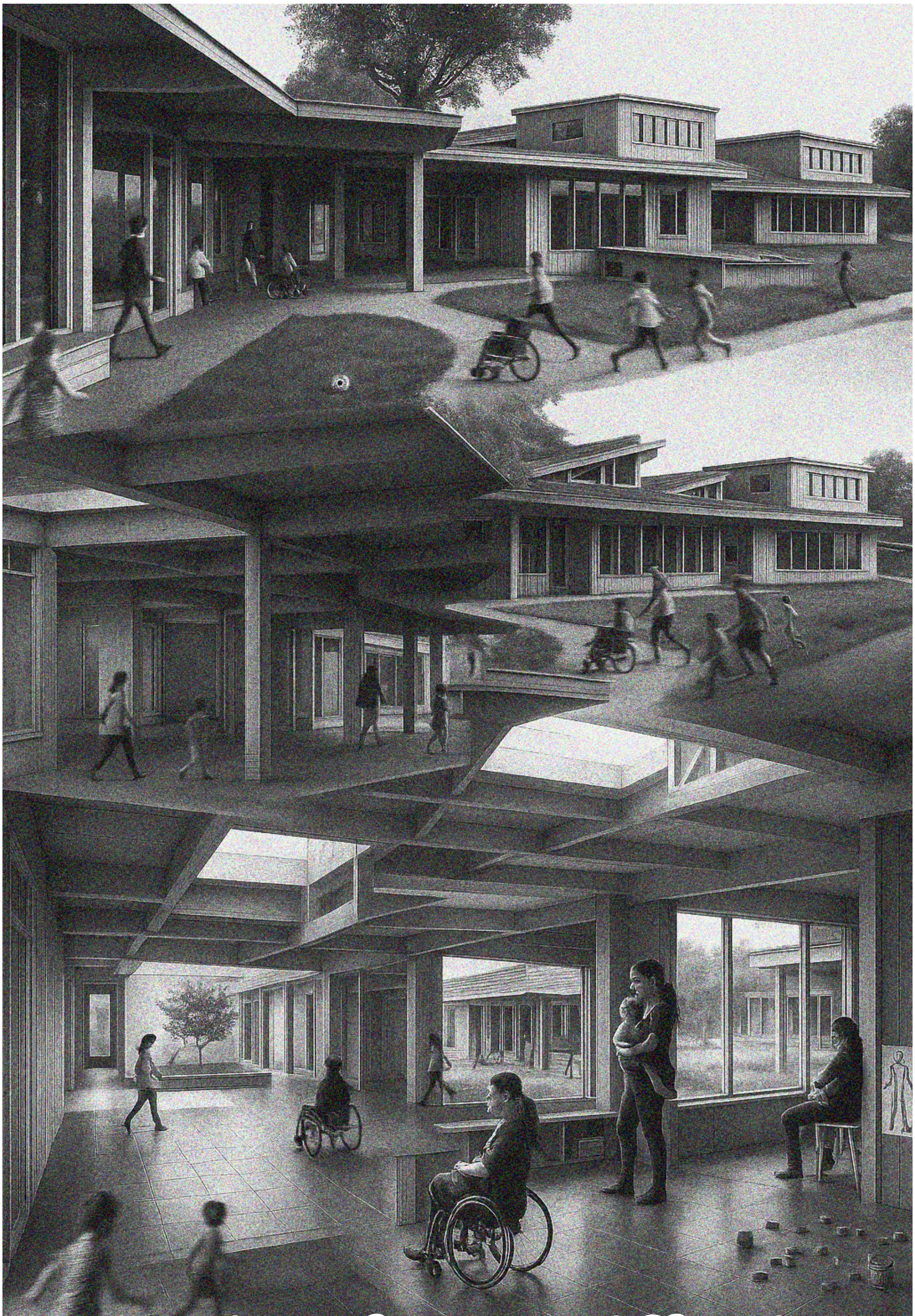


Drawing from Diffusion

The role of image generation amid the rise of AI in Architecture

Thomas Nelissen
Sebastiaan van Drie

Promotors:
Elien Vissers-Similon, Tine Segers

Studio Generative AI

The limits of photography cannot yet be predicted. Everything to do with it is still so new that even initial exploration may yield strikingly creative results. Technical expertise is obviously the tool of the pioneer in this field. The illiterates of the future will be the people who know nothing of photography rather than those who are ignorant of the art of writing.

– W. Benjamin

CONTENTS

1. Introduction

The evolution of generative Artificial Intelligence (AI) is reshaping the way we think about many professions, including architecture. More than just a new set of tools, it signals a shift in how architects approach their work, blending traditional intuition and expertise with computational capabilities. The questions driving this research are as follows: Can AI help us work more efficiently? Can it streamline specific steps in the design process? Perhaps most importantly, can it lead to better outcomes by expanding on the possibilities we explore? These aren't just technical questions; they touch on something more profound, prompting a reconsideration of the architect's role within an increasingly hybrid man-machine creative process.

Of course, integrating AI into the design process does not come without difficulties or friction. The outputs it generates are often highly polished, which can risk erasing the raw, imperfect, and frequently ambiguous qualities that give architecture its human dimension. The challenge is to utilise these tools in a way that respects and preserves the designer's spontaneity and personal expression. In many ways, the more precise the AI's suggestions, the less room there is for the designer's interpretation, something that becomes increasingly relevant as a design moves from concept to execution.

This research examines how generative AI models, specifically diffusion models capable of generating images, can enhance the architectural design process. Rather than treating generative AI as a shortcut or an aesthetic trend, we view it as a design companion, an extension of the hand and mind, capable of navigating complexity while still allowing space for doubt, intuition, discovery, and, perhaps most importantly, serendipity.

Disclaimer: The findings, observations, and conclusions presented in this thesis reflect the current possibilities and limitations of AI diffusion models and applications as they exist at the time of writing. It is essential to note that, given the rapid and continuous evolution of artificial intelligence, functionalities deemed impossible at present will likely become achievable in the near future. Therefore, all references to technical limitations should be interpreted as a snapshot of the current moment, being June 2025, rather than as permanent or definitive judgments.

2. Background

2.1 Diffusion Models

2.1.1 Definition

Diffusion models are a type of generative AI that can generate images. They operate by defining a forward process in which noise is gradually added to the data, and a reverse process in which a neural network learns to denoise this information, generating new samples. As described by Ho et al. (2020) in their foundational work, Denoising Diffusion Probabilistic Models, the model begins with pure random noise and iteratively refines this input through a learned denoising procedure. Trained on real data, the network progressively removes noise at each step, eventually producing a coherent and realistic output, such as a photograph or drawing. This iterative sampling process, visualised in Figure 1, is central to the model's ability to synthesise high-quality images from random initial conditions.

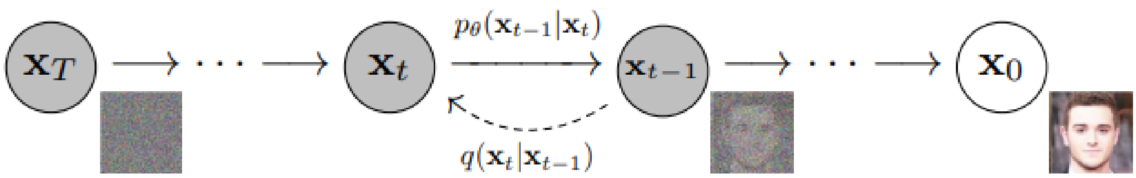


Figure 1. The workings of a diffusion model. (Source: arXiv:2006.11239v2 [cs.LG], https://doi.org/10.48550/arXiv.2006.11239)

2.1.2 Techniques in Diffusion-Powered Generation

Text-to-Image

Text-to-Image is the most widely used technique. A user provides a written prompt, such as “a wooden cabin in a misty forest,” and the model generates a corresponding visual interpretation, as shown in Figure 2. It’s a powerful tool for early-stage concept development, especially when quick iterations are needed.

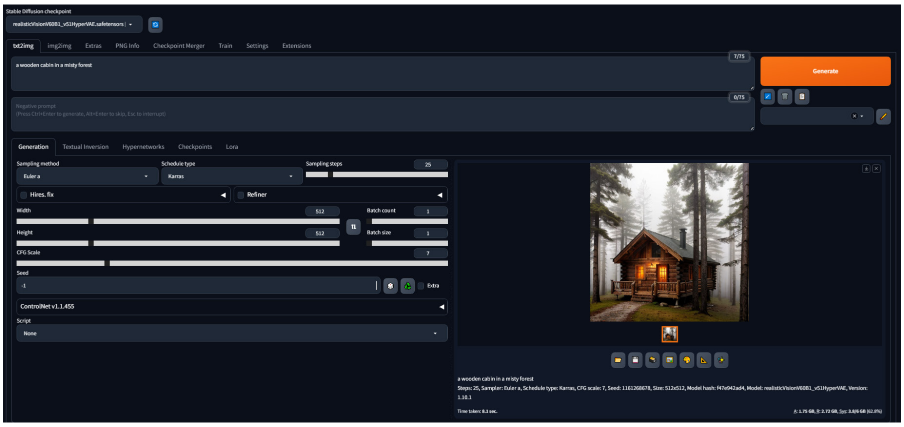


Figure 2. Text-to-Image technique in Stable Diffusion through A1111.

Image-to-Image

Image-to-Image starts with an existing image, combined with a prompt. The model uses the image as a base and transforms it while preserving its core composition. It’s useful for style transfers, refining generated images, or turning drafts into more refined visuals. Figure 3 illustrates how we can transform “a wooden cabin in a misty forest” into “a wooden cabin on a snowy mountain” while maintaining a similar image structure.

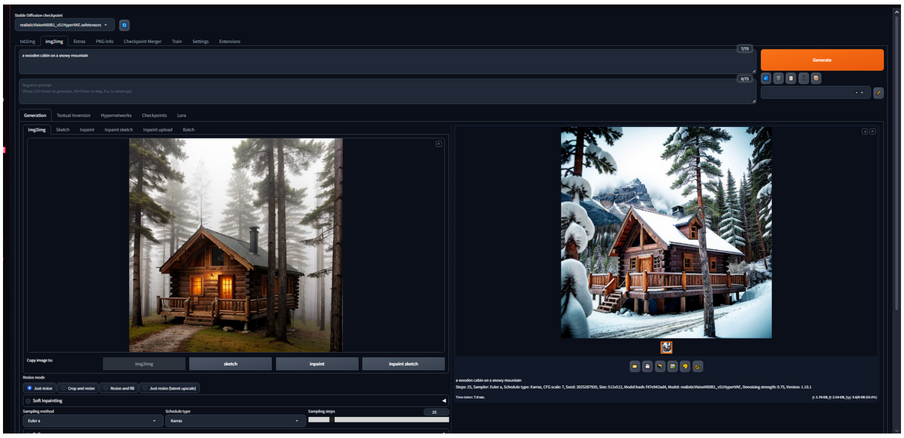


Figure 3. Image-to-Image technique in Stable Diffusion through A1111.

Sketch-to-Image

Sketch-to-image is a technique that utilises a simple line drawing or rough sketch as a structural guide for generating a detailed and coherent image. This process often relies on extensions, such as ControlNet, which allow the diffusion model to follow the input sketch while generating new visual content conditionally. By optionally combining the sketch with a descriptive text prompt, users can control not only the composition but also the style, mood, and context of the final output. The model then fills in the sketch with textures, lighting, and forms, producing a refined image that maintains the original layout while enhancing it with visually rich detail, as shown in Figure 4.

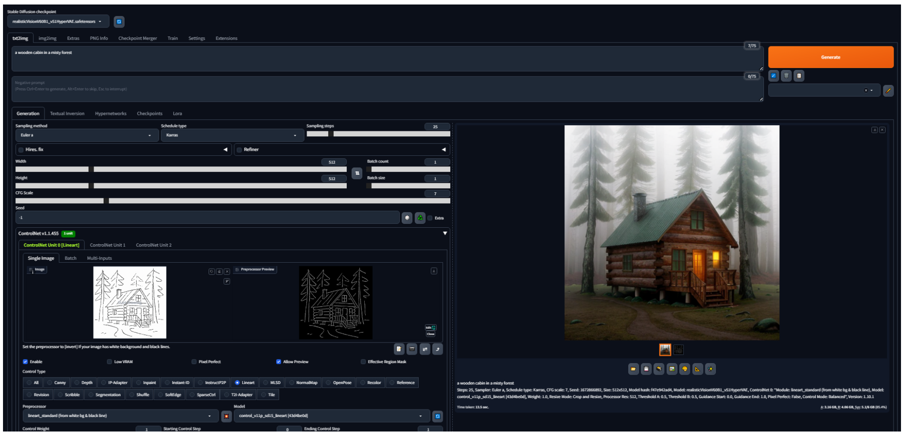


Figure 4. Sketch-to-Image technique in Stable Diffusion through A1111.

Video Generation

Text-to-Video and Image-to-Video techniques aim to create short video clips from text prompts or image sequences. While still being actively improved, they open new possibilities for animations, storytelling, or even conceptual walkthroughs in architectural environments. The lessons and conclusions from this research can almost be directly translated into moving images, possibly offering additional opportunities for animation and storytelling. This technique was not used in this research, but would be worth conducting further complementary research on in the future.[1]

Image-to-3D

Image-to-3D techniques are not native to diffusion models, but they can convert 2D images into 3D models by employing several additional steps. Typically, this involves generating multiple views of an object and then reconstructing it using tools such as mesh-based techniques. As this technology is still under development, it may be worth exploring in future research. [2]

1 Relevant applications supporting these techniques will be mentioned in the appendix.
2 Relevant applications supporting this technique will be mentioned in the appendix.

2.2 Diffusion Models in Architecture

In the field of architecture, the integration of pre-trained diffusion models, such as Stable Diffusion, DALL-E, and Midjourney, could rapidly reshape early-stage design workflows. These models, often accessed through APIs or cloud-based platforms, enable the quick generation of photorealistic images from text prompts or sketches. They can produce high-quality visuals within minutes, or even seconds. This could significantly reduce the time and cost associated with traditional rendering techniques, enabling more fluid transitions between ideation, visualisation, and presentation.

Leading figures and firms, such as Ismail Seleit[3], Cas Esbach[4], and MVRDV[5], are actively exploring the application of these tools, positioning them not as one-off experiments but as integral extensions of the design process itself. The use of diffusion models in the early phases of a project facilitates ideation by providing designers with an expanded range of geometric or atmospheric alternatives. In this context, speed is not simply a matter of convenience but a means to broaden the horizon of possibilities, allowing for the generation and evaluation of multiple variations before committing to a single concept.

Customisation of diffusion models is becoming increasingly common, with some offices fine-tuning existing models using proprietary datasets. This enables generated visuals to more accurately reflect the firm's specific aesthetics, regional context, and client expectations. Outputs can thus be tailored to particular material palettes, typological conventions, or spatial logics, strengthening their relevance in real-world applications. More advanced practices involve the creation of bespoke, in-house diffusion systems. For instance, Heatherwick Studio has been developing proprietary platforms[6] that support not only fast rendering and image editing but also integration with knowledge databases and BIM tools. These models are no longer limited to image generation but are becoming instrumental in lighting simulation, spatial analysis, and the optimisation of material and energy resources. As such, they are increasingly embedded in decision-making processes, enabling more informed evaluations grounded in a broader range of iterations.

3 (https://www.linkedin.com/in/ismailseleit/)

4 (https://www.linkedin.com/in/cas-esbach-176475b2/)

5 (https://www.mvrdv.com/themes/15/next)

6 (https://constructionmanagement.co.uk/how-heatherwick-humanises-buildings-with-technology/)

In the Flemish architectural landscape, there is a clear growing interest in the professional field, as evidenced by the fact that events such as 'Starten met AI als architect,' by NAV Academy[7], are almost instantly sold out. This suggests a widespread desire among architects to engage with the perceived 'hype' surrounding AI and to explore potential efficiency gains. However, despite this enthusiasm, a general uncertainty remains about how to effectively implement these technologies in practice, mainly due to the novelty of the field and the rapid pace at which it continues to evolve.

2.3 Research Question

This initial analysis offers a starting point for reflecting on the potential and limitations of diffusion models in design contexts. While seemingly full of potential, there is a fine line between arbitrary generation and adding actual value to design. Through these explorations, we can formulate a preliminary research question driving this thesis:

How might generative AI, diffusion models in particular, augment or reshape architectural design processes?

This research will focus on creating a 'real' application for these tools to formulate an answer to this question. A design will be developed through a hybrid man-machine process, where any potential applications for these tools are thoroughly explored, tested and documented, enabling us to draw structured and substantiated conclusions regarding this question.

7 (https://www.nav.be/nav-academy/starten-met-ai-als-architect)

3. Methodology

3.1 Case based research

To formulate a well-founded answer to the research question, a case-based research approach was chosen, with personal findings at the core. The cases are structured through a systematic analysis of design workflows, aiming to arrive at a conclusion that applies to as broad an audience as possible within the architectural and design fields. When artificial intelligence is mentioned in this research, it refers to generative AI in the form of diffusion models, unless otherwise specified. In cases where different types of generative AI, such as large language models, are discussed, this form of generative AI will be explicitly mentioned. In the next chapter, we will describe the tools we used; these explanations are not completely exhaustive technical descriptions, as this would fall outside the scope of this research. Instead, we opted for concise descriptions to clarify the roles each tool played in our process. Where relevant, references to more technical documentation and external sources are provided, allowing interested readers to explore the underlying technologies in greater depth.

3.2 Used Models

3.2.1 Stable Diffusion *Used as of October 2024*

Stable Diffusion is a latent diffusion model developed by Stability AI. It is capable of generating photo-realistic images based on textual input, pictures, or sketches. The model has gained significant popularity due to its open-source nature, high customizability, and its ability to run on consumer-grade GPUs. It's accessible through user interfaces such as ComfyUI and AUTOMATIC1111 and supports a wide range of creative and professional applications.

The user interfaces of Stable Diffusion in Automatic1111 (A1111) and ComfyUI reflect two distinct approaches to interacting with these models. A1111 offers a more traditional interface where users can adjust parameters such as prompts, resolution, and samplers through clearly defined input fields, as shown on the left in Figure 5. This makes it relatively accessible for new users and efficient for generating images rapidly.
In contrast, ComfyUI introduces a node-based interface that emphasises

modularity and transparency in the generation pipeline, as shown on the right in Figure 5. Each step, ranging from prompt processing to model execution and output handling, is visualised as a node, allowing users to customise and fine-tune their workflows more experimentally. While ComfyUI has a steeper learning curve, it provides greater flexibility and insight into the underlying architecture, making it particularly suited for advanced users and research-oriented experimentation.

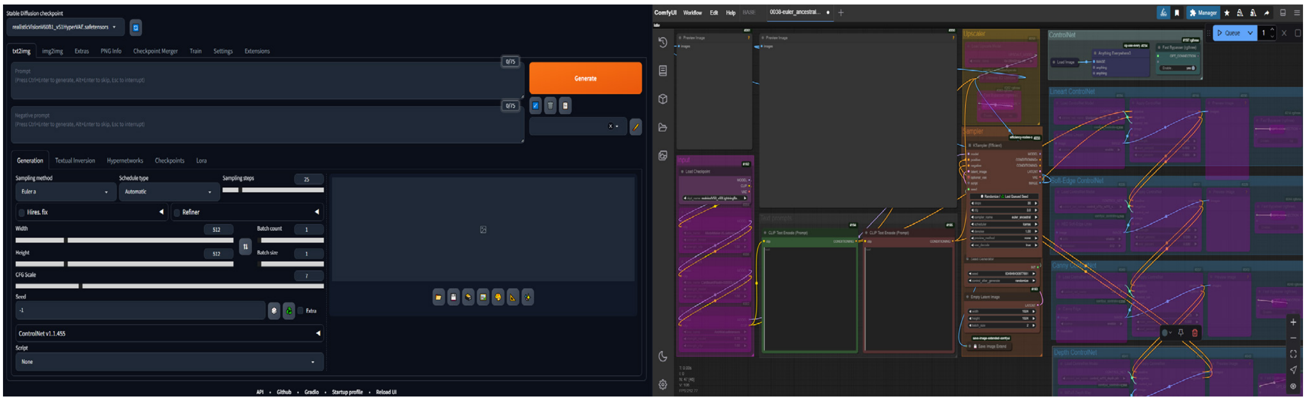


Figure 5. User interfaces of A1111 (left) and ComfyUI (right).

