AI-Driven Pedagogies in Architecture: A framework for early-stage design education

Angelica Paiva Ponzio[1], Ava Fatah gen. Schieck[2]

[1]Propar - Federal University of Rio Grande do Sul - UFRGS, Porto Alegre, Brazil,
[2]University College London, Bartlett School of Architecture, London, United Kingdom

angelica.ponzio@gmail.com; ava.fatah@ucl.ac.uk

Abstract. This exploratory study investigates the role of 2D generative Artificial Intelligence (GenAI) applied during early design-stages in architectural education settings. Grounded in Donald Schön's (1983) reflective practice and adopting a Design-Based Research (DBR) methodology, it presents frameworks where **AI functions as a conversational design partner**. Three exploratory phases are presented: (1) a theoretical stage—aimed at building familiarity with tools and refining processes; (2) a test stage— a workshop testing the pipeline; and (3) a design studio, employing a structured AI matrix, to explore the concept of **parameter-based prompt blending** across multimodal platforms. Initial findings indicate that GenAI enhances conceptual stages when anchored in clear design intentions. Moreover, the recent emergence of multi-model AI ecosystems enabled cyclical and layered workflows through Human-AI dialogues producing outputs that transcend generic architectural representation. Furthermore, these systems also foster new forms of pedagogical dialogue between educators and students within the academic environment.

Keywords: Artificial Intelligence, Multimodal Generative Models, Architectural Design, Early-Stage Ideation, Parameter Prompt Blending

1 Introduction

Architectural education has been under significant transformations due to the continuous development of digital technologies. Within this context, new representation, fabrication, and simulation technologies have expanded towards offering creative methodologies that foster innovation. In 2014, Neil Spiller (2014) described a "tsunami of technology" reshaping architectural education, predicting that digital fabrication techniques would revolutionize how architecture is "built, composed, and procured." Similarly, Sedrez and Celani (2014) emphasized the importance of adopting new technologies in design to address contemporary architectural challenges. They proposed a pedagogical framework consisting of three key elements: conceptual content, experimental methodologies, and digital skills—a framework that remains relevant today.

Advances in computational storage and processing have enabled the rapid deployment of AI models, such as neural networks and deep learning (Mitchell, 2019), and more specifically, Generative Artificial Intelligence (GenAI), which rely on vast datasets to create new content. This includes Large Language Models (LLMs) such as GPT-4 and Gemini, and Large Visual Models (LVMs) such as Diffusion Models (DM)—like DALL·E, MidJourney, and Stable Diffusion. These models introduce a new technological paradigm for creative disciplines, promising significant potential for architectural design (Montenegro, 2024; Onatayo et al, 2024; Li et al, 2024; Vissers et al, 2024). However, alongside new possibilities, pedagogical challenges emerge. The "black box" nature of GenAI—where internal decision-making processes are opaque—raises concerns about transparency, authorship, and control, presenting a pedagogical challenge in architectural education. Nonetheless, the capacity of GenAI to generate contextually relevant content within high-quality images makes it powerful when used responsibly. Recent studies (Jang et al., 2025; Li et al., 2024; Fang et al., 2025; Onatayo et al, 2024) have focused on the potential of multimodal GenAI in combining different inputs such as text-to-image, image-to-image, image-to-text in architectural design. Vissers-Similon et al. (2024) further classified AI techniques according to their applications in early architectural design stages, highlighting transformer-based models – including LLMs and LVMs– as effective in conceptual design due to their capacity to combine diverse concepts into unified outputs – such as blending a biological reference with an architectural element. More recently, new AI multi-model systems such as Glif (2024), give users access to the backend of the workflow, enabling more agency to both educators and students, which will be explained in the research presented here.

Building on Schön's (1983) concept of reflective practice, this exploratory research critically investigates the role of GenAI, particularly text and image-to-image models, as a collaborative agent for creative and pedagogical advancement in the early stages of architectural design. By studying different frameworks for conceptual ideation, **the study positions AI as a reflective design partner**, aligning with recent developments in the field. While these technologies are still evolving, the research emphasizes the importance of not only multimodal, user-friendly interfaces but also emerging multi-model platforms that integrate various AI models into a unified workflow, and their application particularly within design education. This approach aligns with the broader goal of fostering human + AI agency, where designers actively engage with AI tools to co-create and iterate during the design process (Yu, 2025). The integration of GenAI—anchored in structured prompting, multimodal inputs, and conceptual clarity—offers new perspectives for architectural design education.

