Proceedings of the Design Society, Volume 6: DESIGN 2026
https://doi.org/10.1017/pds.2026.10608

Human-AI co-creation: why, what, and how?

Runtao Ren[1,✉], Jian Ma[1] and Jianxi Luo[2]

[1] *Department of Information Systems, City University of Hong Kong, Hong Kong S.A.R. (China),*
[2] *Department of Systems Engineering, City University of Hong Kong, Hong Kong S.A.R. (China)*

✉ runtaoren2-c@my.cityu.edu.hk

ABSTRACT: The rapid diffusion of generative AI is pushing creative work toward human–AI co-creation (HAIC). This paper designs a conceptual HAIC model that specifies several indispensable elements of effective co-creation: Human, AI, Artifact, Instruction, and Interaction. We demonstrated through a case study of a large-scale management information system development project how the HAIC model helps organizations implement HAIC. The proposed framework offers both an analytical lens for researchers and prescriptive guidance for practitioners seeking to engineer reliable human–AI collaboration.

KEYWORDS: artificial intelligence (AI), generative design, collaborative design, design theory

1. Introduction

For a long period, creative work was understood mainly as a human-centered activity, grounded in individual insight and experiential knowledge (Perry-Smith & Mannucci, 2017). With the rapid progress of generative AI, this assumption is being challenged. AI systems can now create texts, images, code, design alternatives at a scale and speed that were previously unattainable. More and more studies are beginning to investigate human–AI co-creation (HAIC) (Rai & Sarker, 2019; Song, Zhu, & Luo, 2024; Wu et al., 2021). HAIC here refers to humans and AI working together synergistically to accomplish creative tasks (Song, Zhu, & Luo, 2024; Wu et al., 2021). In the process of HAIC, both humans and AI leverage their unique strengths, engaging in dynamic interaction to achieve tasks. HAIC is already visible in practice, such as prompt-based creation, model-assisted design, agent-supported planning (Sarica & Luo, 2024; Ren, Ma, & Luo, 2025; Zhu & Luo, 2023). The purpose of this article is to clarify the necessary elements in HAIC. Hence, we propose a conceptual model of HAIC that distinguishes five core elements: Human, AI, Instruction, Interaction, and Artifact. By defining these elements and their relationships, we contribute a theoretical lens consistent with design-oriented research. We treat the co-creation process itself as a design object, providing developers and researchers with a specific checklist to diagnose why a collaboration fails.

2. Theoretical background

Creativity, traditionally defined as the creating artifacts that are novel, valuable, and surprising (Sternberg & Lubart, 1999), has historically been viewed as an exclusively human faculty. However, the deep impact of AI into the design domain has destabilized this anthropocentric view, forcing a re-examination of the locus of creativity. From a historical perspective, the role of AI in creative tasks has undergone three transformations, which are caused by changes in the division of labor between human and machine (as shown in Figure 1).

In the early phase, AI served usually as a tool. The creative process was entirely human-driven, with AI acting merely as executors to improve efficiency or precision within fixed boundaries. While this ensured that design outputs were deeply grounded in human intent, the process was inherently bound by human cognitive limitations (Jansson & Smith, 1991; Hatchuel, Le Masson, & Weil, 2011). Humans, while

https://doi.org/10.1017/pds.2026.10608 Published online by Cambridge University Press

excellent at assigning meaning, are often slow to explore complex combinatorial spaces (Boden, 2009), limiting the scope of innovation to familiar solutions.

With the rise of Generative AI, AI could create vast amounts of content with a few human prompts. However, this autonomy introduces a critical barrier to design practice. Without human intervention, AI-generated outputs face significant legal and ethical challenges, particularly regarding copyright infringement and intellectual property ownership (Ren et al., 2025). Since generative models derive outputs from vast datasets of existing works, purely AI-generated artifacts often exist in a legal gray area, lacking the clear "human authorship" required for copyright protection in many jurisdictions. Furthermore, the opacity of the generation process raises unresolved questions about accountability: if an AI design infringes on a patent or causes harm, the absence of a human agent makes liability difficult to assign.