One of the model's key strengths lies in its flexibility and the ability to fine-tune outputs using extensions such as Low-Rank Adaptation, or LoRA for short. Initially proposed by Hu et al. (2021) in the context of natural language processing, LoRAs have since become widely adopted in diffusion-based workflows. Instead of retraining the entire base model, which may contain over a billion parameters, LoRA introduces small trainable matrices into specific layers. These lightweight additions enable efficient fine-tuning with significantly reduced computational demands and storage requirements. During inference, LoRAs are applied on top of a base model without altering its core parameters.

This approach proves particularly valuable in scenarios that demand rapid iteration and stylistic control, such as emulating the visual identity of an architectural firm or generating stylised imagery for material studies and conceptual explorations. By training a LoRA to replicate a specific artistic or rendering style, users can produce highly tailored outputs without needing to retrain the entire model.[1] However, excessive or poorly executed use of LoRA can negatively impact output quality, introducing overfitting or visual artefacts.

1 LoRAs are inherently linked to the base model on which they were trained, meaning that transitioning to a different base model may require retraining the adaptation. More information on Loras: https://www.nextdiffusion.ai/tutorials/how-to-install-and-use-lora-models-for-stunning-images-in-stable-diffusion.

Several complementary techniques further expand the design potential of diffusion models in architectural workflows. Inpainting enables users to select and mask specific parts of an image, allowing them to regenerate only that area. This is particularly useful for replacing the environment, redesigning building elements, or correcting visual details. Outpainting, on the other hand, expands the canvas by imagining what could exist beyond the original image's borders, an effective strategy for extending scenes or recontextualising an image within a new layout.

ControlNet[2], also known as guided generation, introduces a layer of precision to the generative process by incorporating elements such as depth maps, edge outlines, or pose references. This technique offers a balance between creative freedom and structured output, which is especially valuable in architectural contexts. For example, a design's geometry can be preserved while exploring alternative materialities.

Additionally, the possibility of using multi-input workflows is a decisive factor in tool selection. When multiple images can be provided, such as one image defining geometry and another indicating natural lighting, the refinement in output quality becomes substantial. In contrast, workflows limited to single-image inputs often involve a higher degree of guesswork on the part of the AI, leading to less desirable or inconsistent results. While the versatility of Stable Diffusion makes it an attractive tool, the quality of the generated images remains highly dependent on the user's ability to construct effective prompts. Inexperienced users may struggle to obtain consistent results without first acquiring sufficient experience. Even with our exploration lasting weeks, achieving desired results often required a very technical approach and several attempts. Moreover, its open-source nature raises ethical concerns[3], including the potential for misuse in generating NSFW content or copyright infringement, which reinforces the need for responsible deployment.

Disclaimer: It is essential to recognise the complexity of determining the presence of intellectual property in the outputs of generative models. Given the vastness of the datasets on which these models are trained, it is difficult, if not impossible, for us to make definitive claims about potential instances of IP infringement. In the context of this research, we assume that such concerns do not pose a direct problem for the specific use cases explored. However, we recognise that questions surrounding authorship, ownership, and the ethical use of training data are highly relevant and could merit dedicated investigation beyond the scope of this thesis.

2 https://stable-diffusion-art.com/controlnet/ https://github.com/lllyasviel/ControlNet
3 https://milvus.io/ai-quick-reference/what-ethical-considerations-are-involved-in-deploying-diffusion-models

3.2.2 ChatGPT *Used as of February 2025*

ChatGPT is a transformer-based large language model developed by OpenAI, designed to generate coherent and contextually relevant text from written prompts. It can also support image generation workflows by refining prompts or, when integrated with models like DALL-E 3, generating images directly within the chat interface.[4]

Built on the GPT-4 architecture, ChatGPT can assist with a wide range of tasks, including writing, coding, summarising, translating, and answering questions. The model has become widely adopted due to its conversational fluency, broad general knowledge, and accessibility through both free and professional interfaces. ChatGPT is a valuable tool in image generation, both for its ability to enhance prompts for use in other software, such as Stable Diffusion, and for generating images itself in chats.

Since February 2025, image generation in ChatGPT-4o has been powered by DALL-E 3, the most recent version of OpenAI's text-to-image model, which introduces notable improvements in prompt understanding, image quality, and semantic consistency. ChatGPT-4o excels at translating complex and nuanced prompts, which DALL-E 3 can convert into visual output that is both coherent and aesthetically refined. Integrated directly within chats, it allows users to refine prompts conversationally and modify generated images afterwards, for example, by adding or removing elements, enabling an iterative process.

Unlike Stable Diffusion, ChatGPT is a closed-source model, meaning users do not have access to its internal architecture and cannot run or fine-tune the model locally. In addition, the system includes built-in content filters that block the generation of specific topics regarded as NSFW, which can occasionally affect seemingly innocuous prompts. While such restrictions may appear limiting to advanced users, they reflect OpenAI's broader efforts to ensure safe and responsible use of its products.

In architectural and design contexts, ChatGPT-4o, combined with DALL-E 3, provides an efficient and accessible means to generate concept visuals, explore stylistic directions, or communicate ideas to clients and stakeholders. The integration within chats facilitates a natural and accessible user interface where image generation is closely tied to textual analysis and iteration.

4 OpenAI. (2022). ChatGPT [Large language model]. https://chat.openai.com/

3.2.3 Sora *Used as of April 2025*

Sora is a video-based latent diffusion model developed by OpenAI, designed to generate high-fidelity video clips by denoising spatiotemporal latent representations conditioned on textual prompts. It incorporates large language models to enhance semantic understanding and can also generate still images via its integrated image-based systems.[5]

Like Stable Diffusion, it employs a latent denoising process to generate visuals, but extends this approach to the temporal domain by treating video as a 3D structure. A transformer backbone allows it to model complex motion, lighting, and camera dynamics across frames.

Where ChatGPT generates text and Stable Diffusion generates images, Sora combines both modalities to produce time-consistent, coherent, and optionally moving images. It is trained on large video-text datasets and enables new forms of visual storytelling, including in architecture, where atmosphere, usage, and movement can be communicated through generated filmic sequences. In this research, Sora was utilised for its singular image generation capabilities, enhanced by its extensive contextual understanding of prompts.

3.2.4 Photoshop Generative Fill *Used as of November 2024*

Generative Fill is a context-aware generative feature powered by a diffusion model based on Adobe Firefly[6], integrated into Adobe Photoshop, enabling users to generate, extend, or replace image regions based on selection and textual input. The system considers the surrounding context to ensure seamless blending and realistic outputs.[7]

The tool functions by analysing masked regions and generating new content based on user input and the surrounding context. Unlike traditional clone or patch tools, Generative Fill synthesises new pixels using a text-conditioned generative model, offering more semantically aware and visually coherent results. This allows for quick, highly targeted edits, or content-aware scene expansions.

While it shares conceptual roots with inpainting tools in models like Stable Diffusion, Generative Fill distinguishes itself through its user-friendly interface, real-time rendering capabilities, and seamless integration

5 OpenAI. (2024). Sora. https://openai.com/sora
6 Although Adobe Firefly was not used in this research, it can be a valuable tool for generative design. The decision to exclude it was based on observed limitations in output specificity and the relatively low degree of fine-grained control compared to other diffusion-based tools. For more information on Adobe Firefly and its capabilities, see: https://www.adobe.com/sensei/generative-ai/firefly.html.
7 Adobe Inc. (n.d.). Adobe Photoshop. https://www.adobe.com/products/photoshop.html

within the broader Adobe Creative Suite. It operates through a cloud-based diffusion model trained on Adobe's datasets, emphasising ethically sourced and licensed content. In architectural design, this tool proves helpful for quick visual iterations, environment substitution, or post-production enhancement of renderings and visualisations.

3.2.5 Summary

The tools selected for this research were chosen based on their accessibility, level of control, and compatibility with architectural workflows. Stable Diffusion served as the primary model due to its high degree of customisability compared to other commercial alternatives. Its open-source nature, the ability to integrate LoRAs for efficient and style-specific fine-tuning, and support for extensions such as ControlNet made it a highly flexible environment for experimentation. It is important to note that leveraging these features requires a certain level of expertise with this tool.

ChatGPT proved valuable as an accessible, complementary tool, particularly in refining prompts and translating architectural concepts into precise textual inputs. Its language structuring capabilities contributed to more predictable and semantically rich outputs when used in conjunction with diffusion models, as well as in the images generated by DALL-E 3.

The same criteria that motivated the choice for Stable Diffusion, such as multi-input capability, also justify a gradual shift towards Sora, as it combines visual coherence with high image quality, a strong contextual understanding, and a large amount of control in an increasingly accessible interface.

Photoshop's Generative Fill feature stood out for its intuitive integration into familiar design processes, offering a highly accessible option for localised edits, scene expansion, or quick post-processing. However, it lacks the deeper contextual and prompt-based understanding that models like ChatGPT or Sora provide, making it more suitable for minor changes rather than conceptually driven image generation.

In general, time and efficiency became a practical threshold; if a generative tool slows down rather than supporting ideation, efficiency, or communication, it risks undermining its intended value. It is essential to note that every tool we explored was given a few weeks for testing and learning, allowing us to form an informed opinion on its potential in various applications. First impressions were not a defining factor; we don't want to miss out on potential added value because a tool seems impractical. This can be seen in the choice of Stable Diffusion, for example; a few weeks were necessary to overcome the steep learning curve of creating qualitative content with these models. Still, the added value of some techniques or functions can outweigh the time or efficiency constraint. The integration of AI tools into our processes follows key industry releases and advancements. Figure x illustrates the evolution of this landscape. As previously mentioned, we began actively incorporating these tools starting in October 2024 with Stable Diffusion, followed by Photoshop Generative Fill in November 2024, ChatGPT-4o in February 2025, and Sora in April 2025

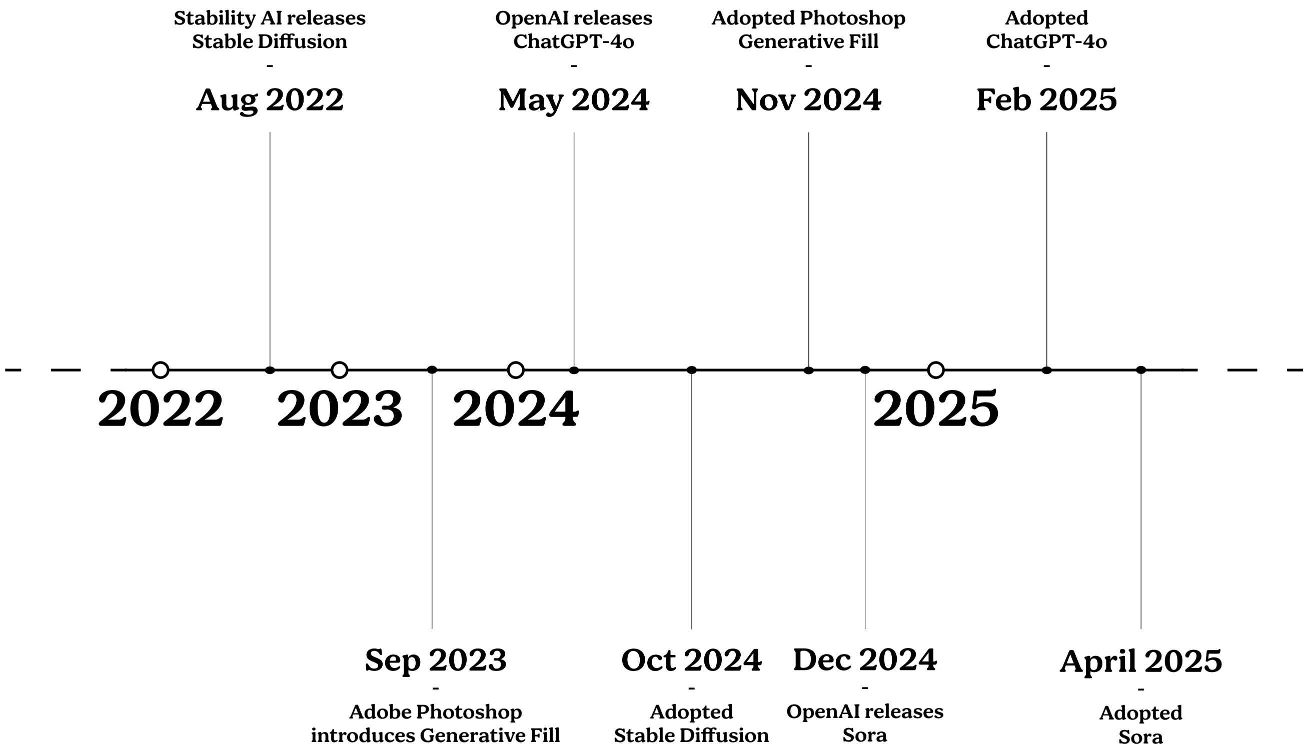


Figure 6. Timeline of the release date and adoption date of used tools.

DISCLAIMER: The selection of diffusion models and interfaces in our work is based on the specific demands of our creative and technical processes. Different models come with distinct training datasets and optimisation, making them more suitable for different tasks. While we favour specific tools for their compatibility with our current objectives, we recognise that other models may be better suited to different applications or user preferences. As the generative AI landscape evolves, we remain open to adapting our toolkit accordingly.

3.3 Exploration

This research begins with an exploratory phase focused on three architectural fragments: Promenade Architecturale, Inside-Outside, and Transparency. These fragments serve as guides for investigating the spatial, material, and experiential aspects of architecture. At the same time, we aim to create scenes that could also be part of our future design brief: a primary school for students with type 2 special needs. This is a choice to guide future generations, ensuring their usability in a later design phase, rather than randomly creating images of spaces with no further application.

Each fragment follows three steps. First, a conceptual inquiry asks what the fragment could encompass in architectural terms. Second, strategic generation is applied to visualise the fragment using text-to-image, image-to-image, and sketch-to-image techniques. Third, the resulting images are annotated to extract design values embedded within the visual output.
It is essential for us that the interpretation of these fragments is personal. There is no clear or defined way to view these architectural elements as a designer. This step is done in parallel, and thus, both of us choose our architectural references based on our interests and interpretations. This initial analysis and the resulting list of architectural elements and qualities in the studied works are crucial in creating a subjective frame of reference and a basis for prompting in the second step.
The third step of analysing the generated images is based on traditional annotation techniques, such as sketching by hand. The annotation method is more than just marking up an image; it's a reflective process that starts from the generated visual material and builds towards critical insight and interventions. It combines analysis with design intuition and unfolds in three steps:

1. Selecting the Right Image

The process begins by choosing a generated image with architectural or conceptual richness, informed by our experience and education. It may reveal interesting proportions, material use, or hint at a structural concept. The image is never complete but has the potential to trigger interpretation and development.

2. Annotating with Depth

Using sketches, diagrams, and short notes, the image is explored from different angles, such as structural, spatial, and experiential perspectives. Annotations raise questions and reveal relationships, rather than merely describing what is visible; for example, how does light enter the space? How is circulation guided? What atmosphere is created?

3. Intervening Through Design

The final step is speculative. Based on the earlier observations, the architect suggests extensions or adjustments, a floor plan fragment, a possible continuation, or a different reading of the space. These proposals are not arbitrary; they follow from a personal design logic and way of thinking.

We also include external annotations from fellow students in this analysis. This approach serves two key purposes. First, it introduces an additional layer of critical insight to the imagery, helping us move beyond our own biases or interpretations. Second, it provides qualitative reflections on what is visibly present in each image, especially in cases where architectural elements or material choices may be ambiguous or abstract due to the generative nature of the visuals. By inviting others, primarily bachelor's students of interior architecture at the University of Antwerp, to annotate, we can document their observations, test their interpretations, and reduce the potential for overly subjective conclusions. The act of annotating volumes or architectural components also encourages readers to engage with the logic and composition in a more tactile manner, even if some visual cues, such as human figures, are deliberately vague to avoid critique based solely on their generative quality.

3.4 Design

3.4.1 Design Process

In this thesis, we simulate a real design process by virtually participating in an open call, an architectural design competition organised by Team Vlaams Bouwmeester[8]. As designing is inherently non-linear and deeply personal, we believe that demonstrating the potential of diffusion models merely on paper is insufficient; a genuine design experiment is necessary. Through this approach, we aim to capture the process as accurately as possible in both written and visual material. The structure of the design process we will carry out is primarily informed by the key insights gained during the exploratory phase.

8 Team Vlaams Bouwmeester is the support team behind the Flemish Government Architect. https://www.vlaamsbouwmeester.be

To support this experiment, we focus on producing conceptually strong and visually compelling material rather than exhaustive technical elaboration, aligning both with our timeframe and the objectives of our thesis. The selected competition for this experiment, Het Mispeltje in Sint-Niklaas, reinforces the relevance of this approach. This small-scale primary school for children with type 2 special needs, which means they can have mild to moderate disabilities[9], calls for a response grounded in empathy, material sensitivity, and spatial clarity. Through this project, we explore how workflows utilising image generation can contribute to detailed, intimate, and socially meaningful design, thereby challenging the narrative of AI-generated architecture as necessarily large-scale or generic.

Early in our design process, we will focus on five distinct architectural 'Moments' that, in our view, encapsulate the essence of this project. Rather than attempting to design every aspect of the building in equal detail from the beginning, we identify these moments as anchor points, scenes through which the core architectural intentions could be communicated and tested. This approach aligns with a broader design philosophy, as seen in the work of Rem Koolhaas[10] , and these curated moments serve as narrative fragments that guide both users and designers through the project's conceptual and spatial landscape. More recently, this can be observed in practices such as the Ghent-based office AJDVIV, where architecture is curated through carefully composed sequences rather than an exhaustive totality. These Moments provide a framework for decision-making, enabling us, as architects, to maintain focus on the experiential quality of the architecture, rather than being overwhelmed by its totality. Each moment is textually defined early on and used as a lens through which spatial decisions can be evaluated. The idea is not to design everything at once, but to concentrate the architectural experience on several particular and evocative encounters, which, in our case, allows for a clear objective when generating images.

In the later stages of the design process, we shift from our exploratory and diffusion model-powered workflows to more traditional architectural interventions, allowing us to consolidate and refine the design into a coherent proposal. These interventions, while grounded in familiar architectural methods, are still strongly informed by the speculative outcomes of previous steps, including generated images and textual descriptions.

9 *https://www.belgium.be/nl/Leren/onderwijs/buitengewoon_onderwijs*
10 https://architecture-history.org/architects/architects/KOOLHAAS/biography.html

3.4.2 Design Evaluation

We have chosen to virtually participate in a competition because it provides real-world constraints and offers the opportunity to engage in dialogue with Team Vlaams Bouwmeester[11], allowing for authentic feedback on the quality of the design. This creates a simulated interaction with a client, enhancing the realism of the design scenario. Team Vlaams Bouwmeester has limited experience with AI tools and a limited understanding of the subject. Consequently, they raise questions about the actual benefits of using such models in terms of speed, quantity, quality, or other advantages.

The competition format also provides a concrete benchmark through the predefined timeframe of 600 hours for the design. We will be able to make a comparison to this limit by tracking our design hours and gather information about the additional time typically required by other experts in a conventional design team. This facilitates comparison with an average team that does not use these image-generation tools, helping to reflect more critically on the role and potential impact of generative AI within architectural practice.

This format aligns with our ambition to deliver a strong and coherent architectural concept within a defined scope. Competitions tend to emphasise clarity, atmosphere, and vision over technical detail, which suits the design we aim to explore. Working within the context of a competition allows us to clearly distinguish between two analytical layers: the design layer, which focuses on producing an architecturally convincing and visually refined proposal, and the research layer, which demonstrates our critical control over the generative processes. We document how these tools were applied, adapted, and evaluated, with authorship remaining firmly in our hands. Additionally, to critically assess the effectiveness of the proposed design workflows, an expert evaluation is conducted as a final step. This evaluation serves a dual purpose: on the one hand, it provides qualitative insights from professionals across architecture and adjacent disciplines, and on the other, it enables a quantitative comparison of key performance indicators to measure how the approach might be applied in a real-life architectural context.

Given that the design output of this project takes the form of an initial, competition-style proposal developed by students, it is necessary to acknowledge certain limitations in experience and domain expertise. While the exploration and design phases offered valuable insights, the qualitative dimension of the project requires assessment by those with a more established professional background. Experts with diverse profiles, including a practising architect, an engineer, and a government official, will critically review elements such as coherence, conceptual depth, feasibility,

11 Team Vlaams Bouwmeester is the support team behind the Flemish Government Architect. https://www.vlaamsbouwmeester.be

and contextual sensitivity. This evaluation enables a more nuanced understanding of the project's strengths and shortcomings. Their feedback reflects the assumptions we made during the design process and helps validate the qualitative aspects of using generative AI within our design workflow.