This paper, therefore, presents ongoing investigations and experiments that explore how multimodal GenAI platforms, through image and text-based workflows, can enhance concept generation in the early stages of architectural design, while contributing to the development of pedagogical frameworks centered around designers' agency.

1.1 Setting the Scene: reflective design dialogue through multi modal environments

Among the possible purposes for AI in design, a design assistant seems to be by far the most promising. Some examples of what such a computer-based design assistant might do are (…)provide a system that extended the designer's repertoire of prototypes and enhanced his/her ability to explore them and bring them into transaction with particular design situations (Schön, 1992, p.13).

In conventional architectural design practices and education, precedents and references are selected, drawn, interpreted, and contextualized and their curation are closely tied to the designer's intent, experience, and memory. In AI-driven workflow, curation unfolds during the generation process: prompts guide the model to produce or reassemble dynamic referential outputs, **making prompting itself an act of design**. This shift underpins three main ways GenAI is currently integrated into 2D architectural workflow (Montenegro, 2024; Fang et al, 2025; Jang et al, 2025; Li et al, 2024): (1) as a style transferring and rendering tool for visualizing different material and aesthetic qualities, while preserving spatial composition; (2) as an ideation tool, for conceptual exploration and creative divergence; and (3) as a hybrid process that combines both visual refinement with conceptual development.

Bolojan (2022) indicated, when working with hybrid workflows combining GANs (Generative Adversarial Networks; Goodfellow et al., 2014) and LLMs through text and visual inputs it is important to adopt **multimodality in design thinking** to address the complex and interrelated nature of design (Alexander, 1968). Through a "disentangling" methodology, Bolojan indicates that **AI must be adapted to the cyclical, layered nature of design**. While native multimodal models such as GPT-4o and Google Gemini can directly interpret and respond to both textual and visual inputs, **multi-model environments** often integrate separate models (e.g., a text-based LLM and an image-generating LVM) in a single workflow. In such cases, the LLM supports prompt creation and refinement, while LVM manages visual generation outputs in a controlled manner. Platforms like Glif (modular block-based) and XFigura (canvas-based AI sandbox) exemplify this integration, enabling designers to orchestrate iterative, multi-model dialogues that leverage both linguistic reasoning and visual creativity.

2 Methodology and Multi-Stage Analysis

Instead of focusing on programming or custom code development, the research tested existing user-friendly interfaces and platforms within structured workflows to assess their usability and integration into the design process within architectural education. The objective is to analyze whether these tools: (1) facilitate dynamic iterations compared to non-AI procedures and (2) enable designers to address diverse considerations and aesthetic preferences while exploring innovative solutions. In order to do that, a qualitative and exploratory Design-Based Research (DBR) methodology is adopted (Wang & Hannafin, 2005, Anderson & Shattuck 2012). While linking iterative design propositions with theoretical reflection in design learning contexts, it aligns with DBR phases—analysis and exploration; design and implementation; evaluation, reflection, and refinement— developed so far across three stages:

- **Stage 1 - Theoretical exploration** – developed in three stances (2022/23), the aim was to build familiarity with Image generation of DALL·E 2 and LLM – ChatGPT, focusing on generating climate-responsive conceptual designs.
- **Stage 2 – Graduate workshop** - refine the framework using ChatGPT (GPT-4) + DALL·E 3 and image-to-video generation. The framework was applied during a one-day workshop for Graduate studies (2024) and focused on a brief for biological concept generation for a climate-responsive pavilion using a structured pipeline.
- **Stage 3 - Undergraduate design studio** (2024/25) - adopted a comparative approach in which students designed a small hotel in nature using a multi-model matrix to generate data for reflection and to inform the design workflow development.

2.1 Stage 1 - Theoretical Exploration

The first stage was divided into three sequential experiments (1.1, 1.2, and 1.3) that aimed at building familiarity with AI tools and to explore iterative processes in text-and-image-to-image generation. These experiments explored the generation of climatic responsive referential images of small habitable structures in different locations and contexts. The first two experiments, 1.1 and 1.2, employed a common LVM while the third added a LLM interaction. At 1.2 and 1.3, a 3D independent sequential workflow was also tested on a computational platform (Rhino/Grasshopper).

