, J. and Wolf, K., 2023. *Frontier AI regulation: Managing emerging risks to public safety*.

Baltà-Salvador; R. , El-Madafri, I., Brasó-Vives, E. and Peña, M., 2025. Empowering engineering students through artificial intelligence (AI): Blended human–AI creative ideation processes with ChatGPT. *Computer Applications in Engineering Education*, 33(1). Wiley.

Bharadhwaj; H. , Klemmer, S.R. and Agarwala, A., 2021. *DesignScape: Design with interactive layout suggestions*. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp.1–13.

Birkstedt; T. , Minkinen, M., Tandon, A. and Mäntymäki, M., 2023. AI governance: themes, knowledge gaps and future agendas. *Internet Research*, 33(7), pp.133-167.

Brynjolfsson; E. and McAfee, A. (2014) *The second machine age*: New York, N.Y.: W.W. Norton & Company.

Chen; L. , Song, Y., Jia, G., Sun, L., Childs, P. and Yin, Y., 2025. How generative AI supports human in conceptual design. *Design Science*, 11. Cambridge University Press.

Chen; Z. , Luo, Y. and Sra, M., 2025. Engaging with AI: How interface design shapes human-AI collaboration in high-stakes decision-making. *arXiv preprint arXiv:2501.16627*.

Chen; V. , Liao, V., Vaughan, J., and Bansal, G, 2023. Understanding the Role of Human Intuition on Reliance in Human-AI Decision-Making with Explanations. 10.48550/arXiv.2301.07255.

Chiou; E. K. and Lee, J.D, 2021. Trusting automation: Designing for responsivity and resilience. *Human Factors*, 2023-02, Vol.65 (1), p.137-165.

Chauhan; K. D.S., 2023. From 'What' and 'Why' to 'How': An imperative driven approach to mechanics of AI regulation. *Global Jurist*, 23(2), pp.99-124.

Dellermann; D. , Ebel, P., Söllner, M. and Leimeister, J.M., 2019. Hybrid intelligence. *Business & Information Systems Engineering*, 61(5), pp.637-643.

Demirel; H. O., Goldstein, M.H., Li, X. and Sha, Z., 2024. Human-centered generative design framework: An early design framework to support concept creation and evaluation. *International Journal of Human-Computer Interaction*, 40(4), pp.933–944.

Doshi-Velez; F. and Kim, B., 2017. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*.

Dove; G. , Halskov, K., Forlizzi, J. and Zimmerman, J., 2017. UX design innovation: Challenges for working with machine learning as a design material. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp.278–288.

Ege; D. N., Øvrebø, H.H., Stubberud, V., Berg, M.F., Steinert, M. and Vestad, H., 2024. Benchmarking AI design skills: Insights from ChatGPT's participation in a prototyping hackathon. *Proceedings of the Design Society*, 4, pp.1999-2008. Cambridge University Press.

Gmeiner; F. , Yang, H., Yao, L., Holstein, K. and Martelaro, N., 2023. Exploring challenges and opportunities to support designers in learning to co-create with AI-based manufacturing design tools. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, pp.1–20.

Grech; A. , Mehnen, J. and Wodehouse, A., 2023. An extended AI-experience: Industry 5.0 in creative product innovation. *Sensors*, 23(6), p.3009. MDPI.

Guidotti; R. , et al., 2018. A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), pp.1–42. DOI: 10.1145/3236009.

Guo; X. , Xiao, Y., Wang, J. and Ji, T., 2023. Rethinking designer agency: A case study of co-creation between designers and AI. In: De Sainz Molestina, D., Galluzzo, L., Rizzo, F. and Spallazzo, D., eds. *IASDR 2023: Life-Changing Design*, 9-13 October, Milan, Italy. DOI: 10.21606/iasdr.2023.478.

Hinsen; S. , Hofmann, P., Jöhnk, J. and Urbach, N., 2022. How can organizations design purposeful human-AI interactions: A practical perspective from existing use cases and interviews. In: *Proceedings of the 55th Hawaii International Conference on System Sciences (HICSS)*. University of Hawai'i at Mānoa, pp.1–10.