Complementing this qualitative layer, a quantitative evaluation addresses the practical metrics that determine the viability of the proposed workflows in real-world applications. These include measurements such as the time invested in different stages of the process, the number of individuals or roles required to execute the work, the accuracy of outputs from both architectural and engineering standpoints and the potential cost implications. These metrics enable a more informed comparison between traditional workflows and AI-assisted alternatives. They help determine whether the introduction of generative tools leads to measurable efficiency gains or instead adds new layers of complexity. Additionally, this evaluation includes a personal reflection on the design, assessing the extent to which our design intentions were achieved.

3.5 Cataloguing Workflows

We will systematically analyse the various workflows tested, evaluating them not only in the form of general takeaways from the exploratory phase but also as they emerge during the actual design process. Here we implement the analytical framework that will be used consistently throughout this research. It takes the form of a comparative analysis table, as illustrated in Table 1. It includes a fixed set of criteria: the time required to achieve a result, the use case, the tools it uses within steps, and additional relevant criteria These criteria are not selected arbitrarily, but because they collectively contribute to answering the core research question: How generative AI, diffusion models in particular, might reshape or augment architectural design processes.

	Descriptive name to identify the workflow
Software	Tools used during the workflow
Duration	Estimated time required for this workflow, ranging from Short (5-15 minutes) or Medium (30-60 minutes) to Long (60-120 minutes) and Extended (120+ minutes)
Use Case	Specific phases or moments in a design where this workflow is useful
Input Type	Format and nature of input (images, text, plans...)
Man-Machine Ratio	Ratio of human involvement compared to automated generation
Limitations	Main challenges, constraints, or downsides of the workflow

Table 1. Structured table for workflow analysis.

Disclaimer: The estimated durations and evaluations included in the comparative analysis tables are based on our own experience and familiarity with the tools used. We assume that users engaging with these workflows possess at least a similar level of proficiency with generative AI tools, image editing software or architectural design methods. Readers with significantly less or more experience may find the time investment and complexity of specific workflows to differ accordingly.

3.6 Reflection

Our role as the designers strongly influences this research, as we essentially served as the test subjects throughout the process. Beyond the inherent individuality and non-linearity of any design trajectory, this personal involvement has significantly shaped the course and outcomes of the experiment. For this reason, we will include our reflection in line with the main conclusions of the thesis. This reflection aims to frame and nuance the findings through our perspectives, acknowledging that the conclusions drawn here may not universally apply to every designer or design context. Other designers can interpret and apply the conclusions drawn here to their processes.

4. Exploration

4.1 Fragments

4.1.1 Fragments: S. van Drie

Promenade Architecturale ***October '24***

For me, the Promenade Architecturale is about designing with movement in mind. A great example of the quality I perceive in this fragment is James Stirling's Neue Staatsgalerie in Stuttgart (Fig. 7), which serves as my primary reference for this analysis. The Promenade Architecturale is not just a way to connect spaces, but a way to shape how architecture is experienced over time. I see it as a sequence that unfolds gradually, guided by changes in light, material, scale and direction. These shifts help define transitions and shape the perception of space. Circulation and function often blend; a hallway might feel like a room and vice versa. Pauses, in the form of courtyards, double-height spaces, or framed views, create moments of clarity or reflection. What makes it powerful is the mix between large and small, abstract and specific. It draws both the body and the eye through space, turning architecture into a lived experience rather than a fixed image.

Inside-outside ***November '24***

For me, the relationship between inside and outside is a central architectural theme; it's about how a building interacts with its surroundings, both physically and experientially. In James Stirling's Staatsgalerie in Stuttgart, this connection is active and layered. Ramps, terraces and open courtyards extend circulation outdoors, while materials often flow from façade to interior, blurring boundaries. Rather than a sharp divide, I see inside and outside as a continuum. Through thresholds, openings and spatial overlaps, architecture can either open up to its context or guide how that context is experienced. This creates a dynamic, public quality that invites movement and engagement, visualised in Figure 7.

Figure 7. Facade of Neue Staatsgalerie. (Source: ©Flickr User: pov_steve)

Transparency *December ‘24*

Transparency in architecture, in my view, goes beyond visual openness; it’s about how a building communicates its spatial logic, structure and intent with clarity. It’s not just about glass or open views, but about how the organisation is revealed, how movement is guided and how materials express their role. In James Stirling’s work, as seen in the Staatsgalerie in Stuttgart or the University of Leicester Engineering Building (See Fig. 8), transparency is layered and dynamic. He uses ramps, framed views and contrasting materials to create depth and orientation. There’s a play between classical references and modern elements, which brings complexity but also legibility. I view transparency as a means to help users understand a space, not by exposing everything at once, but by inviting exploration and gradually unfolding the design.

Figure 8. Facade of Neue Staatsgalerie.
(Source: ©Canadian Centre for Architecture)

4.1.2 Fragments: T. Nelissen

Promenade Architecturale *October ‘24*

I see the Promenade Architecturale as inseparable from the human body and its sensory engagement with space. My understanding of this concept was shaped by Juhani Pallasmaa’s The Eyes of the Skin[1], which made me realise that architecture is not something we merely see; it’s something we feel, hear, and physically inhabit. Especially in our design brief, a primary school for children with special needs, this embodied approach is not optional; it’s essential.

The promenade is not just a route through a building, but a sequence of bodily experiences. The way light shifts through a corridor, the feel of a handrail, the contrast between wooden floors and cool concrete underfoot. These are not details, but fundamental parts of how architecture communicates. Movement becomes narrative, and every transition becomes a moment of engagement.

Designing this means paying close attention to scale, texture, and rhythm. A hallway must be more than just circulation; it must offer pauses, encounters, perhaps a glimpse outside, or a place to sit. The architecture must meet the child at their level, both physically and emotionally.

1 Pallasmaa, J. (2012). The eyes of the skin (3rd ed.). John Wiley & Sons.

Inside-outside

November ‘24

For me, the relationship between inside and outside is not a boundary but a carefully shaped transition. Marie-José Van Hee’s architecture (See Fig. 9), particularly her house in Gent, captures this with clarity, a central courtyard that creates calm through rhythm, repetition, and proportion. It’s not just about openness, but about structuring serenity.

These values resonate with what a type 2 primary school ought to nurture, a place where overstimulation must be avoided and orientation must be gentle and gradual. I view architecture as a means to guide rather than frame the context through layered thresholds, covered walkways, and the gentle filtering of light and air. These spaces don’t separate the inside from the outside; they allow the two to blend in a way that feels natural and reassuring. Inspired by Van Hee, I aim for a quiet architecture where the structure brings peace, and where the transition from public to private becomes an experience, not abrupt but unfolding, steady, and clear. The courtyard becomes more than an open space; it is a spatial pause, a moment of breath between inside and outside.

Figure 9. a+u magazine on MJVHA.
(Source: De Vocht & Van Sande, 2021)

Transparency

December ‘24

Transparency in architecture, to me, is about readability, not visibility. It’s not about exposing everything at once, but about making a building’s logic legible through geometry, order, and the rhythm of space. In the work of Pezo von Ellrichshausen (See Fig. 10), I find this clarity in their use of strong, simple forms and structured plans. Their architecture doesn’t shout; it guides. This idea of transparency resonates deeply with our school design. Children with special needs benefit from spaces that are clear, consistent, and easy to navigate. I believe a good building offers orientation through calmness, utilising spatial hierarchy, repetition, and material honesty to help users understand where they are and where they’re going.

Figure 10. The architecture of Pezo von Ellrichshausen. (Source: El Croquis 214, 2022)

4.2 Resulting Images

4.2.1 Resulting Images: S. van Drie

Promenade Architecturale

To explore the concept of the Promenade Architecturale, I began by examining reference material that demonstrates how movement can be integrated into architectural form. Rather than simply replicating these projects, my approach was to reinterpret their underlying spatial strategies, focusing on qualities such as rhythm, sequence and structure. This involved abstracting the compositional logic and translating it into new structures that would guide a visitor through the space. Key to this was the preservation of spatial cues, including framed perspectives, changing scales, structure, and transitional moments that encourage a gradual unfolding of space. The resulting images are therefore not static representations but spatial propositions, explorations into how architecture can invite movement, define pace or construct a narrative through their built form.

The first process begins with a reference image of the HfG School of Arts by XDGA, as shown in Figure 11. Step 2 involved transforming the hallway into a wooden structure through Image-to-Image, retaining the original geometry. Step 3 shows the result of another pass in Image-to-Image, recontextualising this structure to a pavilion in a park area. After this, Photoshop's Generative Fill feature was used to correct any undesired elements, as shown in step 4. The final step consisted of expanding the scene with Generative Fill.

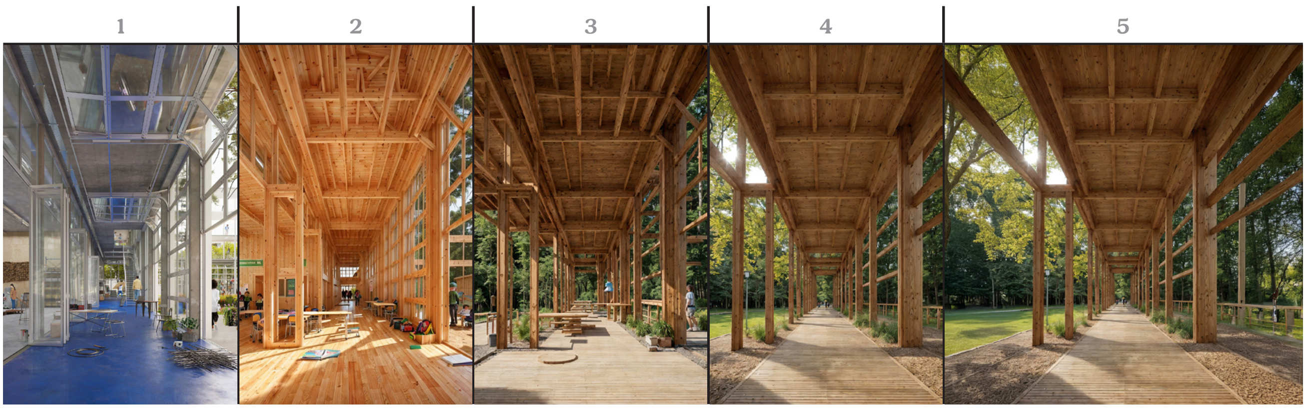


Figure 11. Process of reinterpreting a school as a pavilion. (Source base image : XDGA)

The extended colonnade constructs a clear path that emphasises movement. The regular rhythm of the timber structure generates a cadence, inviting the body to walk while framing views toward the surrounding landscape. Light filters through the open sides, activating the space and reinforcing a sense of progression. The architectural sequence is not static; instead, it encourages variations in pace or perception through its materiality and layering.

The second process begins with a reference image of the RDO House by Cechvala Architects, as shown in Figure 12. A first pass through Image-to-Image reinterprets the building as an open structure in a park area. Step 3 shows the result of an expansion of the scene by using Photoshop Generative Fill. The resulting image was created by reinterpreting this structure as a new construction using the Image-to-Image technique.

Figure 12. Process of reinterpreting an open structure as a building. (Source base image : Cechvala Architects)

The building similarly engages a notion of movement through a series of thresholds and articulated transitions. The use of timber reinforces continuity between architecture and site. At the same time, deep overhangs and recessed volumes blur the line between inside and outside, offering a gradual unfolding of space rather than an abrupt entrance.

Inside-Outside

To investigate the theme of inside-outside relationships, I aimed to test how effectively the following images could be generated through prompting alone, thus without relying on extensive post-processing or refinement. The aim was to create images that directly reflected the spatial principles described in my textual analysis. By focusing on the formulation of clear and intentional prompts, I was able to produce an extensive series of variations. This large output enabled a careful curation process, selecting those images that best embodied the qualities I aimed to explore. The selected visuals reflect a nuanced interplay between interior and exterior conditions and were chosen not for their representational perfection, but for their capacity to spatialise the concepts of threshold, extension and contextual responsiveness in a legible way.

In this first image (Fig. 13, left), the covered walkway forms a porous threshold, where a sequence of slender columns and open glazing creates a soft transition between the interior and the courtyard. The overhanging concrete slab provides spatial depth while offering shade, reinforcing a sense of shelter without enclosing the space entirely. Visual transparency across the courtyard and into adjacent rooms allows for a layered reading of activity, extending the experience of the interior into the outdoors.

This second image (Fig. 13, right) similarly constructs this continuity through a series of spaces. A red structure articulates the passage from one courtyard to the next, guiding movement while framing views and anchoring the building in its landscape. Openings are generously sized and often flush with the façade, reducing the threshold to a minimal gesture. The focus here seems not to separate, but to integrate, the architecture serves to regulate interaction with the surroundings.

Figure 13. Resulting images for Inside-Outside fragment.
Prompt: 'illustration of a courtyard of a low, single story, small primary school, connection to the interior through three facades, composition of the windows on the ground floor makes the rhythm of the canopy on the second floor, architecture'

Transparency

While exploring the theme of transparency, I chose to work with the representation of architectural models rather than scenes within the architecture. This decision was grounded in the idea that transparency encompasses not only visual openness but also the legibility of architectural intent. Just as in the traditional design process, where physical models often help clarify spatial organisation and volumetry, their image-based counterparts can serve a similar analytical role. This exploration also served as a first experiment in combining different LoRAs to create a desired result. By combining multiple qualitative models focused on generating architectural models, section models, and cardboard, I was able to make a consistent style reminiscent of a photo studio for architectural models. The generated images reflect this dual focus: the larger massing model highlights the organisation and hierarchy of exterior volumes, while the sectional model invites a reading of the internal logic of the building, revealing circulation, structure and spatial layering. By treating the image of a model itself as a medium of communication, I was able to test how clarity and transparency can be conveyed without relying on excessive detail or an actual physical model.

In this first image (Fig. 14, left), the composition around a courtyard clearly articulates the organisation of the building. The arrangement of volumes around a central void allows the logic of circulation and programmatic zoning to be intuitively grasped. Thresholds, overhangs and framed openings suggest permeability and rhythm, while the orthogonal layout creates a calm and readable system. The spatial hierarchy is not hidden, but instead structured in a way that subtly guides users through the ensemble.

In this second image (Fig. 14, right), a sectional model reveals how transparency can emerge through structural clarity. The stepped roofline and exposed circulation elements, such as the staircases and voids, articulate movement and vertical relationships within the building. Rather than a continuous whole, the image of the model communicates transparency through the sequencing of spaces, material layering and spatial overlaps.

Figure 14. Resulting images for Transparency fragment.
LoRAs: ModelMaker-XL, Cardboard Style

4.2.2 Resulting Images: T. Nelissen

Promenade Architecturale

Having amassed a series of reimagined references, this exploration culminated in a collage that consolidated generative outputs into a single proposition. Three photographs were selected for relevance to the promenade theme: the timber canopy at Charles-Henri Tachon's school in Lyon (Fig. 15, rows A-B) and the Loggia Baseliana pavilion (Fig. 15, row C). Each image was first translated with image-to-image in Stable Diffusion, then lightly extended and clarified using Photoshop Generative Fill (Fig. 15, columns 2-3). I extracted the most resonant components from these resulting images: the sun-lit decking and vegetated rafters (Fig. 15, 3C), the emphatic axial roof (Fig. 15, 3A), and the disciplined colonnade rhythm (Fig. 15, 3B). These fragments were composed into a new canvas (Fig. 16, top), intentionally juxtaposing roof, deck and column sequences to heighten the bodily cadence central to the promenade architectural. A brief, final Image-to-Image pass homogenised and produced the final image, a coherent timber walkway that retains Tachon's underlying geometry, yet communicates a sensory narrative. (Fig. 16, bottom) The architecture must meet the child at their level, both physically and emotionally.

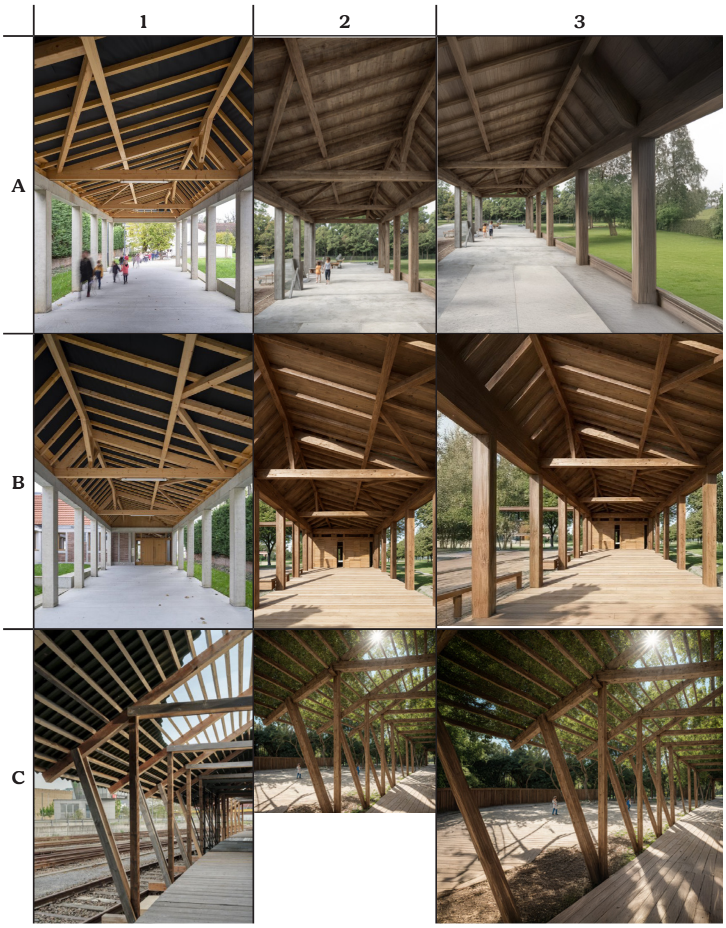


Figure 15. Process for Promenade fragment.

Figure 16. Collage (top) and Resulting image (bottom) for Promenade fragment.

Inside-Outside

To explore the second fragment, inside-outside, I began with an image of Marie-José Van Hee's own house in Ghent, in which the courtyard sequence embodies the quiet, rhythmical transition we seek for a special-needs primary school. A single reference photograph (Fig. 17, left) was transferred into Stable Diffusion's Image-to-Image function to generate a cluster of reinterpretations (Fig. 17, center). From these images, I selected one frame that best preserved the green courtyard, yet suggested a warmer, timber-based tectonic, ensuring the values of Van Hee's architecture were preserved in the image. The Retouche Technique then unfolded: the image was opened in Photoshop for surgical edits; black mullions were added to formalise glazing, a protective plinth was created to aid wheelchair navigation, and additional visual cues were provided for the visually impaired. Human figures were added to imply scale and programme. This file was returned to img2img for a brief consolidation pass, allowing the model to create a final, stylised image with a harmonised atmosphere. (Fig. 17, left)

1 2 3

Figure 17. Process for Inside-Outside fragment.

Transparency

To interrogate the third fragment, transparency, I abandoned photographic references and generated a set of scale-model images directly from text prompts. Using the ModelMaker-XL LoRA and a prompt derived from Pezo von Ellrichshausen's rigorous concrete language, I produced a large batch of hand-model images in Stable Diffusion's Text-to-Image. Applying the canvas-storm workflow, I sifted through the batch for what I saw as transparency. For example, in Figure 18, the concept of the design is communicated well. Stripped of detail, the grey plaster volumes illustrate geometry and rhythm, echoing the pedagogical models once favoured by Petillon Ceuppens: clarity through reduction.

This process reinforces the fragment's premise: transparency emerges when the form is immediately readable, whether cast in monochrome plaster or rendered through a different medium. It invites users to navigate through calm spatial cues rather than overt spectacle.

1	2	3

Figure 18. Process for Transparency fragment.

4.3 Analysis

4.3.1 Annotation

To illustrate the analytical process in this step, we present a concrete example (Fig. 19) of how we engage with a selected image through annotation. This step-by-step approach demonstrates how we interpret visual material, enrich it through critical questioning, and ultimately begin to speculate on its architectural potential.
1. Selecting the Image: This image was chosen for its quality, which presents an outdoor space surrounded by parts of a building. IMG This can be interpreted as a courtyard, but its ambiguity allows for greater freedom in interpreting the image.
2. Annotating with Depth: Using sketches, diagrams and short notes, we analyse the image. We raise questions about the human scale in this space; we seek connections between disparate areas within the image. We look for elements that stand out in the image.
3. Intervening Through Design: We speculate what the dimensions should be. We define the width and depth of this courtyard, as well as the ceiling height or roof structure. An additional step could be to consider how this space is connected to the surrounding spaces.