Experiment 1.1 (Ponzio, 2022) used DALL·E 2 (OpenAI), a multimodal GenAI system that pairs a diffusion model with CLIP (Contrastive Language–Image Pre-training). This model enabled in-painting (erase-and-replace) and accepted image inputs for stylistic or compositional variation. Prompts combined (i) a short textual description and (ii) a **fragment technique** with a small image fragment supplied as a texture or form cue. Results showed that this dual-input strategy yielded high-resolution outputs closely aligned with the intended design goals: the inserted fragment anchored composition and style, while the text steered overall content. Without this fragment, DALL·E 2 tended to default to lower-resolution or generic architectural imagery; with it, the model produced sharper, more context-specific results. Fig.1

Experiment 1.2 (Ponzio et al, 2023) refined the prompt engineering process by introducing a CAPT (Climatic Archetypal Parameters Table - cold, temperate, hot and dry, hot and wet climates) as a validation tool, while continuing to use the **fragment technique** in DALL·E 2. In both experiments, a 3D parametric modelling generation was independently employed to test viability of the design process. Two key limitations emerged: (1) the reliance on generating single façade views in order to facilitate translation into a full 3D model, and (2) the manual effort required to integrate the 2D AI image into the 3D pipeline. Fig.2

Experiment 1.3 (Chornobai et al., 2023) combined DALL·E 2 (via Microsoft Bing, **without the fragment technique**) with text interaction using ChatGPT, in a two-step process. First, **the conversational capabilities** of the LLM were used to generate prompts by cross-referencing climatic archetypes (using CAPT as reference) with biological design principles (biomorphic, bioclimatic, biophilic, biomimetic, and bio architecture). In the second step, these prompts were tested in DALL·E 2 to generate images. While some architectural terms—particularly culturally or technically specific ones—did not clearly translate into visuals, most outputs conveyed meaningful formal and conceptual variations. Limitations included reduced coherence in extended LLM interactions. Nevertheless, both tools demonstrated strong potential for creative exploration when guided by “active human agency”. Fig. 3

The outcomes of stage one and reflections on the findings contributed to the development of the second phase of this research which was conducted in a workshop setting (Figs.4 and 5).

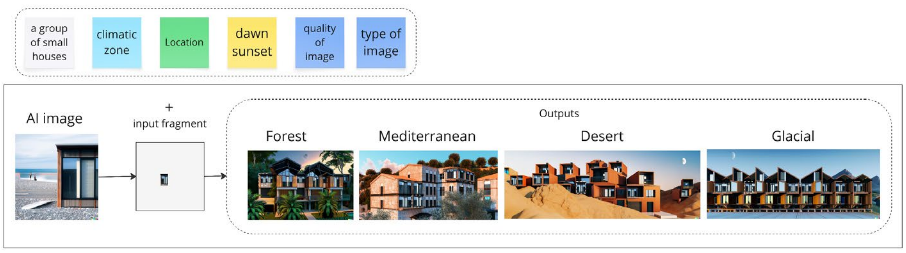

Figure 1: AI Design Process. Author: Ponzio, 2022.

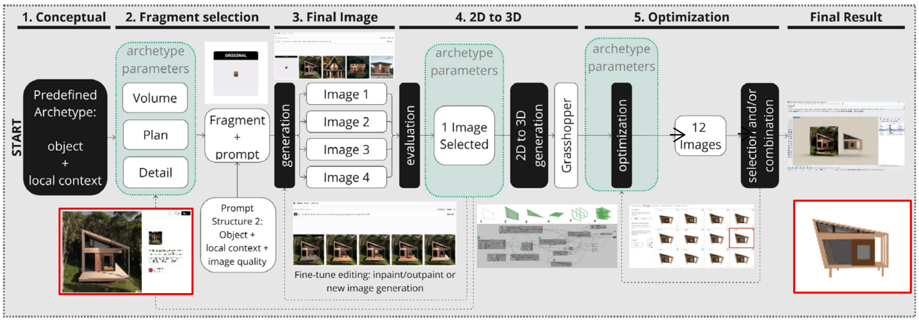

Figure 2: Design Process. Authors: Ponzio et al., 2023.