Consequently, the AI-led creation, despite its generative power, is insufficient for professional design practice. The inability to secure legal ownership or guarantee ethical accountability necessitates a transition to HAIC. Rather than treating AI as a mere evaluator or executor, this paradigm of HAIC positions AI as a co-creator, co-generator, co-explorer, co-evaluator, and even co-explainer (Yu, 2025). Through dynamic interaction, human and AI collaborate to accomplish creative tasks. In this paradigm, the goal of HACI is not to replace human creativity by AI but to expand the boundary of creativity by human-machine collaboration—achieving breakthrough innovation neither human nor machine could achieve alone.

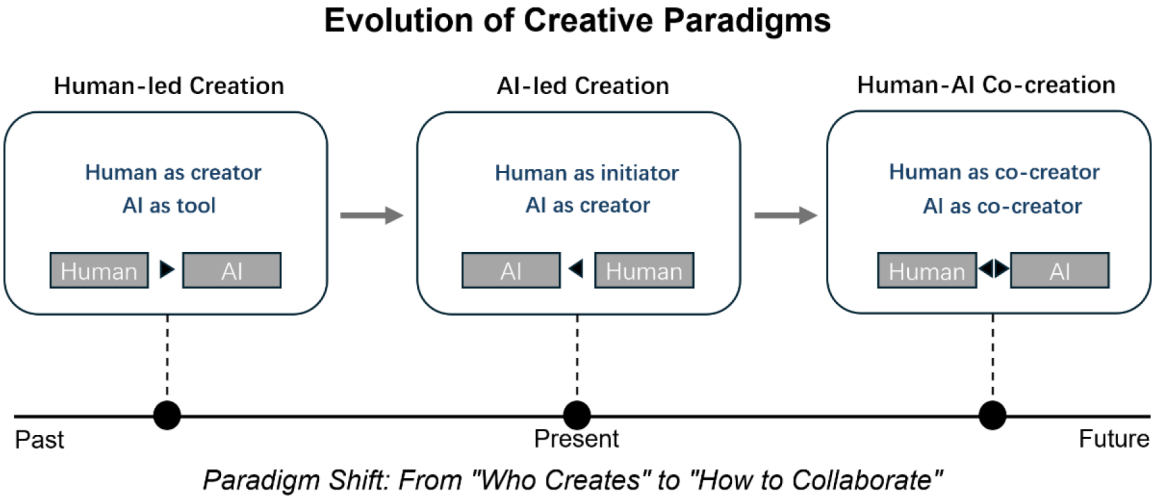


Figure 1. Evolution of creative paradigms

3. Proposed framework

3.1. Requirements to define a theoretical model and relevance of design theory

In the context of design science research, a design theory is not simply an observation of "what happens," but a prescriptive framework that needs to answer two questions: "What is the design object?" and "How is this design constructed, instantiated, and evaluated?" (Hevner et al., 2004). Therefore, we posit that a valid conceptual model for HAIC must satisfy three critical theoretical requirements:

(1) The design object of the model must be defined. In information system design science, the design object is usually a type of IT artifact (Hevner et al., 2004); here, our design object is the process of HAIC. We are not designing a piece of code or an interface, but a collaborative method of human-machine that can be repeatedly executed. By treating the "co-creation process" itself as object can we discuss subsequent improvements, comparisons, and evaluations.

(2) The elements and relationship of the model must be explained. The reason why a theoretical model can guide design is because it clearly states: if this type of design activity is to work normally, which elements and relationship need are required. In HAIC, the relationship between various elements is that humans, AI, and artifacts all require instructions to achieve internalization. So we propose five elements in concept model of HAIC: human, AI, instruction, interaction, and artifact, based on the consideration of the "minimum feasible collaborative unit": without people, there is no purpose, without instructions,

https://doi.org/10.1017/pds.2026.10608 Published online by Cambridge University Press

there is no alignment; without interaction, there is no intermediate correction; and without products, there is no evaluation basis. That is to say, these elements form a necessary set, which together constitute the operational process of the conceptual model of HAIC; By reducing any option to optional, the model degenerates from a designable framework to an accidental usage scenario.

(3) A theoretical model must also be evaluable and communicated. Design theory does not encourage remaining only on conceptual diagrams; it requires models to be implemented in construction and evaluation: Can we use it to design a specific system or product? Can we use it to check which element is missing in an existing method? The five elements of proposed HAIC model are precisely for operationalization: developers can use it as a checklist for human-machine co-creation system design (e.g., Does the interface have explicit interactions?), and researchers can also use it as an analytical framework to explain why the same model behaves differently in different organizations (e.g., What unclear instructions cause interaction failure?).