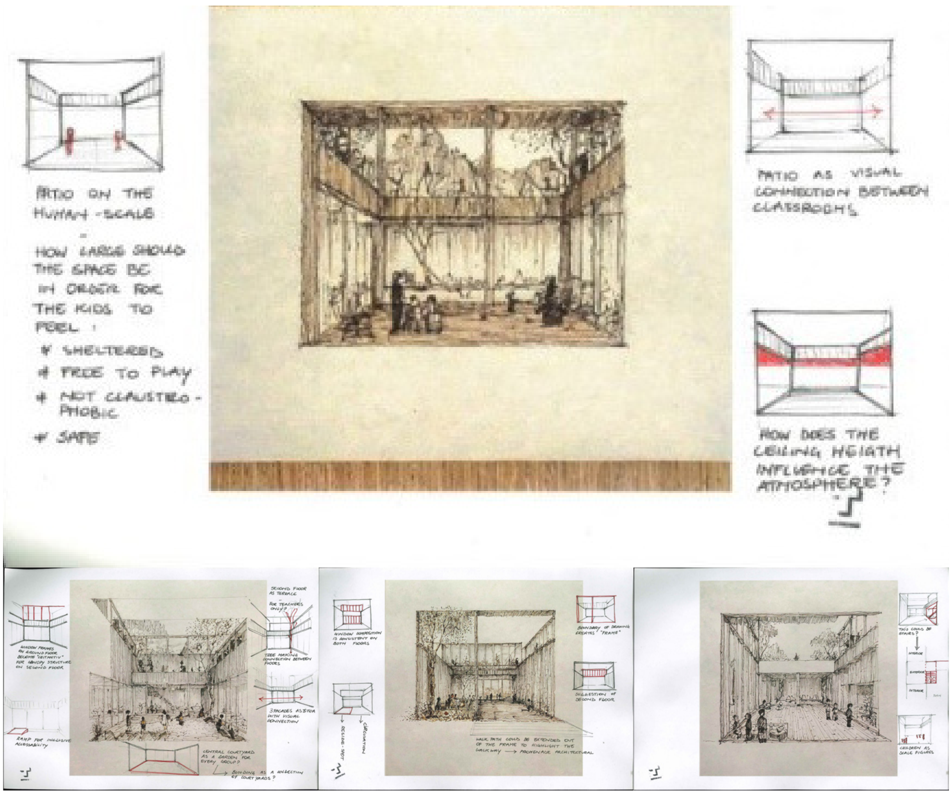


Figure 19. Example of annotation on a generated image.

Figure 20. Additional example of annotation on a generated image.

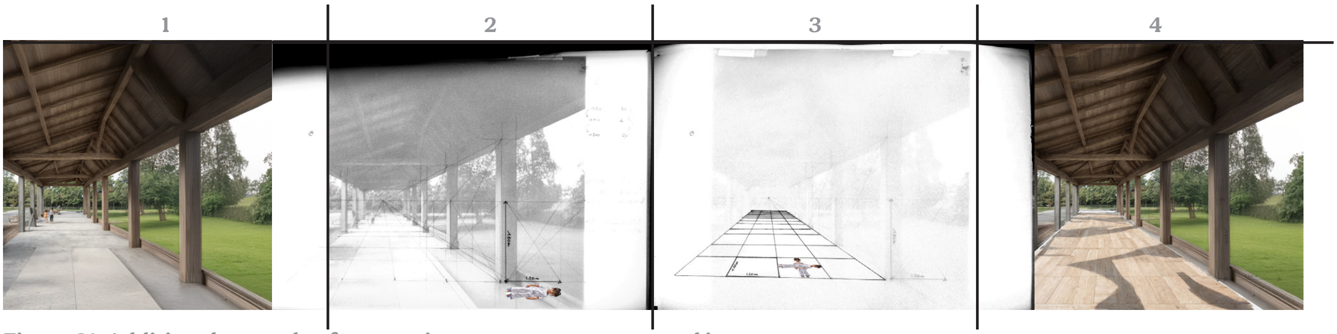


Figure 21. Additional example of annotation process on a generated image.

4.3.2 Fiches

To exemplify this method of visual analysis, we include the following fiche (Figure 22, left) as a representative case within the broader survey. This annotated example demonstrates how respondents critically engage with generated images, not only by observing formal elements such as repetition, scale, and materiality, but also by speculating on the underlying spatial logic and architectural potential.

The architectural design in the image suggests a sense of continuity primarily through the repetition of window frames and materials. However, these alone offer only a limited understanding of spatial progression. Additional design cues, such as the external 'staircase', inspire further speculation, hinting at the potential to create a multi-level outdoor space. In this fiche, the annotator proposed the fusion of the two roof volumes, reinforcing the idea of architectural coherence and prompting a reconsideration of how the generated space might evolve in built reality.

We are told the image appears to depict a school-like programme, inferred from contextual clues including the presence of children, visible backpacks and the overall enclosed, protective nature of the courtyard. This programmatic reading aligns with our notion of a safe, inward-facing environment tailored for young users. Architecturally, the arcade serves as a significant element, clearly marking the threshold between the interior and exterior and emphasising the transitional quality of this space.

Yet, despite the overall clarity of spatial layout, the image demonstrates some shortcomings in terms of human scale. The figures' eye levels are aligned unnaturally, which disrupts the perceived depth and flattens the scene. Moreover, the scale and placement of the furniture elements do not contribute effectively to understanding spatial proportions. In terms of materiality, while the repetition of columns successfully introduces a visual rhythm and sense of depth, the exact nature of certain materials remains unclear, which reduces their communicative impact; however, this also allows for more freedom in interpretation.

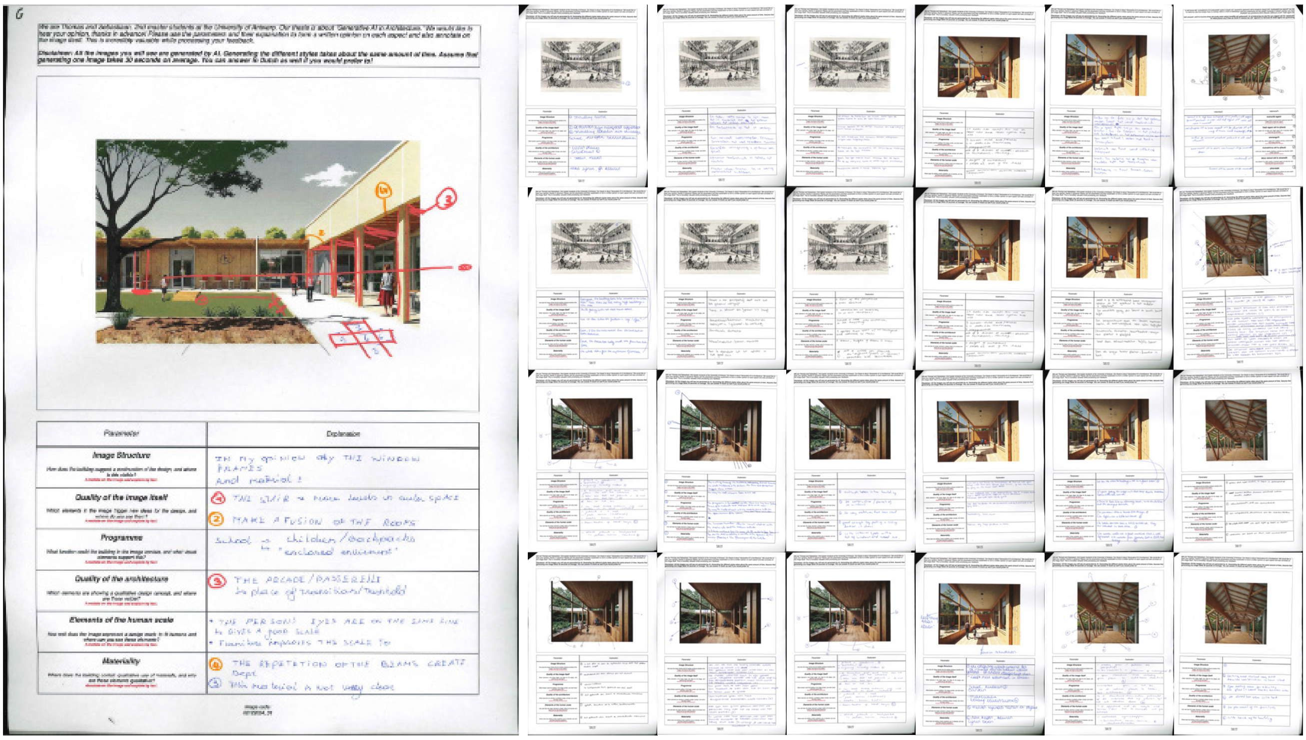


Figure 22. External annotation process on a generated image through fiches.

4.4 Takeaways

4.4.1 Place in process

We observed that AI-generated images in architecture can play two leading roles: as final visuals and as tools within the design process.

As final products, these images are beneficial in early design phases or competition entries, where atmosphere and concept take precedence over technical detail. Tools like Stable Diffusion can quickly create photorealistic or stylistically rich visuals that communicate the core idea with emotional impact.

At the same time, these images can be used as part of an iterative design process. It helps generate spatial ideas, test volumes or explore variations. These outputs are then evaluated, edited and refined by the designer. In this way, diffusion models become a creative assistant, not a replacement. The designer remains in control, utilising AI to expedite exploration while preserving authorship and design intent. This approach supports faster iteration without losing direction, as the designer remains responsible for analysing, interpreting, and annotating these images.

This analysis led to a desire to align the visual tone with the corresponding design phase. This is supported by the results of a survey we conducted among students and professionals. Respondents consistently favoured more expressive, sketch-like imagery during early ideation, while photorealistic renderings were seen as most effective for communicating final proposals. This reinforces the need for stylistic differentiation (Fig. 23) across design phases, validating the use of multiple LoRA styles in Stable Diffusion to match evolving communicative intentions.

Figure 23. Stylistic differentiation on the same base image.

Table 2 summarises the Phase Styling workflow, developed in response to the evolving representational needs across different stages of the design process. By applying distinct LoRA styles in Stable Diffusion, we were able to tailor image aesthetics to the specific phase of design, whether rough sketches in early ideation, illustrative concept visuals, or more realistic renderings for proposals. This short-duration workflow operates at a 10-90 man-machine ratio, relying heavily on pretrained visual presets or custom-trained LoRA models. As the table indicates, its effectiveness depends on having stylistic presets already available or investing time in their development beforehand. By aligning the visual tone with the design phase, the designer maintains familiarity with a specific aesthetic throughout each step of the process.

	Phase Styling
Software	Stable Diffusion
Duration	Short
Use Case	Converting to sketches in sketch phase, illustrations for concepts or storytelling, realistic images for proposals
Input Type	Images to be styleswapped
Man-Machine Ratio	10-90
Limitations	Dependent on available presets or LoRA, preset development takes time

Table 2. Analysis of the Phase Styling workflow.

4.4.2 Prompt Engineering

Prompt engineering plays a crucial role in the generative design process, as carefully selected keywords, clear instructions, and relevant stylistic references help guide both the aesthetic and conceptual direction of the outputs. In addition, the use of tool-specific settings and parameters, such as adjustable controls, predefined presets or more advanced tools like ControlNet, allows for further refinement. These elements add structure and consistency to the results, enabling designers to steer the generation process with greater precision and coherence.

Translating architectural observations into a promptable vocabulary requires a methodical reduction of spatial complexity into a distilled yet meaningful set of terms. Rather than merely listing architectural elements, this method aims to capture the underlying intentions and effects of a design, making them transferable across contexts. For example, we analyse a project that employs deliberate material transitions to structure movement or indicate boundaries between spaces. Instead of describing the materials directly, the prompt encapsulates the spatial strategy, such as juxtaposition, that gives the architecture its identity. Similarly, qualities like layered spaces or natural light are not extracted as superficial descriptors, but rather as condensed interpretations of how users experience depth, transparency, or rhythm in a sequence of spaces.

This analytical approach enables a conceptual rearticulation of architectural ideas, which is crucial when working with generative models. The act of prompting is not neutral; it requires the designer to take a position, to decide which qualities are worth amplifying and which nuances are fundamental to the spatial experience. By translating observations into carefully chosen terms, the designer builds a bridge between the complexity of architecture and the abstraction of language-based generation, ensuring that the output remains grounded in architectural intent, rather than generic form.

A key observation concerns the importance of input quality. The conditions required for meaningful generation, such as contextual awareness, spatial logic and narrative depth, are typically the result of a thorough creative process. In that sense, diffusion models appear more capable of reproducing or reinterpreting existing design work than of initiating it. It is essential to note that while referencing established designers in generative prompts may seem like an effective strategy to evoke a particular architectural language, it often yields outputs that are superficial or reductive. Rather than capturing the depth and nuance of a designer's oeuvre, the generated images tend to reproduce a watered-down version of their most iconic and widely recognised features. This was particularly evident when we were prompting with names such as Marie-José Van Hee (Fig. 24) or Frank Lloyd Wright, which frequently led to generic interpretations stripped of the contextual sensitivity or spatial quality that defines their work. The outcome is typically a bland and predictable aesthetic that fails to reflect the complexity, materiality and atmosphere embedded in the original designs. This highlights the need for caution when referencing historical or contemporary figures too directly in AI workflows, as it can limit creative potential and reinforce surface-level imitations rather than fostering genuine architectural exploration.

Figure 24. Bias in image generation through prompting. Left: Generated image, Right: Source Image.

4.4.3 Collage Technique

The collage workflow consolidates multiple outputs into a single, architect-curated vision through 5 successive operations (See Fig. 25 for 3 examples with each a different amount of steps, but all with the same operations as listed below).

Input image

The process starts either with generated AI imagery or with a screenshot of a 3D model. This primary frame anchors perspective and geometry.

Iterative Image-to-Image

The base image is passed through Stable Diffusion's Image-to-Image several times, with each run focusing on a specific variable, such as materiality, lighting, or programme. The objective is not a finished render but a set of partial successes; each variant contains at least one convincing element.

Selection

The resulting batch is evaluated on architectural and aesthetic qualities. Specific elements are selected and extracted in Photoshop while the remainder of each frame is discarded.

Montage

Selected fragments are combined in Photoshop. At this stage, the designer may optionally perform micro-adjustments to achieve spatial coherence, such as masking seams, rescaling figures, or nudging horizon lines.

Generative consolidation

The manual collage is used as input for Image-to-Image with a short, harmonising prompt, allowing the model to fuse texture, colours and geometry across the stitched boundaries.

Figure 25. Examples of Collage Technique.

Across all three illustrated cases, the collage method reduced exhaustive full-scene re-prompting to targeted combinatory curation, accelerating the exploration of material strategies and atmospheres while preserving the geometric integrity of the underlying design to a reasonable extent.
The Collage Technique workflow, summarised in Table 3, is a process in which selected visual qualities from multiple generated images, or actual built references, are composited into a single, cohesive frame. This medium-duration workflow blends architectural authorship with targeted generative operations, operating at a 70-30 man-machine ratio. Interesting fragments are extracted from multiple images and assembled in Photoshop. A final harmonisation in Stable Diffusion using light prompting ensures textural and chromatic consistency across composited elements.

	Collage Technique
Software	Stable Diffusion, Photoshop Generative Fill
Duration	Medium
Use Case	Collaging qualitative parts of images and reprocessing them into a singular image
Input Type	Generated images
Man-Machine Ratio	70-30
Limitations	Needs images with qualitative elements to be extracted

Table 3. Analysis of Collage Technique workflow.

4.4.4 Canvas Storm Technique

The Canvas-Storm workflow is a method that exploits the speed of text-to-image diffusion to build a library of images before any commitment is made. Its operation unfolds in three deliberate stages.

Prompt formulation
A concise, but atmospherically charged text prompt is drafted. In this example, the prompt describes "a small building of which a horizontal window reveals a park area," while also specifying raw natural materials and calling for a hand-drawn aesthetic. The wording is calibrated for Stable Diffusion paired with the Handsketcher-XL LoRA, whose training yields sepia-toned, analogue line work.

High-volume generation
The prompt is executed once, but with the batch parameter set to produce forty images (Fig. 26). This scale balances variation against run-time, creating a matrix of micro-iterations. We observe shifts in framing, sectional cuts, and lighting, all rendered in under a minute on consumer hardware.

Rapid intuitive selection
Without over-analysis, the designer isolates a handful of images that best align with their expectations or allow for serendipity, which can yield unexpected, yet qualitative results.

The selected tiles can be optionally used for subsequent annotation and possible montage, becoming a ready reservoir for workflows such as Collage Technique or Retouche Technique. In contrast to image-specific refinement techniques, Canvas-storm deliberately postpones judgement, relying on volume first, evaluation later. The outcome is a curated set of ideas that are clear enough to spark conversation and abstract enough to avoid premature fixation.

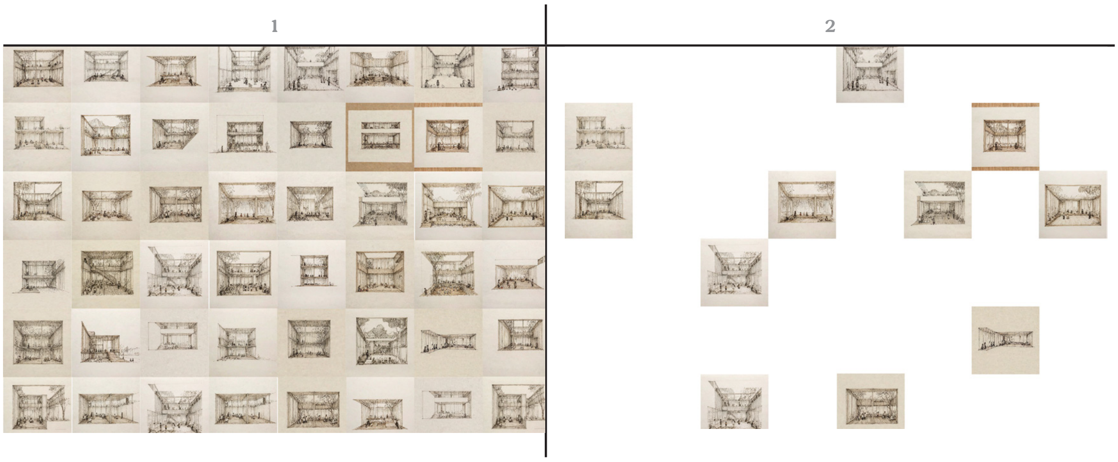


Figure 26. Example of Canvas Storm workflow,

The Canvas Storm workflow, outlined in Table 4, describes an approach to image generation that emphasises exploratory volume over immediate refinement. This medium-duration workflow allows for rapid creation and intuitive curation of image sets. At a 50-50 man-machine ratio, it begins with the careful formulation of a prompt and proceeds to generate large batches of images in a single cycle. These selections serve either as stand-alone stimuli or as raw material for subsequent workflows.

	Canvas Storm
Software	Stable Diffusion
Duration	Short
Use Case	Exploring different options and curating specific desired images
Input Type	Textual prompts, images
Man-Machine Ratio	50-50
Limitations	Technical knowledge is required to produce at least some desired results within a larger batch

Table 4. Analysis of Canvas Storm workflow.

4.4.5 Retouche Technique

In this workflow, Photoshop is more than just a post-production tool; it becomes a curatorial step for editing and refining generated images. Instead of accepting the AI output as final, the designer actively reworks the image to better align with their design intent. The original annotation of images might influence this intent, but elements that don't suit this intent can be refined within this workflow. The integrated Generative Fill feature in Photoshop is also a valuable tool for quickly generating these changes to the original images. The process (Fig. 27) includes four key steps:

1. Retouching for clarity
Unwanted elements or unclear details from the AI output, such as misplaced furniture, are removed or corrected. For example, benches might be taken out to create a more neutral and readable space.

2. Architectural detailing
Missing or unclear features, such as window frames or geometry, are added to enhance spatial logic and support the architectural narrative.

3. Adding human figures
Figures are introduced to suggest scale, use and atmosphere. These additions help bring the scene to life and clarify how the space is meant to be experienced.

4. Reprocessing in Stable Diffusion
The edited image is sent back through Stable Diffusion for harmonisation. This step brings everything together visually; here, we reimplement uniform textures, lighting and overall composition while keeping the designer's changes intact.

Figure 27. Process of Retouche Technique workflow.