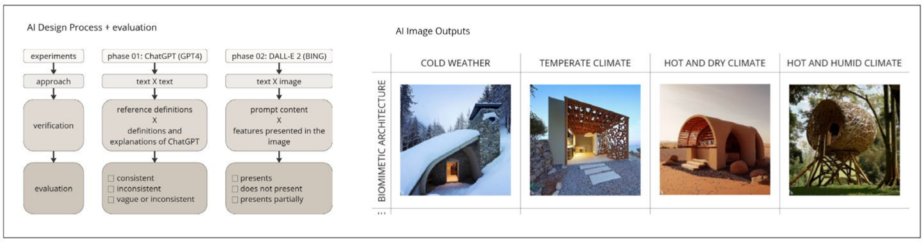

Figure 3: Design Process. Authors: Chornobai et al., 2023.

2.2 Stage 2 – Workshop for Graduate studies

This stage aimed to build on Stage 1 outcomes through practical application with a small group of architectural students. The focus was on translating concepts into formal design solutions within a contextual framework. The workshop was conducted in 2024 with 14 participants from the Architectural Computation MSc at the Bartlett School of Architecture, UCL, London. While the choice of AI models was left open, meaningful results emerged when students' teams used the **structured multimodal prompt** (Fig. 4) with DALL·E 3 in ChatGPT (GPT-4), which automatically **rewrites prompts** into more detailed textual descriptions, differently from DALL·E 2 (Fig. 5 box 04). The workshop explored nature-inspired concepts and climate-responsive solutions for a pavilion in Regent's Park, London. After receiving the suggested prompt structure (Fig. 4) and site images from the tutor, a biomimicry design framework was introduced to support a nature-driven approach, using local flora or fauna as conceptual references, while addressing seasonal weather conditions. Expected outputs included 2D images and a short AI-generated animation showing dynamic interactions between structure and site. The hands-on session (2 hours) was followed by an introductory lecture on generative AI (1 hour). During the workshop, students were divided into four groups of three to five; while three groups followed the suggested prompt structure, one group did not. Despite the availability of DALL·E 3 built-in prompt generation, it was essential to guide the students in crafting their initial inputs. **Well-structured multimodal prompts enabled more intentional and concept-driven design results.**

Two groups used the site image as input, demonstrating DALL·E 3's capacity to interpret and reproduce visual references while simultaneously supporting concepts generation. This exercise exemplified **semantic prompt engineering** in which carefully curated visual and textual tokens steered the model toward context-specific outcomes. One group successfully used a local bird image input (fig. 5), producing AI-generated outputs within a concept inspired by the bird´s form. Initially, the AI bias towards the word "pavilion" resulted in a modernist structure. To refine the output, the group added a specific style ("parametric pavilion") and then introduced the bird image, morphing its form and colors into the design.

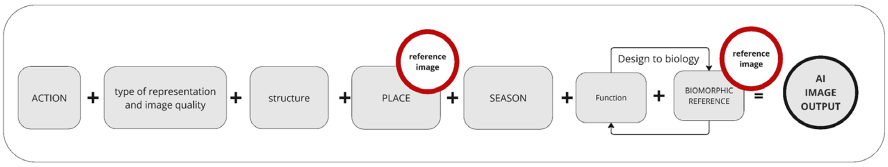

Figure 4: Prompt framing. Author: Ponzio.