A good design outcome should be "purposeful, innovative, formalizable, searchable and communicated"; HAIC corresponds to these points: it formalizes the human-machine collaboration process into a describable cycle, and clearly communicates it to developers, researchers, and managers through five elements.

3.2. Modelling human-AI co-creation as concept model

We argue that HAIC should be rendered as a purposeful design process. From in-depth observation, abstraction, and systematic consolidation of collaboration patterns, we formulate the HAIC model as a conceptual representation of this process. This model conceptualizes the HAIC as the process in which humans and AI collaborate efficiently and interactively under instructions to jointly create novel and valuable artifacts. As, shown in Figure 2, the concept model of HACI has five core elements: Human, Artificial Intelligence, Artifact, Instruction, and Interaction. These elements are key to achieving efficient and high-quality co-creation. Next, we will elaborate on the five core elements, clarifying their roles, indispensability, and unique contributions in co-creation process.

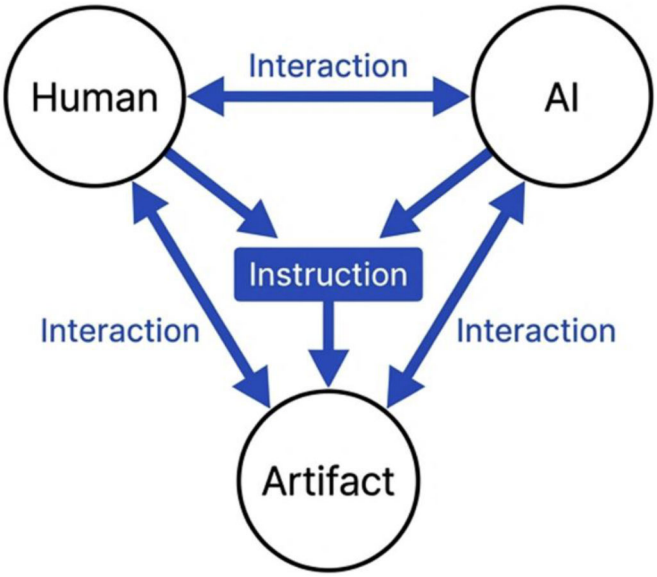


Figure 2. The framework of HAIC Model

3.2.1. Human

The goal of HAIC is to collaboratively generate useful and novel artifact in a controllable, iterative, and evaluable process, rather than allowing the AI unilaterally to output plausible but are disconnected from real-specific context. Under this objective, the reason why human is considered as an element is that only humans can define what is 'useful', what is 'novel', and what exactly needs to be generated in the current real-specific context. In other words, HAIC aims not to produce abstract content, but to collaboratively create content tailored to specific object, tasks, and contexts, this this direction can only be injected by people. In achieving this goal, humans play two crucial roles: specification and evaluation. Specification refers to humans transforming a vague need into an instruction that AI can understand and execute, such

https://doi.org/10.1017/pds.2026.10608 Published online by Cambridge University Press

as refining "write a description of iPhone" into "write a description of iPhone for non-technical readers, emphasizing application value, and under 200 words." Evaluation refers to humans judging whether AI outputs align with the intent, identifying deviations, and determining whether further iteration is warranted. These two actions transform HAIC from "the model outputting content once" into a "collaborative process that continuously approaches human intent."

When human-machine collaboration is truly in operation, humans are not merely abstractly "present," but rather play a role through two entry points: On one hand, humans write, select, and modify instructions—that is, transforming their intentions, standards, tone, and requirements into commands that AI can execute—clearly stating "what I want." On the other hand, humans use interaction to see what the AI has generated, whether it has understood, whether it needs to be restarted, and whether partial modifications are needed—clearly stating "how well it was generated." Humans therefore repeatedly switching between these two entry points constitute the human control chain of HAIC.

3.2.2. AI

The goal of HAIC is also to enable co-creation to work effectively, efficiently, and diversely, which means transforming the work that originally required a lot of manual designing, searching, writing, checking, comparing and expanding into a generation process that can be executed quickly and repeatedly, while keeping human intent intact. Under this goal, AI's role is not simply "helping to write a paragraph," but rather acting as a "generation and transformation engine" for the collaborative process: given great instructions from humans, AI can produce multiple versions, perspectives, and contents in a very short time, laying out all the parts that previously required slow thinking, research, and organization from humans at once, giving humans choices, adjustable results, and comparable evidence.