The virtue of the Retouche technique is versatility. A nearly identical process (Fig. 28) was performed on a Rhino capture during a PAAcademy workshop with Cas Esbach IMG. A rough model was refined to presentational quality in under twenty minutes by alternating Photoshop precision with diffusion-based tone setting. The Retouche Technique focuses on one promising image and incrementally perfects it, allowing the architect to steer specific elements, such as window compositions, plinth height, or human scale. It is essential to note that the two-day PAAcademy workshop's live exercise concluded after the initial transformation of a Rhino screenshot into the first img2img output. Subsequent refinements, namely, the Photoshop interventions and the return pass through Stable Diffusion, were executed independently afterwards, once it became clear that the workshop image did not yet meet the desired architectural and atmospheric expectations.

Figure 28. Process of Retouche Technique workflow based on PAAcademy input.

To further evaluate the Retouche Technique, we began with a rudimentary SketchUp scene (Fig. 29). The view was exported as a flat image and used as input in Stable Diffusion's img2img four times, each run driven by the same descriptive prompt and defined geometry, but with one material change each: terracotta, brick, timber, and in-situ concrete. Out of these four outputs, the strongest elements were extracted. These fragments were composited in Photoshop to form a single, hybrid image, with some retouched elements. After this step, a brief consolidation through img2img aligned all present elements. This exercise confirmed the additional value of this workflow as a swift test for materiality.

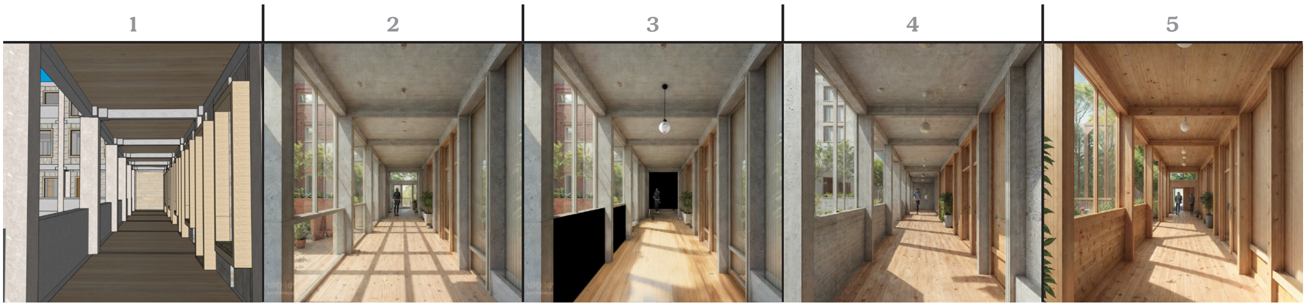


Figure 29. Process of Retouche Technique workflow. (Source: Gauthier Meyers)

The Retouche Technique workflow, outlined in Table 5, describes a medium-duration process in which generated images are refined to enhance spatial clarity and conceptual alignment during exploratory design phases. Rather than accepting AI outputs as final, this approach treats Photoshop, particularly its Generative Fill feature, as an active design tool, enabling targeted interventions that reinforce the architectural narrative. While operating at a 50-50 man-machine ratio, these iterative modifications allow the designer to filter, augment, and synthesise imagery in a way that supports communication without undermining authorship. While this workflow introduces an additional step in the design process, it ultimately strengthens the role of the image as a tool for design reasoning rather than a fixed product.

	Retouche Technique
Software	Photoshop Generative Fill, Stable Diffusion
Duration	Medium
Use Case	Refining specific elements in exploratory phases to communicate ideas or concepts
Input Type	Generated images
Man-Machine Ratio	50-50
Limitations	Additional step in the design process

Table 5. Analysis of Retouche Technique workflow.

4.4.6 Reference library

The extensive collection of images generated during the exploratory phase serves as a valuable reference library for future design development. Since all outputs were produced with the typology of a primary school in mind, the 'dataset' is already embedded with relevant spatial, functional and atmospheric considerations. Rather than being discarded as mere by-products of experimentation, these images become a curated visual archive, offering a wide range of interpretations and stylistic directions that can be revisited and repurposed throughout the design process. This allows us to systematically categorise and evaluate the generated material, identifying recurring spatial logics, architectural expressions or compositional strategies that align with our evolving design intent. By selecting and annotating key images, the library becomes an active design tool rather than a static collection, enabling us to draw from previous explorations in a targeted and informed manner. In doing so, the generative process is not limited to the moment of creation but continues to inform decision-making as the project progresses.

4.4.7 Finalised Research Question

After this thorough exploration, we can refine our initial research question to one that aligns more closely with the actual scope of this research. In the earlier phases, we identified opportunities and limitations with these diffusion models, which now shape our approach to these tools. This hybrid man-machine process has an impact on the way we develop and explore new ways of designing to implement and test the capabilities of these diffusion-powered workflows. As these tools have both an influence on the design process and the final product, the research question should consist of two main parts as well: the process and the evaluation. With this acquired information, we modify the research question:

What are diffusion model-powered workflows that can reshape architectural design processes, and what is their added value?

5. Design

5.1 Het Mispeltje

5.1.1 Project Brief

This project involves the design of a new school facility for type 2 special needs primary education in Sint-Niklaas, accommodating approximately 50 pupils with intellectual or physical disabilities. The location on Mispelstraat 45 currently comprises a disused neighbourhood school and childcare centre. The project allows for the possibility of demolition and redevelopment to address specific functional and accessibility-related demands, which Team Vlaams Bouwmeester[12] even recommended that we assume in our design. A nearby city-owned chapel, presently without a defined use, may be considered for integration following consultation. Site development prioritises the preservation and activation of green outdoor spaces, which serve both recreational and educational purposes. Full accessibility for wheelchair users and smooth circulation for school transport is crucial, as well as clearly indicated and safe drop-off areas. The educational environment must be adaptable, enabling a range of pedagogical approaches while encouraging autonomy and socio-emotional development. Generous and secure communal areas are essential, alongside designated zones for support staff, including therapists and administrative personnel. The architectural expression should create a domestic and inviting atmosphere, achieved through the careful selection of warm materials and colours. A coherent and legible spatial layout is crucial for enhancing user comfort. The entire educational program must be situated at ground level, with auxiliary functions preferably located on the same floor to ensure inclusive access. Sustainability objectives must be embedded throughout, aiming for low environmental impact through energy-efficient and ecologically responsible construction practices.

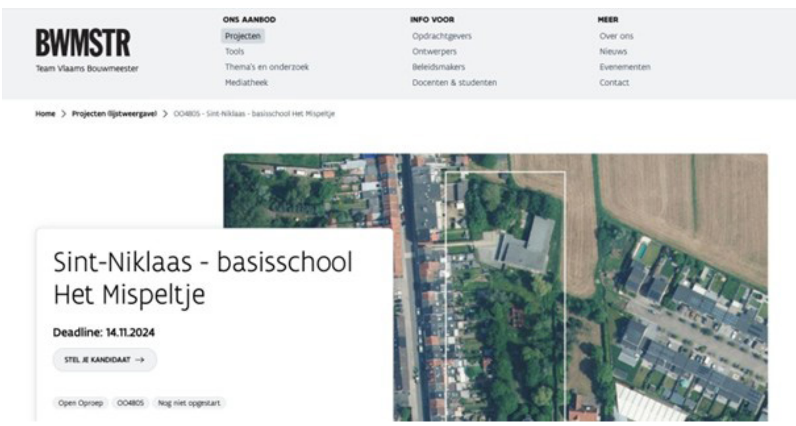


Figure 30. Open Call competition brief.

12 https://www.vlaamsbouwmeester.be/nl/projecten/oo4805-sint-niklaas-basisschool-het-mispeltje

5.2 Design process

This scheme (Fig. 31) outlines the steps of this hybrid design process. The legend clarifies the typology of each step, highlighting the interaction between human input, AI contribution, and hybridized workflows. Additionally, it shows how each step relates to the others.

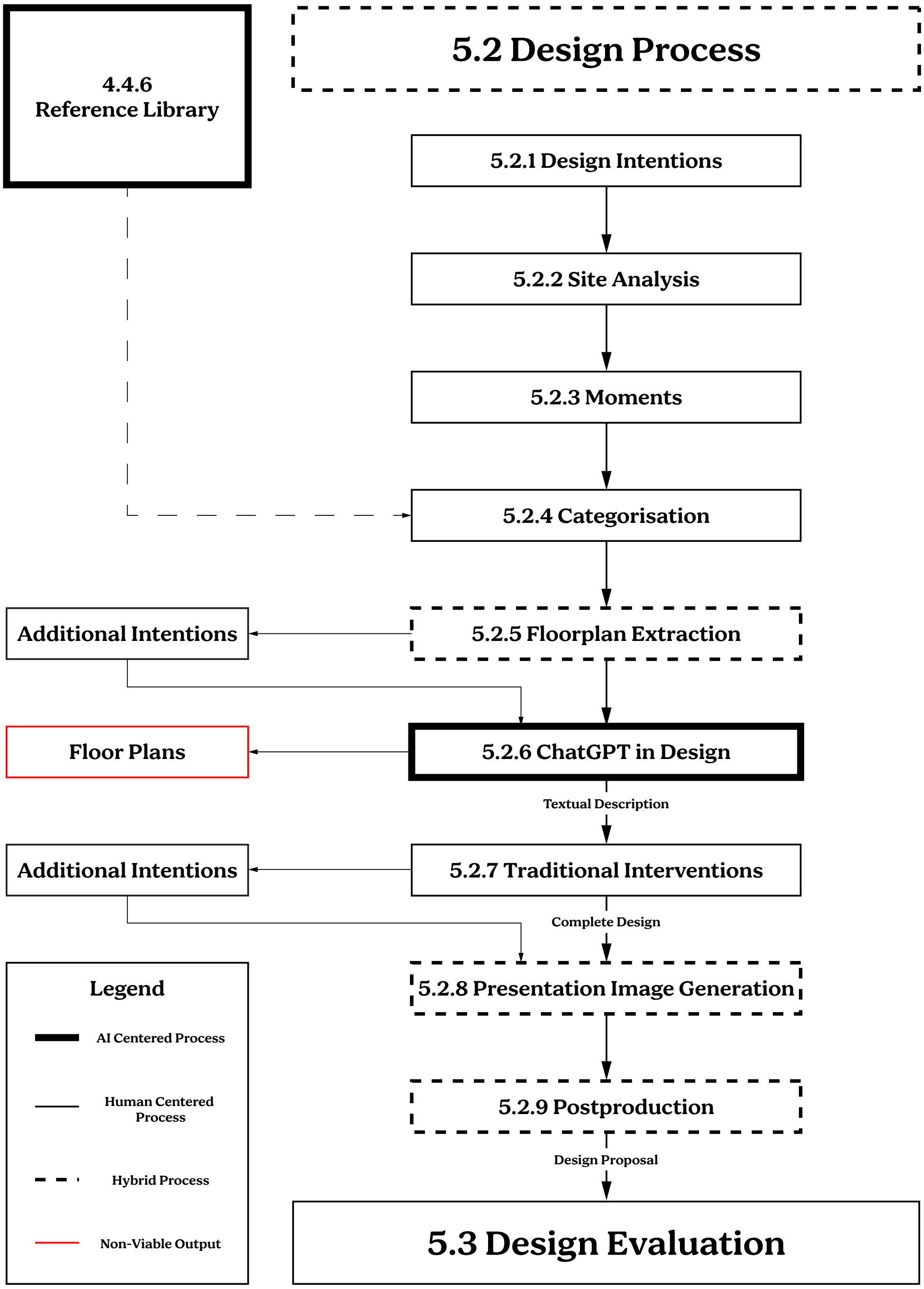


Figure 31. Scheme of hybrid design process.

5.2.1 Design Intentions

Based on the project brief outlined in the previous section 5.1, a series of design intentions were distilled to provide a clear framework for the architectural response. These intentions synthesise the core functional, spatial, and experiential ambitions of the client, users and us, as designers, into a structured overview. Table 6 categorises them into three thematic groups, surroundings, quantitative parameters, and qualitative parameters, offering a concise yet comprehensive summary of the priorities that guided the subsequent design process.

A	Surroundings
A1	Thoughtful integration of a new access road, ensuring smooth flow and a safe drop-off zone
A2	Possible integration of the adjacent city-owned chapel
A3	Continued availability of the site as a public play area outside school hours
A4	Inclusion of outdoor green spaces and playgrounds that support both relaxation and education
B.	Quantitative Parameters of the Building
B1	All need-to-have programme spaces are included in the design.
B2	All nice-to-have programme spaces are included in the design.
B3	Delivery of the project within a budget of € 2,880,000 (excl. VAT and fees)
B4	Entire learning environment on ground floor, support functions on +1
B5	Integration of generous play and sports facilities to foster interaction and collaboration
B6	Aim for a minimal ecological footprint through sustainable and energy-efficient design.
C.	Qualitative Parameters of the building
C1	Flexibility to accommodate diverse pedagogical methods
C2	Architectural expression aligned with the values of care, inclusivity and educational excellence
C3	Use of warm, inviting materials and colours to create a homely and secure atmosphere
C4	Design relates to the human scale

Table 6. Design intentions derived from design brief.

Disclaimer: This list is subject to change and evolution throughout the design process.

In addition to the detailed spatial program, we also received a provisional programme of requirements (POR) from Team Vlaams Bouwmeester. This document was explicitly identified as incomplete during the first contact with Hedwig of Team Vlaams Bouwmeester. As such, some aspects of the POR were either missing or incorrect. We acknowledged this limitation, and our design decisions were made independently of the inconsistencies in this document.

To establish a coherent spatial configuration, we conducted a comparative analysis of six scenarios, each representing a different combination of programmatic elements and spatial ambitions. The scenarios ranged from a minimal footprint, based solely on essential requirements, to a more expansive layout that incorporated additional, non-essential functions. This approach allowed us to critically evaluate the balance between 'need-to-haves' and 'nice-to-haves', aiming for a pragmatic and spatially efficient outcome.

Function	Units	Surface area per unit	Surface area per function
NEED TO HAVE			
Classrooms with integrated sanitary facilities (washbasin, toilet, changing table) and a First Aid room	7	72	504
Sensory room	1	36	36
Seclusion room	1	8	8
Physiotherapy room	1	35	35
Occupational therapy room	1	25	25
Speech and language therapy room	1	20	20
Teachers' lounge	1	40	40
Office for the school management and orthopedagogical support	1	25	25
Administrative office	1	25	25
Staff kitchen	1	20	20
Shower facility (within staff toilet area)	1	6	6
Staff restroom	1	8	8
Maintenance room	1	9	9
Multipurpose room (for movement education, various group activities, etc.)	1	100	100
NICE TO HAVE			
Covered outdoor playground	1	100	100
Additional therapy room (e.g., yoga, music therapy)	1	25	25
Bicycle parking for staff and visitors (20 spaces)	1	60	60
Total surface area			1046

Table 7. Programme Of Requirements.

The selected scenario proposes the entire learning environment at ground level, ensuring accessibility and spatial coherence, while all supporting functions for teachers are situated on the upper floor. In addition, the adjacent chapel, currently without a defined use, would accommodate the extra therapy space and the seclusion room. This not only reduced the volume of new construction but also offered the opportunity to activate the chapel in a meaningful way within the project context.

5.2.2 Site Analysis

Upon visiting the site for the first time, we were struck by a spatial quality that exceeded our expectations. The terrain between the Mispelstraat and the Kleibeekstraat in Sint-Niklaas presents a surprisingly layered composition: long and relatively narrow, yet generous in its spatial sequencing and potential. The plot includes two distinct access points. The first, a narrow path, leads east through a passage between two existing houses on the Mispelstraat. This entrance, associated with the former childcare centre's official address, feels secondary in both spatial and functional terms. In contrast, the second access from the Kleibeekstraat is more expansive. It runs north alongside the full length of the site, guiding visitors and future users past its key components, towards the existing building at the rear. Given its legibility and potential for spatial hierarchy, this avenue now assumes the role of principal access.

The site unfolds sequentially from this entrance, beginning with a striking architectural element: a small chapel located directly along the street. Though currently unused, its presence anchors the site with a formal and symbolic significance. Its proportions and detailing are modest yet refined, reminiscent of classical precedents, albeit in a restrained and vernacular way. Beyond the chapel lies an open grassy area, formerly adapted as a small football pitch, followed by a compact playground. Further along, the landscape transitions into a more enclosed zone, where a small creek intersects the site. A wooden bridge facilitates passage to the final portion of the terrain, where the old building is located. Notably, a remnant fence indicates the perimeter of the former childcare facility. Throughout the site, a consistent border of medium-sized trees frames the central open space, enhancing the longitudinal experience and reinforcing the site's natural qualities. The presence of the creek and the gradual buildup from open field to enclosed architecture establish a spatial rhythm that is both intuitive and layered. This subtle sequencing, combined with the generosity of the green surroundings, informed our design strategy.

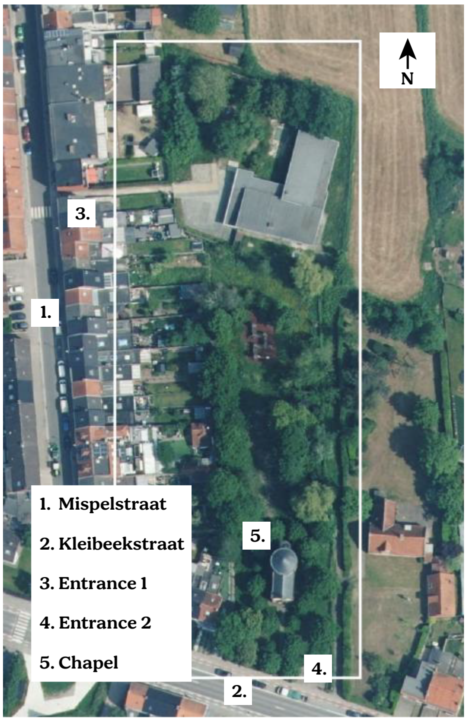


Figure 32. Site analysis.

Although the existing childcare centre will be demolished, the adjacent chapel offered a compelling architectural reference. Its calm proportions and careful siting provided a sense of orientation and continuity that we sought to preserve and reinterpret. By abstracting its geometry, a single circle and a rectangle, we introduced a conceptual gesture that acknowledges the chapel's role within the site. While the verticality and function of the chapel could not be replicated due to the ground-floor spatial logic required for special education, its influence remains embedded in the design's spatial language and organisation.

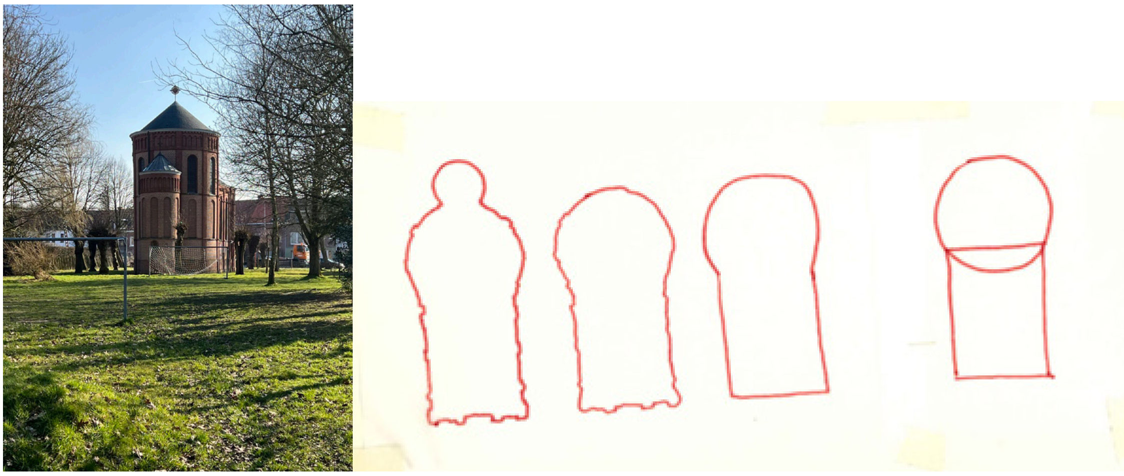


Figure 33. Chapel on the site and sketch of its geometric simplification.

Neighbourhood Survey

To better understand the local context and ensure that the architectural proposal responds meaningfully to its surroundings, we conducted a small-scale neighbourhood survey. We chose to perform this in a traditional method because the information that, for example, ChatGPT provides is not necessarily accurate or nuanced enough. The objective of this survey was to gain insight into how the site is currently perceived and used by nearby residents and to identify recurring concerns or expectations that could inform and enrich the design process. This form of early engagement allowed us to move beyond abstract site analysis and begin translating local sensitivities into tangible design considerations.