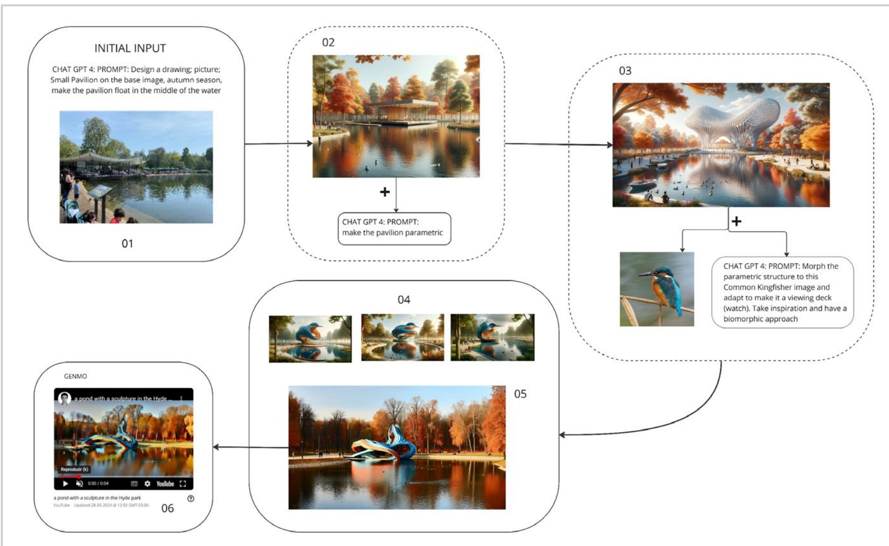

Figure 5: Design Process. Authors: Group 01, 2024.

2.3 Stage 3 – Undergraduate Design Studio

The third phase was developed within a 5th-semester undergraduate design studio (2024/2nd sem. – 12 students; 2025/1st sem. – 15 students). Unlike in stage 2, the studio brief focused on designing a group of small cabins for a nature-based hotel complex located by a river. GenAI exercises were introduced prior to form development, serving as tools for conceptual exploratory ideation. This time, students worked with a **more constrained prompt framework** (fig.6) and a **comparison matrix** (fig.7), testing both text-to-image and image-to-image multimodal GenAI models, including Adobe Firefly, ImageFX, Meta AI, Craiyon, Gamma, Leonardo AI, Flux, LookX, DALL·E 3, and ChatGPT. The studio also leveraged the **API-driven modular multi-model platform—Glif**, which supports **controlled creative agency** through an editable block-based structure. In Glif, each block operates as an input parameter, enabling users to prioritize specific aspects of the generative process—including the ability to assign weights to text and image inputs within a ControlNet block. This structure mirrors parametric design logic, where dynamic relationships between variables are explicitly defined to influence outcomes, thereby enabling a form of "**parameter prompt blending**." Students analyzed the images and how the varying parameters—program, concept, materiality, typology, location, and persona— could directly influence the outcomes when composing a **multi-attribute prompt** (see Fig.7). This reinforced the understanding that **prompting is in itself a design act**, requiring iterative refinement and strategic decision-making. Key observations include:

1. **Longer, descriptive prompts** consistently produced outputs with greater architectural diversity, spatial coherence, and fidelity to the design intention. This was evident when contrasting DALL·E 3 within ChatGPT (2024/2; red box in Fig. 7)—which automatically rewrites user prompts into more detailed internal textual instructions—with ChatGPT's GPT-4o Image Generation (2025/1), which relies directly on user inputs. Sharing GPT-4o's transformer architecture, the latter delivers faster results but often defaults to literal or generic solutions unless guided by equally detailed user prompts.
2. Certain words can carry **built-in bias toward predictable archetypes**—like pitched roofs for "cabin," or curved geometries for "parametric." However, when text and/or image inputs incorporated conceptual **analogical or concrete formal** references (e.g., snail shell, origami, etc.), output shifted toward potentially innovative proposals, demonstrating AI's capacity for creative divergence.

The models also worked as **rendering tools** when students uploaded sketches, CAD drawings, and site photographs alongside text to test design concepts. Finally, the comparison matrices served not only as output catalogues but rather as **reflective tools for design thinking**.

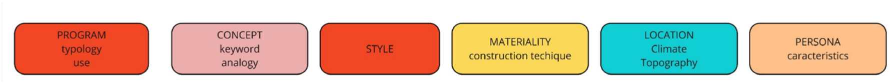

Figure 6: Prompt Framework, 2024/2; Author: Ponzio.

Figure 7: Comparison Matrix. Authors: Group 3, 2024/2.

3 Results and Discussion

This research explored how prompting within AI models can function as a **design conversation**, reframing the role of image referencing in design thinking and positioning prompt-crafting itself as a **core design act**. Across the experiments, besides AI's speed and variability accelerated the generation of diverse iterations, three primary uses of AI-generated 2D imagery emerged: (1) producing referential images, (2) rendering input images, and (3) combining both strategies. Results are as follows:

The experiments in **stage 1** revealed the potential of 2D GenAI tools to augment both creative ideation and workflows. While DALL·E 2 fragment inputs (**fragment-based prompting**) in 1.1 and 1.2 captured contextual nuances, the employment of ChatGPT **conversational capabilities** improved the user's ability to prompt writing. It also underscored the importance of incorporating conceptual analogical or concrete formal references in the workflow. Results from **stage 2** demonstrated an improvement in multimodal prompting with DALL·E 3 embedded in ChatGPT (GPT-4), when exploring structured prompts and referential imagery. This iteration marked a shift from fragment input to **semantic prompt engineering**, supported by the model's ability to generate detailed textual expansions and stronger visual coherence with conceptual alignment. The shift from a pavilion typology (stage 2) to a design studio brief such as the hotel cabin in **stage 3**, aimed to test more detailed and multi-attribute prompting strategies—adding materiality and program considerations. This extended the methodology beyond generic object-based outputs toward more specified, program-driven architectural proposals. The results of **stage 3** confirmed that AI-assisted design ideation is highly dependent on designer's agency in crafting prompts. Multimodal prompting—combining text, site images, concept diagrams, and referential images—was crucial for mitigating default biases and steering the AI toward meaningful, site-specific, and conceptually rich outcomes, while also introducing unexpected ideas and expanding design possibilities. The comparison matrices used in this phase functioned not only as catalogues of visual outputs but also as reflective design tools. Platforms like Glif, which enabled users to weigh multimodal inputs and build prompts through modular blocks and supported a process of **parameter prompt blending**—also opened new forms of **tutor's agency** in guiding prompt structure. Nevertheless, some students exhibited hesitation in adopting more innovative formal solutions, likely reflecting their ongoing design vocabulary development and limited CAD tools confidence—a challenge less evident in the workshop, where programmatic and technical constraints were reduced.

Therefore, this study highlights how the emergence of multi-model AI ecosystems enables **cyclical reflection-in-action and layered design processes through Human-AI dialogues**, moving beyond generic architectural representations. These systems foster iterative exploration while also unlocking new pedagogical possibilities—where **design agency is not outsourced but co-constructed**. The augmented agency of the tutor, supported by platforms such as Glif, plays a central role in guiding students' engagement with AI, reframing prompting as a shared and reflective learning process. Therefore, this research calls for the development of:

1. A pedagogical framework for AI-assisted early-stage architectural ideation that integrates pedagogical strategies to scaffold ideation and development.
2. Evidence-based guidelines for integrating AI into design studio pedagogy to support reflective practice, including:

- **Multi-prompt engineering** strategies, to enhance the awareness of the context, emphasizing site-specific constraints, programmatic considerations, and conceptual references.
- **Semantic prompt engineering**, enabled by LLMs like ChatGPT, to support detailed and context-rich instructions.
- **Parameter prompt blending**, as a pedagogical strategy, in which the act of weighing and combining different text and image inputs is made explicit to students as a way of thinking relationally — mirroring parametric design logic and supporting conceptual control in early design stages.
- **Raising AI literacy**, raising awareness regarding the capabilities of the different AI tools available. Tool selection and capabilities play a crucial role, with multimodal and multi-model platforms facilitating workflow transitions workflows from ideation to 2D and 3D materialization.

4 Conclusion

This study extends Donald Schön's (1983) concept of reflective practices and explores new methodologies within the digital design studio driven by GenAI-assisted design workflows. Prompting becomes a design move, with the AI's output functioning as a form of back-talk—revealing latent patterns, biases, and unexpected results—as students adjust their prompts to refine style, context, and composition. This mirrors Schön's (1983) notion of "moves, consequences, and reflection-in-action," and shifts the designer's agency toward curating language, selecting multimodal inputs, and interpreting AI responses—requiring critical prompting strategies to sustain authorship and creative control. By identifying key factors that influence ideas generation, the research tested alternative ways of integrating generative AI into the design process as part of developing a pedagogical framework. Findings highlight the importance of tool selection, structured prompts, and active human agency in leveraging AI's potential for innovative, context-sensitive proposals. Multimodal frameworks foster reflective processes that push AI beyond image creation toward narrative construction, conceptual reasoning, and speculative form-making. Moreover, AI's ability to produce unexpected outputs becomes a catalyst for creativity—when guided by clear conceptual frameworks. Meaningful integration depends on balancing structured with iterative dialogues—especially with the rise of multi-model platforms, which enable controlled creative workflows combining text and image in a continuous feedback loop, and facilitate active agency by tutors and students. Ongoing work includes extending the pipeline into 3D mesh conversion and prototyping. Responsible GenAI use will be further explored through interpretive framing and critical reflection across design stages.