Many real-world collaborative scenarios—such as migrating a description to several styles, writing versions for different audiences based on the same set of facts, generating several alternative solutions based on a single idea, and quickly organizing unstructured data into structured data—are essentially high-frequency, repetitive, batch tasks that require memorizing large amounts of context. If these tasks were entirely done by humans, the cost would be extremely high, and rapid iteration would be impossible. Therefore, AI is key to transforming HAIC from simply "collaborating" to "collaborating efficiently."

3.2.3. Artifact

The goal of HAIC is not simply to "make conversations between humans and AI," but rather to "ensure that this collaboration produces tangible and evaluable results." Without artifacts, HAIC is merely an interactive process; with artifacts, HAIC becomes a creative process. The term "artifact" here should not be understood as just the final solution/product to be delivered, but also includes various intermediate results generated throughout the process, such as ideas, solutions, and even concepts and knowledge generated by human-computer interaction (Kaptelinin & Nardi, 2009).

The reason for considering artifacts as an important element of HAIC is that for co-creation to be controllable and evaluable, there must be an externally medium to accommodate the respective tasks performed by humans and AI. If humans provide instructions to AI within interaction, but the result only remains at the level of "the model was called once," the collaborative process and output is invisible, and it's impossible to say, "what went well in this round, what went wrong, or how to adjust in the next round." The existence of an artifact transforms the abstract "I want the model to generate something" into the concrete "This is the content it has generated." Only when artifacts are created can people use their contextual knowledge, quality standards, and style preferences to judge whether something is right, sufficient, or whether it needs further refinement. In other words, the artifact is the ontology that transforms the process into an object; without artifact, the human, instruction, AI, and interaction in Human-Computer Interaction cannot be verified.

3.2.4. Instruction

The HAIC process is not simply about "humans saying a sentence and AI responding with a sentence," but rather about enabling AI to stably, repeatedly, and predictably execute human intentions. Therefore, there needs an intermediary responsible for "turning what human think or say into what AI can do", which is instruction. The "instruction" here is not a temporary prompt in the narrow sense, but refers to a

https://doi.org/10.1017/pds.2026.10608 Published online by Cambridge University Press

complete set of methods that formalize, structure, and proceduralize human intentions, it can be a carefully designed prompt, a set of step-by-step task specifications, a fixed generation template, or even the process of translating a certain design/innovation method into machine-executable commands. Without great instructions, human-machine interaction even becomes disjointed—like " a chicken talking to a duck "—and fails to produce valuable artifact. Therefore, instruction should be considered as a significant element key for HAIC.

3.2.5. Interaction

Humans decide "what to do," instructions decide "how to do it," AI decide "to what extent it can be done," and artifacts decide "what has been achieved so far," then interaction determines "whether all of these can be seen, whether these can be corrected in time, and whether these can proceed at an appropriate pace." To achieve this, there must be a dedicated channel responsible for "transporting information, sending results back, and relaying feedback," and this channel is called interaction.
In actual operation, interaction typically could manifest in familiar forms: a dialogue website, a selectable and editable interface, and even a canvas where modules can be dragged and dropped. Regardless of the form, it all boils down to the same thing: conveying instructions from one side to another. Therefore, in HAIC, interaction is not just a simple interface, but a link that truly connects the human, AI, artifact, and instruction into a collaborative system.

4. Case study: HAIC model in practice

4.1. Implementation background

ScholarMate is the largest professional research social networking platform in China and has long undertaken the development and operation of research management information systems of many universities and scientific research institutions in Chinese Mainland and Hong Kong. In 2025, this company launches the research management information systems for the Macau region. Compared with existing regions, Macau has several localized differences in scientific research project filing, statistical reporting, and other aspects, which makes it impossible their previously developed systems to be directly migrated and used. They need to design a system prototype with a similar structure but different rules in a relatively short period of time. At the start of the project, the company adopted the existing development approach: with front-end and back-end divided by two teams; The expression of requirements mainly relies on PPT documents created by product personnel, which display the expected interface and process in the form of static screenshots; Every adjustment to the interface, fields, and process requires the design side to update the PPT and then deliver it to the R&D side for front-end restoration. This model was acceptable in projects with a 6-12 month cycle in the past, but in the Macau project, due to the high frequency of requirement changes and the need to quickly align with multiple business personnel, the original approach showed obvious problems of high iteration costs and long feedback cycles.