The survey, distributed among residents living near the project site, yielded responses from eleven different households. While the number of participants was relatively modest, the diversity and specificity of the responses provided valuable insights. One of the main findings was the fragmented relationship between the neighbourhood and the site. Some residents reported regular use of the area, particularly the playground and the green space near the chapel. Others described their interaction with the site as rare or non-existent, often due to perceived inaccessibility or a lack of spatial quality.
The responses regarding the school's future development varied significantly, but they effectively reinforced the intentions developed in the previous step. Some residents expressed uncertainty or indifference, while others articulated a clear preference for a cheerful and welcoming building suited to the needs of children with disabilities. Several respondents, however, opposed any form of new construction on the site, proposing instead the creation of a forested area or a communal parking lot. Overall, there was a preference for a modest, safe and context-sensitive design rather than one that feels oversized or overly institutional.

In terms of suggestions for the reuse of the chapel and the adjacent park, responses reflected a wide range of ideas. These included restoring the chapel for quiet reflection or community use, introducing infrastructure such as a skate ramp, planting flowers to attract pollinators and even establishing a pop-up space for young artists or small entrepreneurs. Conversely, a few participants proposed demolishing the chapel to create a more open and accessible green space, potentially supplemented by the neighbourhood parking. One respondent emphasised the importance of redirecting school access away from the Mispelstraat, citing existing traffic issues.

Concerns related to mobility, safety and traffic were among the most frequently raised issues. Many residents described the current street conditions as problematic, particularly during peak hours and feared that the new school could worsen this situation. Comments referenced difficulties with navigating the street due to parked cars, the general lack of space and past instances of vehicle damage. While several respondents expressed support for the educational function of the project and acknowledged the importance of inclusive facilities, this support was often accompanied by concerns about the impact on the daily lives of residents.

The Communication Styling workflow, summarised in Table 8, emerged from our intention to tailor visual outputs to different stakeholders involved in the design process. Whether addressing clients, neighbours, or municipal officials, each group responds to distinct visual languages and degrees of abstraction. By leveraging Stable Diffusion's LoRA-based styling, we were able to adapt the format of our illustrations, from loose conceptual sketches to photorealistic proposals, depending on the communicative context. This short-duration workflow is characterised by a high machine involvement compared to the designer's input in a 90-10 ratio, where images are "styleswapped" using predefined or custom-trained stylistic presets. Its effectiveness depends mainly on the availability and quality of those presets.

	Communication Styling
Software	Stable Diffusion
Duration	Short
Use Case	Refining output before meeting with different stakeholders in a project, based on what material they would want to see
Input Type	Images to be styleswapped
Man-Machine Ratio	10-90
Limitations	Dependent on available presets or LoRA, preset development takes time

Table 8. Analysis of Communication Styling workflow.

5.2.3 Moments

Moment 1 - Arrival and Departure

Arriving at school is a daily ritual for children, parents, and caregivers. We identified this moment as critical in setting the tone for the day. Children arrive from various directions, so a clearly articulated and safe drop-off zone is essential. This space must facilitate a smooth and low-stress transition from home to school, offering comfort and reassurance to both students and their parents.

Moment 2 - First Impressions and Entry

The entry defines the first impression of the building. A calm and comprehensible approach fosters a sense of safety and a sense of belonging. The positioning of the entrance, its visual legibility and the way it guides people inward are all designed to provide clarity to those who enter, especially important in a school for children with specific needs·

Moment 3 - Play and Social Interaction

The polyvalent space acts as a threshold between structured education and informal interaction. We view this as a space for shared play and spontaneous social bonding, which are essential for emotional development and inclusion. The architecture must support flexibility, openness and a sense of communal presence. In parallel, spaces for therapy and support are carefully integrated and must also remain accessible and legible. Clear circulation patterns allow children and staff to navigate these environments with ease, promoting empathy and clarity throughout the building.

Moment 4 - Classroom

Though seemingly self-evident, the classroom remains one of the most sensitive and complex spatial moments. Here, safety, familiarity and focus come together. We consider the classroom not as an isolated container, but as a nuanced environment that supports varied learning needs while offering spatial clarity and emotional security. Its design contributes directly to the child's ability to thrive, both cognitively and socially.

Moment 5 - Outdoor Play and Safety

The final moment is the connection between inside and outside. The outdoor play area, situated between the school and the adjacent chapel, offers a protected space that allows children to experience nature and engage in movement within a secure setting. This space encourages physical activity and sensory engagement while remaining intimately tied to the architectural narrative of the project.

5.2.4 Categorisation

After defining the five architectural 'Moments', we revisited the expansive image library developed during the exploration phase and systematically categorised the images according to these moments (See Fig. 34). This additional rationalisation allowed us to assess the extent to which our initial, more intuitive visual output aligned with the conceptual framework we had since developed. By mapping our speculative imagery onto this new and structured sequence of experiences, we were able to critically reflect on the potential of each moment, not only in terms of spatial quality, but also in how effectively the images communicated atmosphere, function and emotional resonance.
This differentiation revealed both gaps and strengths in our process. In some cases, the previously generated visuals captured a moment with remarkable clarity. In other instances, certain moments were underrepresented or lacked the necessary spatial cues, which highlighted where further attention was needed. Importantly, this categorisation exercise also exposed the limitations of working purely with images, especially when dealing with human experiences such as movement, routine or care. Nevertheless, the process offered a valuable feedback loop between speculative image-making and architectural thinking, helping us to sharpen our focus for the subsequent design steps.

Figure 34. Selection of categorised images.

5.2.5 Floorplan Extraction

In this phase of the design process, evocative architectural images generated during the earlier explorations were systematically reinterpreted into floor plans through spatial reasoning and the disciplinary knowledge acquired in architectural education. Rather than relying on automated extraction methods, we deliberately opted for an interpretive drawing process. This is an act of architectural authorship, grounded in the designer's understanding of form, scale and programmatic logic. The generated visuals were not considered resolved designs, but rather atmospheric provocations from which narrative and typological ideas could emerge. Architectural clues, such as column rhythm, shadow depth, thresholds, and facade articulation, were visually analysed and translated into floor plan options. This translation allowed us to articulate spatial intentions, including figure-ground relationships, circulation flows, zoning and the interplay with natural light and landscape.

Using the previously generated imagery provided a productive starting point to imagine the potential configuration of the school. But while these images suggested spatial qualities and moods, they lacked the precision and coherence required for architectural planning. To bridge this gap, we incorporated the dimensional constraints and spatial requirements defined in the project brief. This process helped us ground the speculative content in architectural reality. The tendency to translate images into these spatial drawings is closely related to the way we are educated. Our education places strong emphasis on understanding the relationship between plans, sections, and other visual representations, which is a skill that influences our instinct to move beyond the original generated images.

As seen in Figure 35, one interpretation focused on the relationship between therapy rooms and an adjacent outdoor area, such as a patio or courtyard. The image offered an idea of connecting the circulation space directly to the landscape outside, creating a threshold between the indoor functions and the open outdoor space. Upon closer inspection, certain elements, such as door placements, proved to be unrealistic or non-functional. These were corrected in the floor plan, making the proposal more coherent and bringing it a step closer to becoming authentic architecture.

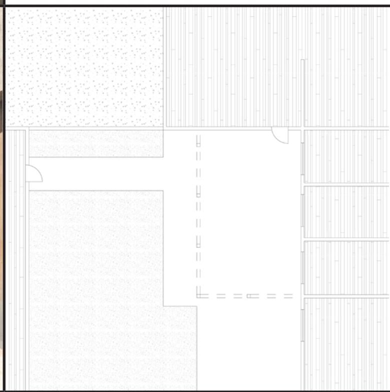
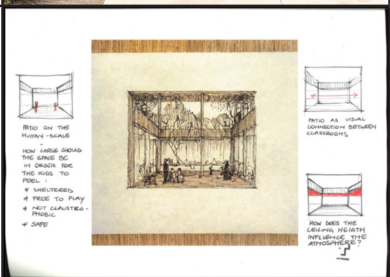
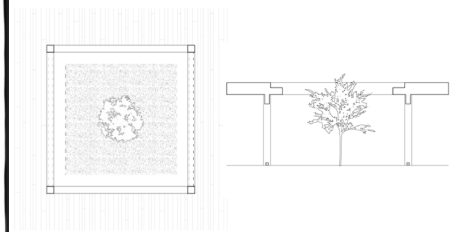
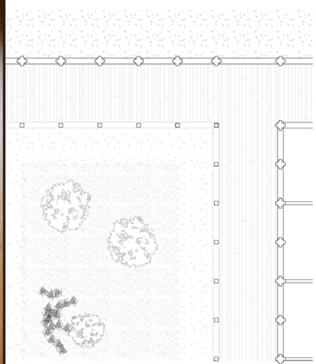
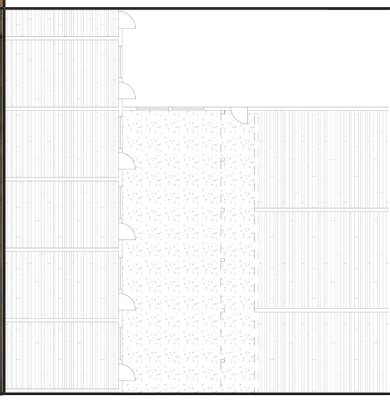
	AI-Image	Floorplan
M1		
M2		
M3		
M4		
M5		

Figure 35. Selection of extracted floor plans.

The Floorplan Extraction workflow, analysed in Table 9, summarises a key phase in our process, where generated architectural images were reapproached through a lens of spatial reasoning and architectural education. This medium-duration workflow relied on intentional prompting within Stable Diffusion, followed by interpretive analysis carried out by the designer. With a 70-30 man–machine ratio, the workflow required an experienced reviewer to extract meaningful architectural qualities from visual suggestions, such as column placement, façade logic, or thresholds. These were then anchored in real constraints from the project brief, allowing for the speculative imagery to evolve into grounded, usable floor plans.

	Floorplan Extraction
Software	Stable Diffusion
Duration	Medium
Use Case	Analysing qualities visible or suggested in a generated image and translating this into usable floor plans
Input Type	Prompting images with a clear goal
Man-Machine Ratio	70-30
Limitations	Needs an experienced and educated reviewer to interpret

Table 9. Analysis of the Floorplan Extraction workflow.

As a result of this step, additional design intentions were developed, as shown in Table 10.

B.	Quantitative Parameters of the Building
B7	Overhang/canopy creates a transition zone between inside and outside
B8	Entrance canopy stretches out like an open arm, generous yet unpretentious
B9	Low gable green roof, sloping northward to bring natural light into each space
B10	Central patio-courtyard that evokes light and nature into the building
B11	Floor plan organised on a clear 8×8m structural grid
B12	Classrooms can be opened up to form one larger classroom.
B13	Each classroom has direct access to the outdoors.
C.	Qualitative Parameters of the Building
C5	Strong legibility and visible boundaries for optimal orientation

Table 10. Additional intentions derived from Floorplan Extraction.

5.2.6 GPT in Design

On March 25th of 2025, a significant update to ChatGPT was released, marking a turning point in our process. Until this moment, our processes have primarily relied on image-based generation tools, such as Stable Diffusion. However, the new capabilities of ChatGPT prompted us to explore a different question: could we use a language-based model to generate a qualitative architectural floorplan?

To test this hypothesis, we conducted a focused experiment using ChatGPT as a generative design assistant. Our input consisted of two key elements: an overview of our hand-drawn floor plans, as outlined in Section 5.2.5, and a detailed text prompt derived from the programme of requirements. The objective was not to obtain a finished plan, but to examine whether the model could synthesise these inputs into a coherent and qualitative layout.

Initial attempts using only these two inputs proved insufficient. Without a defined building contour, the model generated generic floor plans that lacked spatial logic or appropriate scale (Fig. 36). This prompted us to introduce an additional constraint: a predefined outline derived from the existing chapel on site, a conceptual reference discussed earlier in section 5.2.2. By scaling this circular contour (Fig. 37) to a gross surface area of 1030 square meters, we ensured the input aligned with both programmatic and spatial requirements. These dimensions were based on our selection of the fourth scenario, which consisted of 934 square meters of net programme area, plus a ten per cent margin for circulation and structural elements. The contour was correctly oriented to the north on the site layout, and the existing trees were included to establish contextual awareness.

Figure 36. Selection of generated ChatGPT floor plans.

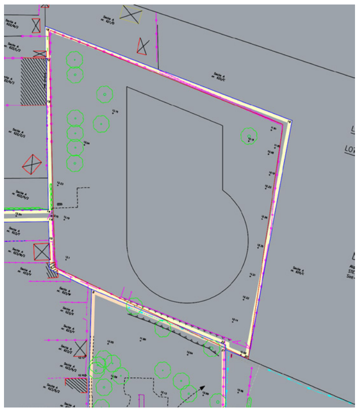


Figure 37. Contour input.

This added framework had a significant effect on the quality of the generative outputs. In several configurations, we observed clear echoes of our earlier design intentions: a central open space surrounded by classrooms, the introduction of a defined courtyard featuring transitional zones or even suggestions for an overhang functioning as a covered outdoor area. (Fig. 38) These proposals, while sometimes unintended, aligned well with our ambitions and demonstrated how generative tools can act as reflective design partners. The circular geometry, anchored in the chapel's footprint, not only grounded the model's spatial logic but also helped embed the building harmoniously into its surroundings. This shape provided the design with a defined perimeter and orientation, while still allowing for variation and exploration within the generative process.

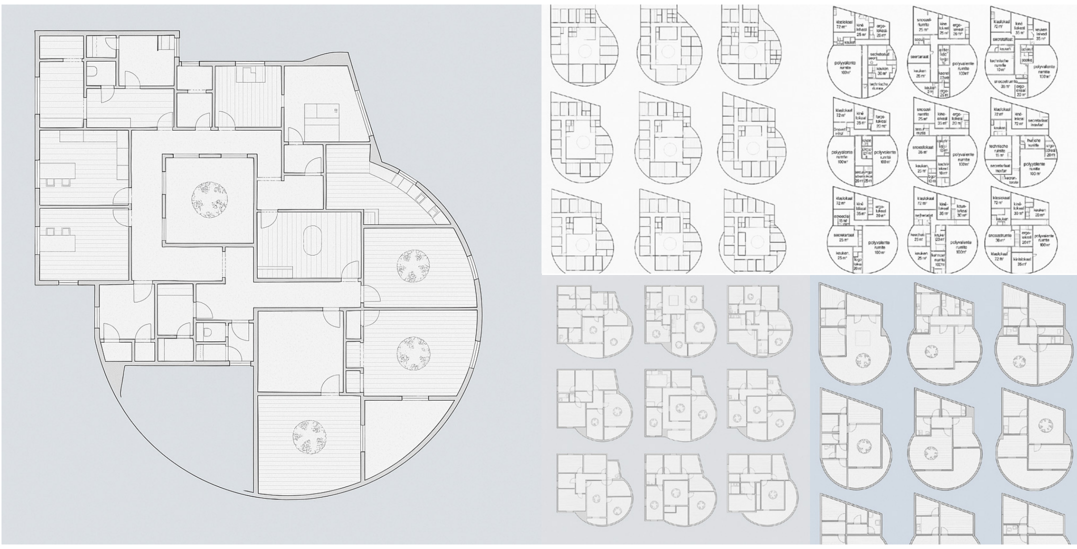


Figure 38. Generated ChatGPT floorplans with contour.

Despite these promising elements, the experiment also revealed apparent limitations. The most notable shortcoming was the complete lack of control over scale. While the spatial logic might have been conceptually sufficient, the practical application in terms of dimensions, proportions and usability was insufficient. As a result, we concluded that generating precise floor plans in this manner is neither efficient nor effective at this stage. Although the current results are limited, we remain convinced of the potential for this approach in the future. With continued development, tools like ChatGPT may become increasingly capable of contributing to early-stage design workflows, not by replacing architectural authorship, but by complementing it with structural suggestions, programmatic logic and creative provocations. Building on this reflection, we also question whether diffusion models are the most suitable tools for generating floor plans. Their underlying architecture lacks the control necessary to define dimensions, relationships or rules with precision. For tasks that require accuracy and spatial coherence, vector-based AI tools may hold greater potential.

To continue the exploration, we opted for a revised approach. Instead of asking GPT to produce a floor plan directly, we shifted focus towards obtaining schematic or textual descriptions. These included spatial orientations, circulation principles and zoning diagrams. This method allowed us to benefit from ChatGPT's narrative and organisational thinking, while maintaining control over the architectural translation ourselves. In this approach, we provided GPT-4o with the complete set of design intentions developed in section 5.2.1, along with the additions made up to 5.2.5. Based on this material, we asked for three descriptive scenarios that aligned with our intended architectural narrative. As illustrated in Figure 39, these scenarios directly connect to the key ideas we articulated in our “5 Moments.” It is important to note that ChatGPT was not consulted sporadically, but rather integrated continuously throughout the design process. This cumulative interaction enabled the model to develop a comprehensive contextual understanding of our project, encompassing orientation strategies, the positioning of existing trees on site, programmatic requirements, and the client’s expectations.

The Storytelling Technique workflow, shown in Table 11, outlines our use of ChatGPT to generate narrative spatial descriptions as a complementary design strategy. This medium-duration workflow, operating at a 70–30 man–machine ratio, enabled us to synthesise text, sketches, and contextual material into cohesive spatial narratives. It proved particularly useful in translating conceptual frameworks into communicable form, supporting both internal reflection and external client communication. However, the workflow is highly dependent on the quality and depth of the input. It also raises important considerations regarding intellectual property, especially when sensitive project data is shared across generative platforms. While our use of ChatGPT was structured and consistent, ensuring continuity and contextual growth over time, its application remained bounded by careful prompt engineering and critical post-evaluation.

Absolutely — here is a **very high-level conceptual description** of how children with special needs (type 2) and their teachers, therapists, and supervisors engage with the building. This focuses on the **emotional, sensory, and spatial interactions** that architecture can support, anchored in the real needs of users.

🧠 **A School That Understands Without Speaking**
This school does not expect its users to adapt to it. It **adapts to them** — silently, patiently, respectfully.
For type 2 students — children with cognitive and/or physical disabilities — every space they encounter speaks in a different language: not of signs or instructions, but of **form, light, material, and rhythm**.
The building helps them **anticipate what comes next**, reduces friction, and gently encourages independence within a framework of care.

🌀 **Arrival as Orientation**
As a child arrives at the entrance — located deliberately at the **curved southeast edge** of the building — they do not encounter a maze, but a moment of **orientation**.
- The entrance is **wide and clear**, with generous overhangs shielding from rain and sensory overload.
- Surfaces are **non-reflective** and tactile: soft paving underfoot, warm timber at hand height.
- No automatic doors suddenly hiss open. Instead, a slow, deliberate passage into a **buffer zone** allows time to adjust — neither outside nor inside yet.

From here, **the polyvalente ruimte** is visible — flooded with morning light — acting as the **visual anchor** for both children and adults. It is **the heart**, spatially and emotionally.

🛋️ **Calm, Structured, Predictable**
Children with special needs often thrive in spaces that are:
- **Logically sequenced**, with clear transitions and legible boundaries.
- **Predictable**, so that anxiety is reduced through consistent layouts.
- **Visually and acoustically calm**, reducing sensory overload and supporting focus.

This is not just a plan. It’s a choreography:
- The **classrooms** are all **the same size and proportion**, arranged in a rhythm that children can learn by heart. Repetition is not monotony — it is **reassurance**.
- The **doors and windows** are always in the same position. Storage is integrated and concealed, avoiding visual clutter.
- **Floor textures change slightly** between circulation and learning zones, allowing even non-verbal children to sense where they are.

🧸 **Spaces That Support Autonomy and Retreat**
The school is designed around a balance between **engagement and retreat**.
- Every classroom has a **niche or corner**, where a child can sit alone, partially enclosed — still visible, but softly separated.
- There are **windows at eye level** for children in wheelchairs, and **quiet side doors** leading to small outdoor patios for one-on-one activities or regulated outdoor time.
- The **snoezelruimte** is not hidden away but part of the care wing — easily accessible but acoustically isolated.

👩‍🏫 **Supportive Architecture for Staff**
Teachers, coaches, and therapists are not stationed in offices — they **inhabit the building alongside the children**.
- The layout allows **visual connectivity** without requiring constant physical movement. From the polyvalente ruimte, a staff member can see both the classroom cluster and the therapy wing.
- **Breakout rooms** for therapy or small group work are always just one door away — no long transitions.
- There are **small resting alcoves**, even for staff — because a culture of care must include caregivers.