Acknowledgements This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. This article was edited using ChatGPT for grammar and style.

References

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? Educational Researcher, 41(1), 16–25. https://doi.org/10.3102/0013189X11428813

Bolojan, D., Vermisso, E., & Yousif, S. (2022). Is language all we need? A query into architectural semantics using a multimodal generative workflow. Proceedings of the 27th CAADRIA Conference: Post-carbon (pp. 353–362). CAADRIA. https://doi.org/10.52842/conf.caadria.2022.1.353

Chornobai, S. R., Ponzio, A. P., Bruscato, L. M., & Fagundes, C. V. M. (2023). Fostering Sustainability in the Early Stages of the Architectural Design process: a Creative Exploration of AI Generative Models. Proceedings of the XXVII International Conference of the Ibero-American Society of Digital Graphics, SIGraDi, (pp. 831–842). https://papers.cumincad.org/cgi-bin/works/Show?sigradi2023_439

Fang, L., Wang, X., Fang, C., Li, J. and Wang, S.J. (2025), Advanced Technologies in Creative Architectural Design: A Systematic Review. Adv. Intell. Disc, 1: 2400035. https://doi.org/10.1002/aidi.202400035

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *arXiv preprint* arXiv:1406.2661. https://doi.org/10.48550/arXiv.1406.2661

Jang, S., Roh, H., & Lee, G. (2025). Generative AI in architectural design: Application, data, and evaluation methods. Automation in Construction, 174, 106174. https://doi.org/10.1016/j.autcon.2025.106174

Li, C., Zhang, T., Du, X., Zhang, Y., & Xie, H. (2024). Generative AI models for different steps in architectural design: A literature review. *arXiv preprint* arXiv:2404.01335. https://doi.org/10.48550/arXiv.2404.01335

Mitchell, M. (2019). Artificial intelligence: A guide for thinking humans. Pelican Books

Montenegro, N. (2024). Integrative analysis of Text-to-Image AI systems in architectural design education: pedagogical innovations and creative design implications. Journal of Architecture and Urbanism, 48(2), 109-124. https://doi.org/10.3846/jau.2024.20870

Onatayo, D., Onososen, A., Oyediran, A. O., Oyediran, H., Arowoiya, V., & Onatayo, E. (2024). Generative AI Applications in Architecture, Engineering, and Construction: Trends, Implications for Practice, Education & Imperatives for Upskilling—A Review. Architecture, 4(4), 877-902. https://doi.org/10.3390/architecture4040046

Ponzio, A. (2022). Tecnologias Digitais para Uma Arquitetura Sustentável: Estratégia Design Intelligence. Anais do I Seminário Nacional de Formação, Atribuições e Atuação Profissional do CAU (pp. 127-129). CAU/SP.

Ponzio, A. P., Chornobai, S. R., Fagundes, C. V. M., Rodrigues, R. C., & Cunha Hafez José, G. (2023). Exploring Creative AI Thinking in the Design Process: The Design Intelligence Strategy. Proceedings of the 28th CAADRIA Conference (pp. 49–58). https://doi.org/10.52842/conf.caadria.2023.2.049

Schön, D. A. (1983). The reflective practitioner: How professionals think in action. Basic Books.

Schön, D. A. (1992). Designing as reflective conversation with the materials of a design situation. Knowledge-Based Systems, 5(1), 3–14. https://doi.org/10.1016/0950-7051(92)90020-G

Sedrez, M., & Celani, G. (2014). Ensino de projeto arquitetônico com a inclusão de novas tecnologias: uma abordagem pedagógica contemporânea. Pós, 21, 78–97.

Spiller, N., & Clear, N. (Eds.). (2014). Educating architects: How tomorrow's practitioners will learn today. Thames & Hudson.

Vissers-Similon E, Dounas T, De Walsche J. (2024). Classification of artificial intelligence techniques for early architectural design stages. International Journal of Architectural Computing, 23(2), 387-404. https://doi.org/10.1177/14780771241260857

Yu, W. F. (2025). AI as a co-creator and a design material: Transforming the design process. Design Studies, 97, 101303. https://doi.org/10.1016/j.destud.2025.101303