4.2. Problem identification

When this company launched the development of a research information management system for Macau, one of its overall goals was to "introduce GenAI into the existing development process to shorten the prototype development cycle." However, when we entered the project, the actual work we observed was still their long-established pattern: Product manager designed static page prototypes in the form of PowerPoint, explained the meaning of the pages to the development team through meetings, and then the developers recreated each page based on the PowerPoint, repeating the above process when encountering business changes. Based on the concept model of HAIC, we examined the existing development process of ScholarMate and constructed a targeted problem matrix (as shown in Table 1) to identify which elements already exist, which elements exist but lack form, and which elements have not yet been made explicit.Overall, the diagnosis can be summarized as: Human and the final Artifact existed, but Instruction was not structured, Interaction was not high-frequency, AI was not embedded in the prototyping stage, and the Artifact was not functionalized to support regeneration.

https://doi.org/10.1017/pds.2026.10608 Published online by Cambridge University Press

Table 1. Problem Identification

Element	Analysis
Instruction	The existing requirements mainly exist through verbal explanations, expressed in a way that is "friendly to people, unfriendly to AI". Instructions have not been organized into prompts that can directly drive GenAI invocation.
Interaction	The team aligns requirements through online and offline meetings. However, this interaction cannot be immediately reviewed, annotated, or assigned in collaborative interfaces.
AI	The existing process has not introduced any GenAI tools or AI-assisted prototype design methods, and all prototypes are manually drawn page by page in slides.
Artifact	The current output mainly consists of static slides, which lack information such as interaction status, page jump relationships, and component hierarchy. This artifact cannot be used as the "sole source of facts" for subsequent development to directly rely on, nor can it be reused as input for the next round of generation.
Human	Although team can clearly articulate what needs to be done, their expression is still primarily aimed at human rather than AI. This means that even with the introduction of AI tools, HAIC will not work smoothly without improving human-AI co-creation skills.

4.3. Solution and effect

Based on the above HAIC-based diagnosis, we did not replace the company's business architecture or team division. Instead, we supplemented the process along the elements in HAIC model so that the entire co-creation loop could be made explicit and executable. The interventions and observed effects are summarized in Table 2. After implementing the above methods, when the company develops new or similar systems, the cycle in the prototyping and front-end determinism phase is reduced from approximately 6–12 months (depending on the frequency of changes) to approximately 3 months to create a complete, clickable, and collaboratively accessible high-fidelity prototype. The amount of rework in the subsequent front-end implementation phase also decreases significantly. However, it is important to note that this efficiency improvement does not stem from the introduction of a single tool, but rather from the simultaneous activation of the five elements of HAIC.

Table 2. Interventions and effects

Element	Intervention	Effect
Instruction	The requirements are summarized into a reusable prompt framework, allowing the team to adjust the order according to business complexity.	The same requirement no longer needs to be repeatedly "heard once and then translated again" by the design team.
Interaction	The practice of verbally proposing changes during meetings has been moved to a real-time co-creation website for online commenting and revising.	Improved collaboration efficiency and reduced centralized rework in the terminal stage
AI	By introducing GenAI tools such as Figma, the summarized instructions can be used to generate page skeletons, component layouts, and basic navigation, instead of manually drawing slides.	Prototype production time was reduced from ≈2–4 weeks per batch of pages (manual) to ≈1–3 days per batch (AI-assisted).
Artifact	All field additions, deletions, and navigation should first be showed on the interaction, allowing the generated prototype to be used as input for next generation.	Upgrading Artifact from a "static diagram for illustration" to an "intermediate product for collaborative and regeneration purposes" not only supports multi-person revisions but also allows it to be reverse-engineered into the next round of generation.
Human	Provide employees with micro-training HAIC: how to write executable instructions, how to review artifact, and how to convert comments back into instructions.	The number of business personnel capable of independently understanding and writing "AI-executable instructions" increased from 0 to 70-80% of the core project team.

https://doi.org/10.1017/pds.2026.10608 Published online by Cambridge University Press

5. Discussion

This paper addresses a specific problem: many individuals or organizations are already using AI for creative work, but most practices remain at the level of one-time model calls and partial replacement of human intervention, failing to guarantee that collaboration is controllable, repeatable, and measurable. Our proposed HAIC model attempts to demonstrate that achieving high-quality human-machine co-creation requires making the elements of the co-creation activity explicit and explaining the relationships between them. Many companies feel that "we're using ChatGPT, but efficiency hasn't improved," not because the model is weak, but because they haven't analyzed the various elements of human-machine co-creation and their relationships, and haven't adopted appropriate methods for the specific task. Overall, in order for HAIC to become a designable, reusable, and assessable process, it is necessary to clarify elements, which can be used to diagnose where the existing process is weak, how to make up for it, and what kind of improvement should be seen after making up for it, just like our case.