🌿 **The Building as a Pedagogical Partner**
This school teaches — not just through curriculum, but through **its physical presence**.
It teaches that:
- **Slowness is okay.**
- **Silence is a kind of communication.**
- **Every child deserves architecture that reflects their reality.**

The building scaffolds learning by reducing noise, guiding the body, and encouraging the senses. It is a **low-stimulation, high-warmth environment**, where **thresholds are softened**, **rooms are understandable**, and **the architecture holds everyone — child and adult — with equal dignity**

Figure 39.ChatGPT’s Textual Output

	Storytelling Technique
Software	ChatGPT
Duration	Medium
Use Case	Using textual descriptions together with generated images to convince clients
Input Type	Text, sketches, web links, images
Man-Machine Ratio	70-30
Limitations	Based on qualitative and extensive input

Table 11. Analysis of Storytelling Technique workflow.

5.2.7 Traditional interventions

At this stage in the process, we opted for traditional architectural processes to synthesise the evolving design. This corresponds to Step 4 in our framework of a traditional design approach (Figure 40), the Creative Iterative Process. The tools and methods used here were primarily analogue and design-focused: hand sketches, conceptual diagrams, floor plans, sections, perspectives, and a digital 3D model constructed in Rhino. Unlike designs in our previous years, we did not create physical models. Given the available time and the clarity already gained through digital representation, the added value of model-making was considered insufficient to justify the time investment.

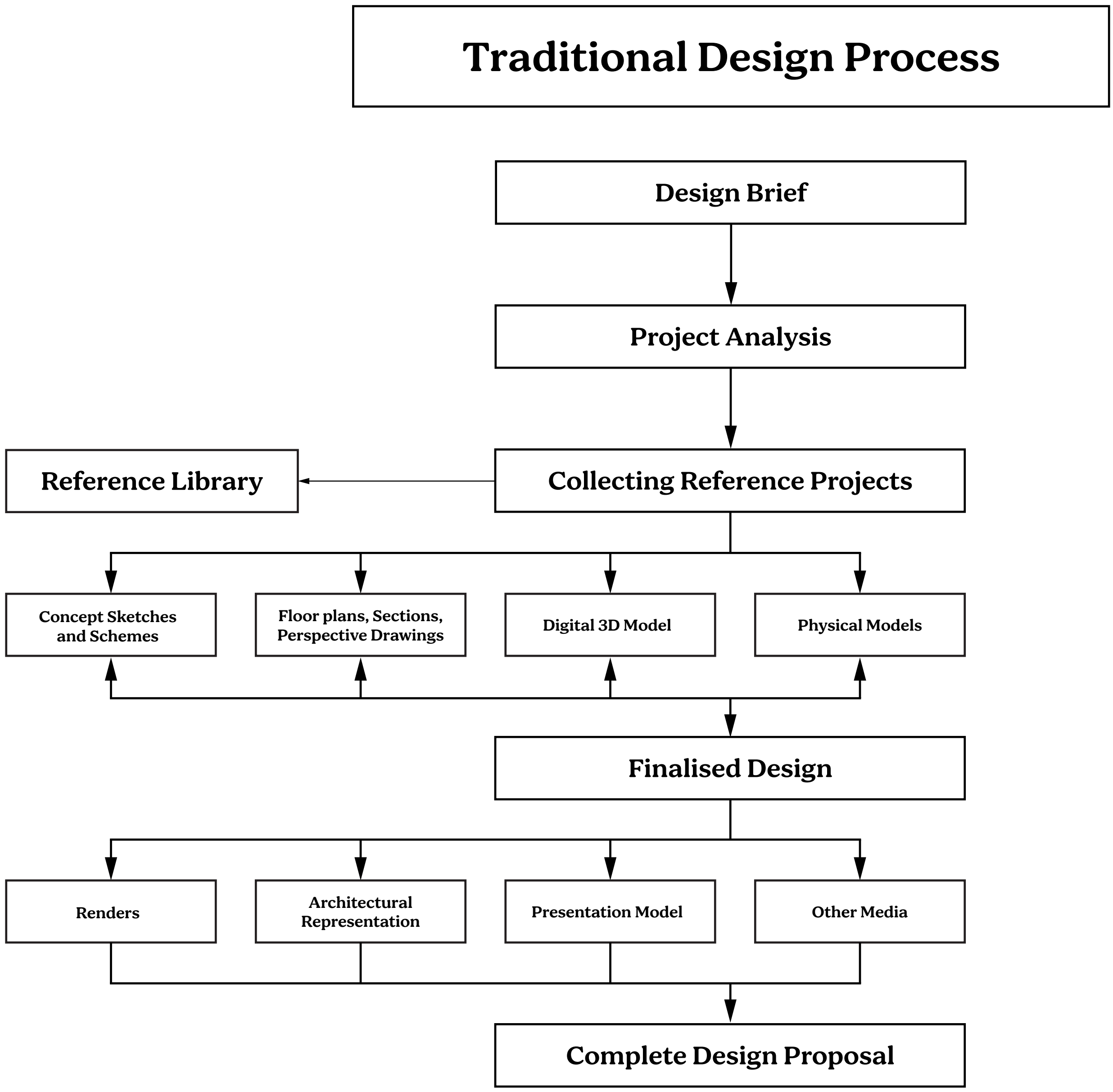


Figure 40. Our traditional design approach.

Despite the shift in approach, the influence of earlier generated imagery remained present throughout. Many of the images produced in earlier steps, particularly those interpreted and translated into architectural drawings, continued to inform our thinking, often in indirect ways. The spatial cues, atmospheric suggestions, and architectural gestures embedded in those visuals were not discarded, but rather absorbed and reinterpreted into architectural decisions.

Floorplans

The development of the final floor plans was carried out in Rhino, utilising the earlier design intentions, as well as the narratives generated by ChatGPT, as a guiding framework. While these textual descriptions and generated references served as inspiration, the act of drawing the plans by hand was a deliberate choice. It required disciplinary reasoning, where spatial relationships, functional adjacencies and programmatic logic were evaluated and constructed with architectural intent. For this reason, this step started with the traditional approach of making sketches to explore a qualitative layout. The process was supported by a broad visual catalogue of reference images and speculative plans, both self-drawn and generated, which helped shape the final spatial configuration.

As a result of this step and further exploration into requirements suggested by AGION [13], additional design intentions were developed, as shown in Table 12.

C.	Qualitative Parameters of the Building
C6	Sequence of spaces balancing structure, openness and shelter with clear, recognisable routes
C7	Spatial transitions as purposeful moments for orientation, reflection or sensory engagement
C8	Accessibility without stigma, universal design, ensuring all spaces and routes are usable by everyone.

Table 12. Additional intentions added during floor plan creation.

3D Modelling

The translation of our design into a detailed 3D model in Rhino marked the moment where all elements of the building were brought together into a single, coherent whole. While creating such a detailed model at this stage of the competition might not have been strictly necessary, a realisation that speaks to the gap between academic practice and professional efficiency, the model nonetheless proved to be an asset. It enabled us

13 AGION stands for Agentschap voor Infrastructuur in het Onderwijs, the Flemish agency responsible for educational infrastructure. agion.be

to test formal decisions, verify spatial coherence and extract various visual outputs. These included volumetric studies, light simulations and segmentation maps, which were then used as structured inputs for the final set of generated presentation images. This modelling step acted as both a verification tool and a bridge between digital experimentation and architectural communication. It was essential to verify dimensions and volumetry. It also resulted in a complete file that we could use to export specific scenes for use in the later visualisation phases.

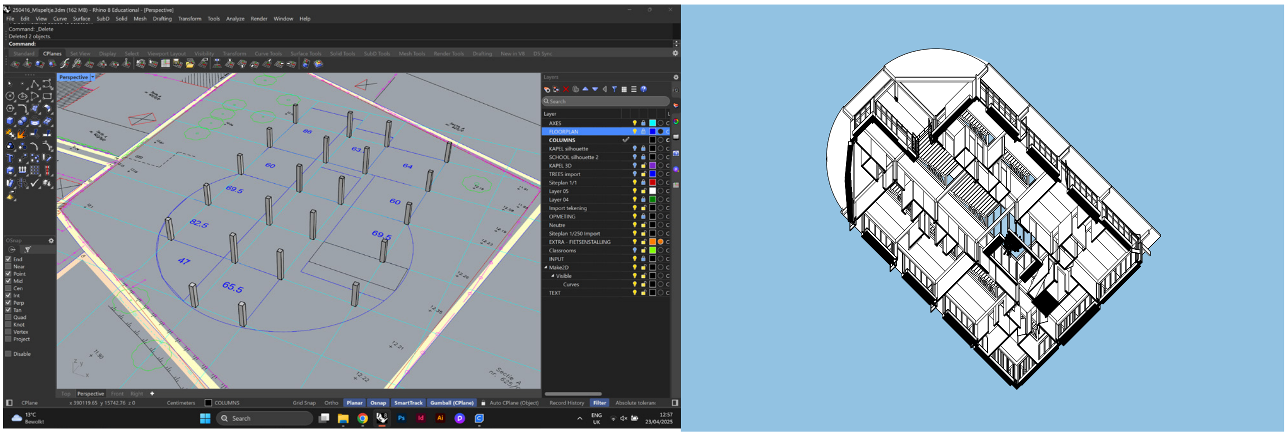


Figure 41. Resulting 3D model and axonometry.

Constructing a detailed 3D model also introduced certain limitations. The high level of detail required significant time and, more critically, compelled us to make architectural decisions that were not yet essential at the proposal stage. As a result, elements that should have remained open for interpretation were prematurely fixed. A more effective approach, as exemplified in section 4.4.2 with the PAAcademy example, would have been to develop a more abstract, volumetric model. In that instance, a simplified Rhino model provided sufficient spatial information to guide the diffusion model, while still allowing for flexible interpretation.

Sections

Constructing sectional drawings introduced a new design challenge, particularly concerning the roof structure. While some of the generated material hinted at spatial gestures such as openings or varying ceiling heights, these cues remained abstract. The actual development of the roof demanded architectural synthesis. We interpreted the suggestive qualities of light and shadow in the imagery and transformed them into an operable and expressive structural concept. Because the building has much of the functionality concentrated on the ground level, we saw the roof as a space for architectural expression. It became an opportunity for us as designers to assert authorship and embed spatial complexity, not to contrast the generated input (Fig. 42), but as a complementary gesture. This intervention exemplifies the hybrid process we employed throughout the project: a dynamic interplay between human creativity and machine-generated content, in which the designer retains agency and decision-making authority. The 3D model functioned as the primary framework through which spatial relationships, structure and proportions were tested and refined. The sections emerged here as a result of our design decisions rather than as extrapolations from generated content.

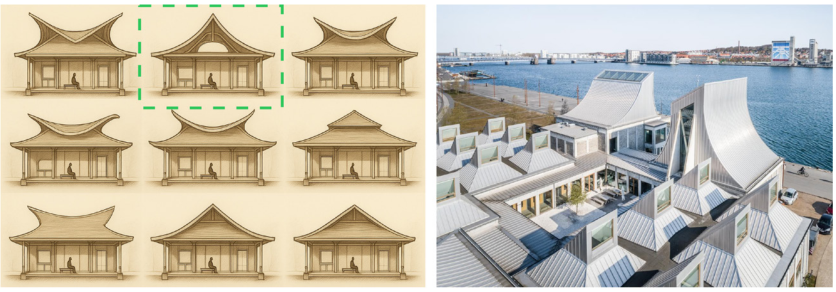


Figure 42. Inspiration through Sora generation reminiscent of Utzol Centre, Aalborg.

As a result of this step, an additional design intention was developed, as shown in Table 13.

B.	Quantitative Parameters of the Building
B14	Ceilings vary in sectional view; all classrooms capture daylight through north-oriented sky-lights.

Table 13. Additional intention derived from sectional experimentation.

5.2.8 Presentation Image Generation

Sora

In the final phase of the project, we used Sora as a generative tool to produce high-quality presentation images. Its intuitive text-to-image interface offered a streamlined and accessible workflow, allowing us to focus on articulating architecture through atmosphere, geometry, and materiality, rather than on the technical complexities typically associated with rendering software. The generation process begins with a textual prompt, and within seconds, the model produces multiple visual outputs that align with the semantic content of the prompt. This enables a responsive, iterative loop: entering a prompt, evaluating the visual output, refining the text and regenerating new images. The immediacy of this cycle dramatically reduces the technical friction often experienced in conventional rendering processes, opening up new space for conceptual exploration.

In addition to its intuitive text-based interface, Sora's ability to interpret multiple visual inputs proved particularly valuable. We leveraged this feature by uploading three complementary views (See Figure 43), exported from our Rhino model: a shaded view, an 'artistic' contour view, and a ray-traced render. Each view type served a distinct purpose within the generation process. The shaded mode visualised layer colours, effectively functioning as a form of segmentation mapping by indicating which surfaces shared the same material properties. The artistic view highlighted geometric contours and cast soft shadows, reinforcing depth and form. Finally, the ray-traced image offered a realistic impression of light behaviour and surface reflectivity. Together, these three inputs provided Sora with a more nuanced understanding of the spatial composition and intended materiality, resulting in image outputs that were not only atmospheric but also closely aligned with our architectural intent. This workflow marked a clear improvement over our earlier experiments with Stable Diffusion, where multi-input functionality was still underdeveloped at the time. At the time of writing this, June 2025, this functionality for Stable Diffusion has improved drastically in interfaces such as ComfyUI.

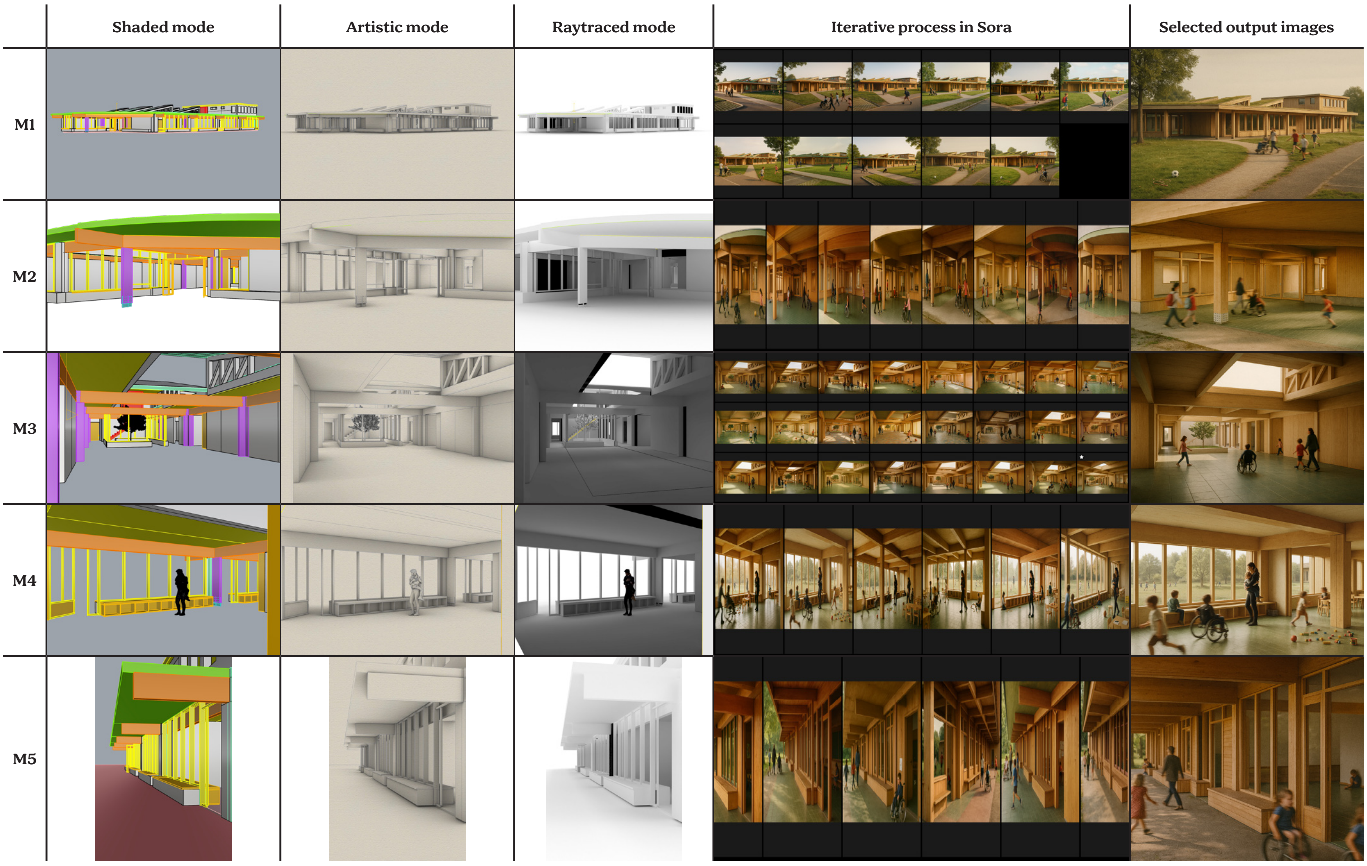


Figure 43. 3D-to-Presentation workflow through Sora.

Table 14 outlines the 3D to Presentation workflow, which we applied during the final phase of the project to generate presentation-ready images directly from a previously developed 3D model. This process combined textual prompting with visual inputs, enabling the rapid production of high-quality visuals through tools such as Sora and, to a lesser extent, Stable Diffusion. By exporting multiple view types from our 3D model, we provided the generative tools with layered spatial and material cues, which significantly improved output fidelity. The resulting workflow operated at a 60-40 man-machine ratio, allowing for quick iterations by refining textual prompts. Its strength lies in reducing technical complexity while maintaining creative control, provided that the vision and the input model are clearly defined. As summarised in the table, this medium-duration workflow demands a balance of conceptual clarity and model detail. Still, it offers a compelling alternative to traditional rendering processes for final-stage visualisations.

	3D-to-Presentation
Software	Stable Diffusion, Sora
Duration	Medium
Use Case	Quickly generating images for final presentations based on scenes exported from an earlier developed 3D model.
Input Type	Textual prompting
Man-Machine Ratio	60-40
Limitations	Requires a clear vision for the output and necessitates a 3D model or a portion of it

Table 14. Analysis of 3D-to-Presentation workflow.

At first glance, the output appears similar to that of manually rendered images produced in tools like Enscape or V-Ray. However, a closer look reveals a fundamental distinction in the process: rather than crafting every detail through manual modelling and lighting setups, Sora interprets and completes the image based on the ambiguity and richness of the prompt. This interpretive nature results in a workflow that is less about precision and more about architectural storytelling.

For example, when generating two images from the same base prompt, we noticed how each version diverged slightly, each interpreting the same textual input with subtle variations. This enabled us to select the one that best matched our design intentions or, if needed, to tweak the prompt and regenerate until the outcome aligned with our vision. In practice, we spent approximately 30 minutes per image, with the majority of that time dedicated to constructing and refining the prompt.

Writing prompts for Sora is a strategic exercise. The foundation of the prompt remains consistent across multiple generations, ensuring visual coherence and reducing redundant effort. Only key variables, such as the scene's activity, time of day or composition, are modified. Stylistically, we aimed to create a distinct atmosphere: one with an analogue, grainy texture, poetic lighting, and a sense of imperfection. Descriptions such as "Shot on an analogue camera," "Grainy aesthetic with dust on the floor and scratches on the walls," and "Lens flare with dreamy atmosphere" proved effective in evoking this desired tone. Human presence was another critical element: prompts like "We see children walking with an adult, one in a wheelchair, but they are not the protagonist of the image" allowed us to suggest occupancy and scale without overpowering the architectural focus. Compared to earlier tools such as Stable Diffusion, Sora demonstrated a significant improvement in language comprehension and prompt responsiveness. Its training on large language models enables more nuanced interpretations of natural language, meaning that minor grammatical or spelling errors no longer significantly disrupt the outcome. This makes the workflow not only faster, but also more forgiving, an essential factor in high-paced design iterations.

Although prompt engineering remains an open-ended discipline, we based our approach on one guiding principle: describe all the elements that one personally values in architectural photography or renders. In our case, this meant textures, lighting conditions, mood and subtle material cues. The resulting images are thus not just visualisations; they are projections of how we, as architects, envision the lived experience of our design. Importantly, the output of this process should not be seen as a definitive visual conclusion. Rather, it serves as a communicative tool, and its use in design presentations can vary depending on the stage of the process and the audience. We distinguish between two approaches:

1. Architect's Selection

In this method, we make a curated selection of one or a few images that we believe best represent the project. This aligns with conventional practice, where the architect aims to convey their interpretation of the client's desires visually. The result is focused and persuasive but also carries the risk of being an "all-or-nothing" proposal, where a single image must succeed entirely on its own.