- Ho; L. , Barnhart, J., Trager, R., Bengio, Y., Brundage, M., Carnegie, A., Chowdhury, R., Dafoe, A., Hadfield, G., Levi, M. and Snidal, D. (2023) International institutions for advanced AI. arXiv preprint. Available at: <https://arxiv.org/abs/2307.04699>
- Katirai; A. and Nagato, Y., 2024. Addressing trade-offs in co-designing principles for ethical AI: Perspectives from an industry-academia collaboration. *AI and Ethics*.
- Karadağ; D. and Ozar, B. (2025) 'A new frontier in design studio: AI and human collaboration in conceptual design', *Frontiers of architectural research* [Preprint].
- LeCun; Y. , Bengio, Y. and Hinton, G. (2015) 'Deep learning', *Nature (London)*, 521(7553), pp. 436–444.
- Li; B. , Karpas, E., Koncel-Kedziorski, R. and Choi, Y., 2022. Tell me a story you like: Generating personalized stories with rewrites. In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*, pp.7647–7661.
- Li; N. , Yan, Y., Yang, Y. and Gu, A., 2022. Artificial intelligence capability and organizational creativity: The role of knowledge sharing and organizational cohesion. *Frontiers in Psychology*, 13, p.845277.
- Makokha; J. , 2023. Future AI systems: An AI outperforms a human as teammate. *TechRxiv*. DOI: 10.36227/techrxiv.21901407.v1.
- Mao; Y. , Rafner, J., Wang, Y. and Sherson, J., 2024. HITAM, a hybrid intelligence framework for training and adoption of generative design assistants. *Frontiers in Computer Science*, 6.
- McGrath; M. J., Duenser, A., Lacey, J. and Paris, C., 2024. Collaborative human-AI trust (CHAI-T): A process framework for active management of trust in human-AI collaboration. arXiv preprint arXiv:2404.01615.
- Murray-Rust; D. , Lupetti, M.L., Nicenboim, I. and van der Hoog, W., 2024. Grasping AI: Experiential exercises for designers. *AI & Society*, 39(6), pp.2891–2911.
- Ogundipe; D. O., Babatunde, S.O. and Abaku, E.A., 2024. AI and product management: A theoretical overview from idea to market. *International Journal of Management & Entrepreneurship Research*, 6(3), pp.950–969.
- Pyaee; A. , 2025. The Human-AI Handshake Framework: A bidirectional approach to human-AI collaboration. arXiv preprint arXiv:2502.01493.
- Rezwana; J. , 2022. Understanding user perceptions, collaborative experience and user engagement in different human-AI interaction designs for co-creative systems. arXiv preprint. Available at: <https://arxiv.org/abs/2204.13217>.
- Rezwana; J. and Maher, M.L., 2022. Identifying ethical issues in AI partners in human-AI co-creation. arXiv preprint. Available at: <https://arxiv.org/abs/2204.07644>.
- Schuett; J. , 2024. Risk management in the Artificial Intelligence Act. *European Journal of Risk Regulation*, 15(2), pp.367–385. DOI: 10.1017/err.2023.1.
- Shetty; D. K., Vijaya Arjunan, R., Cenitta, D., Makkithaya, K., Hegde, N. V., Bhatta B, S. R., ... Pullela, P. K. (2025). Analyzing AI regulation through literature and current trends. *Journal of Open Innovation*, 11(1), 100508.
- Shneiderman; B. , 2020. Human-centered artificial intelligence: Reliable, safe and trustworthy. *International Journal of Human-Computer Interaction*, 36(6), pp.495–504. DOI: 10.1080/10447318.2020.1741118.
- Siau; K. and Wang, W., 2018. Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, 31(2), pp.47–53.
- Sousa; S. , Cravino, J., Martins, P. and Lamas, D., 2023. Human-centered trust framework: An HCI perspective. arXiv preprint arXiv:2305.03306.
- Sun; L. , Qin, M. and Peng, B. (2024) LLMs and Diffusion Models in UI/UX: Advancing Human-Computer Interaction and Design. Center for Open Science.
- Vaithilingam; P. , Zhang, Y., Tang, Z. and Xu, D., 2022. Expectations and experience: Exploring users' perceptions of AI code generation. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp.1–13. DOI: 10.1145/3491102.3501891.
- Wang; D. , Yang, Q., Abdul, A. and Lim, B.Y., 2019. Designing theory-driven user-centric explainable AI. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp.1–15. DOI: 10.1145/3290605.3300831.
- Wang; W. , Yang, M. and Sun, T., 2023. Human-AI co-creation in product ideation: The dual view of quality and diversity. *SSRN Electronic Journal*. DOI: 10.2139/ssrn.4668241.
- Weisz; J. D., Muller, M., He, J. and Houde, S., 2023. Toward general design principles for generative AI applications. arXiv preprint arXiv:2301.05578.
- World Health Organization; 2024. Enhancing trust in AI for health through good practices, transparency and governance. World Health Organization, Geneva, Technical Report.
- Yang; Q. , Steinfeld, A. and Zimmerman, J., 2017. Investigating how experienced UX designers effectively work with machine learning. In: *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp.2817–2828. DOI: 10.1145/3025453.3025913. .