6. Conclusion and future work

This paper proposed a conceptual framework for HAIC that identifies five interdependent elements—Human, AI, Instruction, Interaction, and Artifact—and argues that together they constitute the minimum feasible unit for controllable, repeatable, and evaluable co-creation. We clarified what such a design theory must do to be meaningful: it has to define its design object, identify its necessary elements and their relationships, and be implementable and evaluable in practice. The case of Scholarmate illustrated that, when the five elements were systematically activated, an organization could move from a 6–12 month prototype development cycle to roughly 3 months. However, we also need to acknowledge that the current model is too macro and does not involve specific ways to improve a certain element. For example, we encountered training on how to enhance people's familiarity with AI in the transformation of ScholarMate, and the success of the model is also related to how to use it.

Future research can extend this work in at least three directions. First, comparative studies across domains can test whether the five-element structure remains stable or needs domain-specific refinements. Second, quantitative evaluations can be designed to isolate the effects of individual elements on co-creation performance. Third, longitudinal studies can examine how human co-creation literacy develops over time and how that, in turn, changes the optimal design of the instruction and interaction layers.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (725B2035).

References

Boden, M. A. (2009). Computer models of creativity. *Ai Magazine*, 30(3), 23–23.

Franceschelli, G., & Musolesi, M. (2024). On the creativity of large language models. *AI & SOCIETY*, 1–11.

Hatchuel, A., Le Masson, P., & Weil, B. (2011). Teaching innovative design reasoning: How concept–knowledge theory can help overcome fixation effects. *Ai Edam*, 25(1), 77–92.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 75–105.

Jansson, D. G., & Smith, S. M. (1991). Design fixation. *Design studies*, 12(1), 3–11.

Kaptelinin, V., & Nardi, B. A. (2009). *Acting with technology: Activity theory and interaction design*. MIT press.

Perry-Smith, J. E., & Mannucci, P. V. (2017). From creativity to innovation: The social network drivers of the four phases of the idea journey. *Academy of management review*, 42(1), 53–79.

Rai, A., Constantinides, P., & Sarker, S. (2019). Editor's comments: Next-generation digital platforms: Toward human–AI hybrids. *MIS quarterly*, 43(1), iii–ix.

Ren, R., Ma, J., & Luo, J. (2025). Large language model for patent concept generation. *Advanced Engineering Informatics*, 65, 103301.

Runco, M. A. (2023). AI can only produce artificial creativity. *Journal of Creativity*, 33(3), 100063.

Sarica, S., & Luo, J. (2024). The innovation paradox: concept space expansion with diminishing originality and the promise of creative artificial intelligence. *Design Science*, 10, e11.

Song, B., Zhu, Q., & Luo, J. (2024). Human-AI collaboration by design. *Proceedings of the Design Society*, 4, 2247–2256.

Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity: Prospects and paradigms. *Handbook of creativity*, 1(3–15).

https://doi.org/10.1017/pds.2026.10608 Published online by Cambridge University Press

Wang, P., Khinvasara, Y., Creijghton, G. J., Scholing, T., Wang, Y., Zhou, Z., … & Yin, Y. (2025). Enhancing designer creativity through human–AI co-ideation: a co-creation framework for design ideation with custom GPT. *AI EDAM*, 39, e22.

Wu, Z., Ji, D., Yu, K., Zeng, X., Wu, D., & Shidujaman, M. (2021, July). AI creativity and the human-AI co-creation model. In *International Conference on Human-Computer Interaction* (pp. 171–190). Cham: Springer International Publishing.

Yu, W. F. (2025). AI as a co-creator and a design material: Transforming the design process. *Design Studies*, 97, 101303.

Zhu, Q., & Luo, J. (2023). Generative transformers for design concept generation. *Journal of Computing and Information Science in Engineering*, 23(4), 041003.

https://doi.org/10.1017/pds.2026.10608 Published online by Cambridge University Press