2. Image Collage for Dialogue

Alternatively, we can present a broader range of images, each slightly different in tone, composition or articulation. This approach embraces interpretation and invites feedback. It enables clients, who may not have a fully formed vision, to respond to visual cues and articulate preferences. Here, the multiplicity of images becomes a strength: it reveals the dynamic nature of the proposal and helps the client co-author its development. This method is particularly valuable in the early design stages, while the curated approach becomes more effective once a mutual understanding has been reached.

The iterative nature of this process is shown in our output for Moment 4, where numerous images were generated based on variations of the same core prompt. Each iteration shared a visual and stylistic coherence, yet differed in composition, detail and expression. This not only illustrates the high consistency in visual and aesthetic output quality but also underlines the importance of refinement in achieving the desired architectural communication. These were by no means "one-click" results; instead, they were the product of a controlled and critical prompting process.

Stable Diffusion

Earlier, we focused on utilising Stable Diffusion for generating architectural images. At that stage, we explored various prompt structures and visual styles to investigate the tool's potential for producing atmospheric and conceptually rich visuals. These explorations informed our understanding of Stable Diffusion's possibilities and constraints within an architectural context. For the generation of final images, we attempted to apply the previously developed styles. All settings were kept consistent, and the same textual prompts and base image inputs we used earlier in Sora were implemented here. Since our workflow within Stable Diffusion was primarily limited to single-input generation through ControlNet, we selected Shaded Mode as the input format. This decision was based on earlier positive results, such as the PAAcademy example in section 4.4.5, where the shaded mode yielded outputs with clear spatial qualities and architectural articulation.

However, the resulting outputs from Stable Diffusion (Fig. 44), even when using the same style templates, proved to be less consistent and of lower visual quality compared to the images generated in Sora. Several limitations became apparent, beginning with the observation that Stable Diffusion still demands a significantly more technical approach. Fine-tuning outputs requires constant adjustment of parameters such as denoising strength, image resolution, control types and seed values. This level of manual intervention introduces complexity into the workflow, extending the time required to achieve satisfactory results. Secondly, the generation of a new or adjusted visual style within Stable Diffusion, especially under time constraints, was found to be labour-intensive. Creating a stylistically coherent and atmospheric output demands substantial trial and error, particularly when attempting to reproduce specific qualities such as analogue texture, poetic lighting, or human motion with blur. As such, developing an entirely new style for the final presentation would have consumed a disproportionate amount of time at a stage when the project is being finalised.

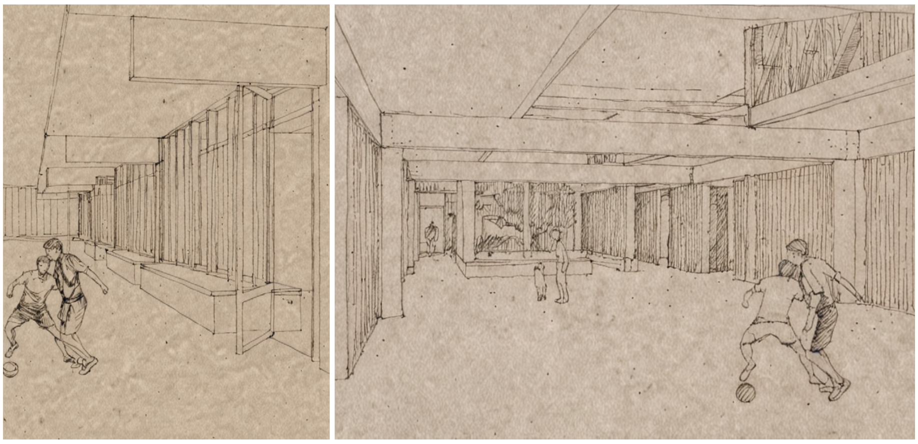


Figure 44. Results using older Stable Diffusion presets.

By contrast, Sora provided a far more efficient and user-friendly iterative process. Adjusting prompts and generating multiple versions of the same scene took only seconds, allowing for rapid refinement based on real-time visual feedback. This accessibility and responsiveness proved to be a key factor in our decision to prioritise Sora in the final stages. The learning curve for Sora is considerably lower than that of Stable Diffusion; yet, the tool quickly delivers outputs of high and more consistent visual quality, especially in terms of atmosphere, materiality, and human-scale integration.

It is worth noting, however, that Stable Diffusion retains certain advantages. Unlike Sora, it operates locally, making it a preferred tool for handling sensitive or confidential project data. Moreover, models such as FLUX offer the potential for generating more photorealistic and stylistically complex imagery, as they are built upon larger and more advanced models. Unfortunately, due to hardware limitations, we were not able to test FLUX ourselves. Nevertheless, it remains a promising direction for future exploration, particularly in contexts where local workflows and complete control over model customisation are necessary.

5.2.9 Postproduction

Once the final generated images were selected, a final step of manual augmentation was undertaken to further enhance the clarity and expressiveness of the architectural narrative. While the outputs generated by Sora were already of high visual quality and stylistic coherence, subtle adjustments were sometimes necessary to reinforce spatial logic, highlight intended programmatic elements or strengthen the emotional impact of the image.

To carry out this process, we employed Adobe Photoshop's Generative Fill function. This enabled us to seamlessly integrate new elements into the image without compromising overall visual integrity. For instance, in Figure 45 we see that the raw output from Sora concerning Moment 4 effectively captured the atmosphere of the learning environment but lacked specific contextual cues that were essential for communicating our spatial and social intentions.

The following modifications were made using Photoshop:

Figure 45. Left: Raw Sora output, right: postproduction.

A partition wall was added between classrooms on the right side of the image to clarify programmatic separation.
A child's drawing was introduced on the adjacent wall to emphasise the appropriation of space by children and evoke a sense of a lived-in atmosphere.
Books were added to the integrated bench in the centre of the image to reinforce the space's identity as a learning environment.
A playing child was inserted in the outdoor area on the left, serving both as a narrative anchor and a tool for scale referencing.

Beyond these compositional adjustments, we applied a black-and-white filter to the final image. This decision stemmed from a personal preference: by desaturating the image, we aimed to enhance its poetic quality and evoke a subtle sense of nostalgia, while also allowing the viewer to interpret the materiality, for example. This tonal shift aligns with our broader design intention, in which atmosphere and emotion are central themes.

The monochromatic aesthetic also unifies the image compositionally, allowing spatial relationships, light and form to take precedence over the distraction of colour.

To refine the architectural narrative in the final visualisations, we implement a post-processing technique, focusing on the targeted enhancement of AI-generated images. This Correction Technique workflow, summarised in Table 15, involved the use of Photoshop's Generative Fill function, which is similar to the Retouche technique workflow described in Section 4.4.5. Operating at a 50-50 man-machine ratio, this medium-duration workflow enabled the integration of context-specific elements without compromising the integrity of the base image. Although this step presents an additional layer to the design process, it proved instrumental in communicating the designer's vision more precisely to clients and peers.

	Correction Technique
Software	Photoshop Generative Fill, Stable Diffusion
Duration	Short
Use Case	Refining specific elements in the final phases to visualise the project
Input Type	Generated images
Man-Machine Ratio	50-50
Limitations	Additional step in finalising the design process

Table 15. Analysis of Correction Technique workflow.

It is worth noting that some of these manual adjustments could now be achieved directly within Sora through its recently introduced inpainting functionality. This tool enables users to select a specific area of an image and regenerate that portion, allowing for targeted interventions similar to those performed in Photoshop. However, in our experience, the results of this inpainting feature often involved subtle but noticeable alterations to the complete image, even outside of the designated edit zone, as seen in the example below. While this can improve coherence in generated scenes, it limits the degree of control when precise preservation of existing elements is required. In this context, our manual retouching method, performed in Photoshop, retains a key advantage: it allows us to maintain the integrity of strong initial outputs while making intentional and limited adjustments.

5.3 Design Evaluation

5.3.1 External evaluation by practising architect

To critically reflect on the quality of our design and its visual presentation, we organised a feedback session with architect and educator Nicolas Petillon of the office Petillon Ceuppens, an experienced practitioner who is familiar with both architectural competitions and public projects. His commentary provided valuable insight into both the spatial qualities of our proposal and the effectiveness of our representational strategies.
In terms of the design itself, Nicolas noted that the floor plan reads clearly and only minor corrections were needed. One of the first points he raised was the positioning of the toilet doors within the classrooms. In our proposed version, the doors opened directly into the main teaching space, which reduces the functional usability of the room. Interestingly, this specific issue had been addressed differently in an earlier version of our plan, where the doors were positioned in the entry zone. Another comment related to the layout of the personnel showers. While we initially provided two showers based on our reading of the POR (Programme of Requirements), Nicolas suggested a more efficient layout with only one shower. He explained that, in reality, additional showers are often among the first elements to be removed from the program due to cost and practicality, particularly in competition settings. In response to this critique, we revised the plan accordingly. This form of experiential knowledge, which is an understanding that transcends purely technical or academic concerns, highlights the value of involving practising architects in evaluative moments. This feedback was not merely a formal critique, but was grounded in an awareness of execution and feasibility.

Regarding the representational imagery, Nicolas emphasised the importance of combining section and plan drawings to convey spatial intent. For instance, the shed roofs gain meaning only when read in conjunction with the accessibility and logic of the floor plan. Presented in isolation, they lose part of their architectural rationale. His critique of our axonometric drawing followed a similar reasoning: while the bottom-up axonometry illustrated many constructional details, it obscured the plan layout due to its heavy layering. Based on this input, we opted to reconstruct the axonometric view from a top-down perspective, removing the roof elements to highlight the spatial organisation.

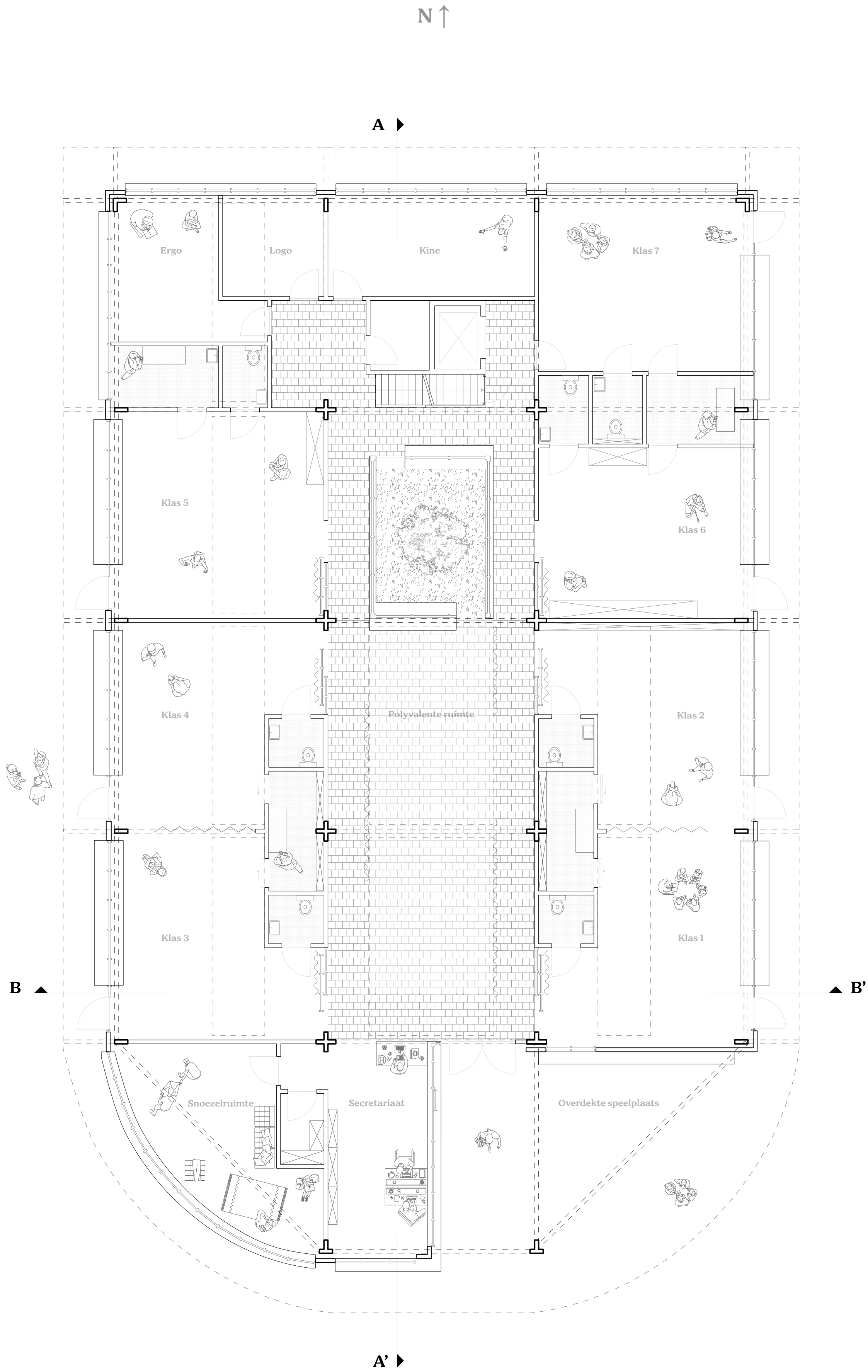


Figure 46. Floor plan +0, 1/200.

N ↑

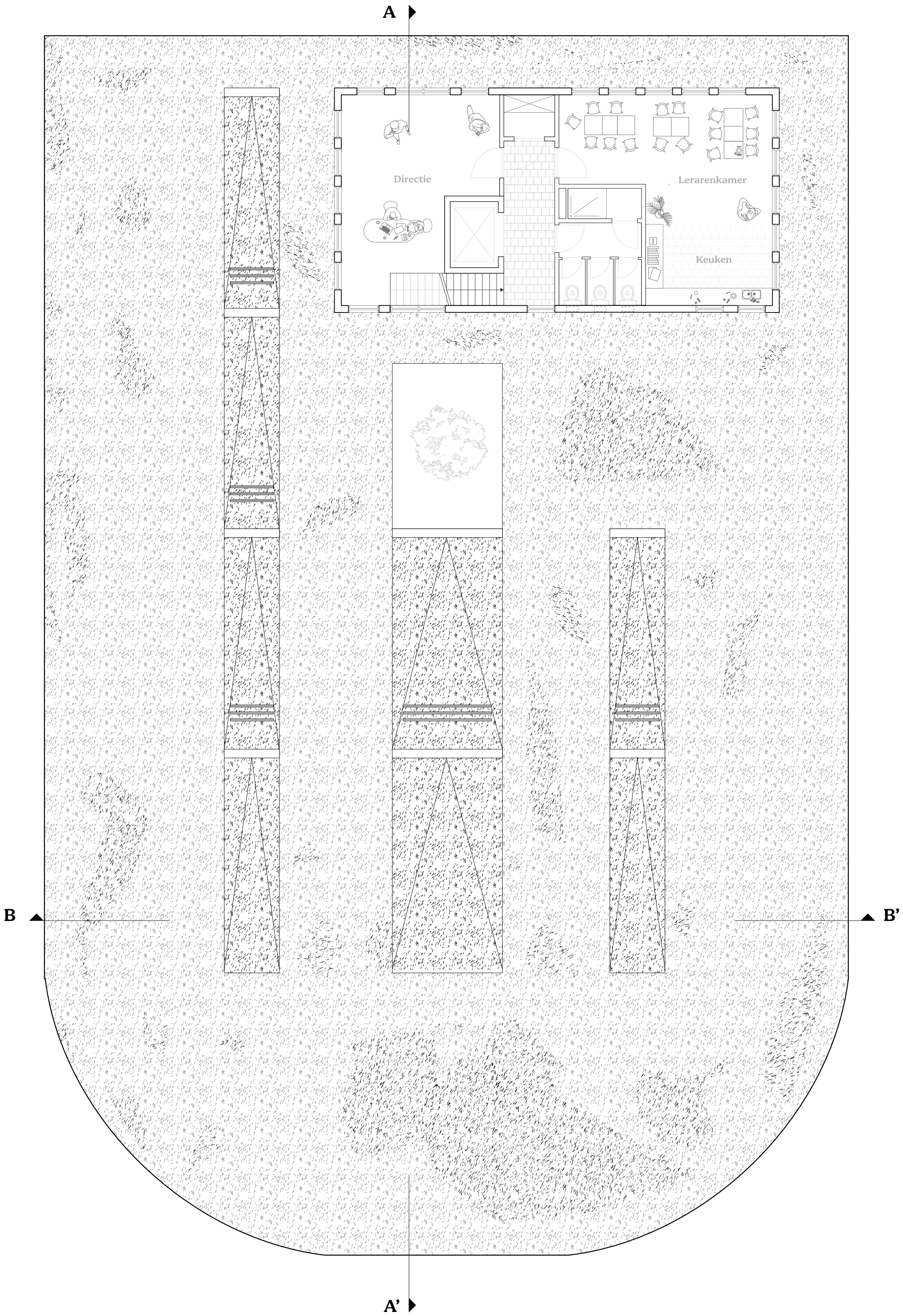


Figure 47. Floor plan +1, 1/200.

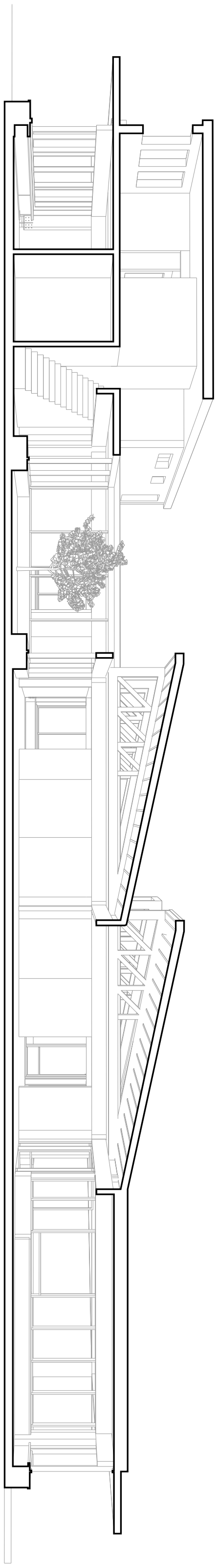


Figure 48. Section AA', 1/200.

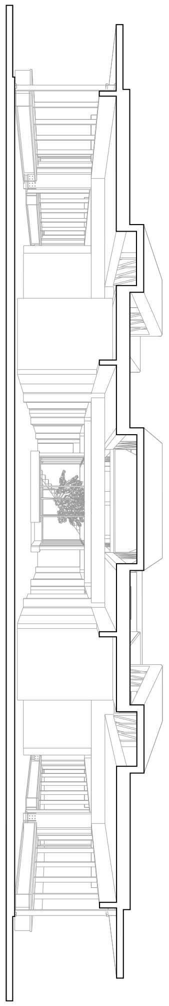


Figure 49. Section BB', 1/200.

One of the most resonant points in Petillon's feedback concerned the presentation images for the key moments. During the earlier exploration phase of this research, we observed that hyperrealistic generated visuals tend to provoke highly critical reactions, often focusing on minor inconsistencies or unrealistic details. Petillon's reaction confirmed this once more. As he rightfully questioned: "Why does the size of the tiles change between the images?" Such comments are entirely valid because architectural images inherently carry a sort of implicit promise. When these visual outputs aspire toward photorealism, they are interpreted as fixed, final representations, even when the design is still evolving.

This realisation steered us toward a deliberate shift in visual strategy. Rather than continuing to refine hyperrealistic outputs on platforms like Sora or Stable Diffusion, we adopted a more abstract and interpretive form of representation. For example, while visualising Moment 3, which depicts the polyvalent space, we struggled to generate a satisfactory image due to the lack of defined materiality and detail at that point in the design process. By applying a sketch style and utilising a pre-trained LoRA model called Handsketcher-XL on Stable Diffusion, we were able to convey atmosphere and intent without being bound to unrealistic levels of visual precision, allowing for interpretation by the client and a better understanding of their needs, without being distracted by off-putting details. This approach aligned well with the narrative structure of our design, where the building is experienced through a series of evolving moments.

An important aspect to note is that all sketch-style images were generated from the same prompt and input image, yet the outputs varied considerably in tone or line quality. This confirms the need for an iterative process in curating visuals. Here, our Canvas Storm workflow proved helpful: by generating a batch of 10 to 20 images with randomised seeds, we could select the version that best aligned with our intended atmosphere. Rather than attempting to control every parameter in the generation process, this approach embraces variation as a strength, focusing on curation rather than correction.

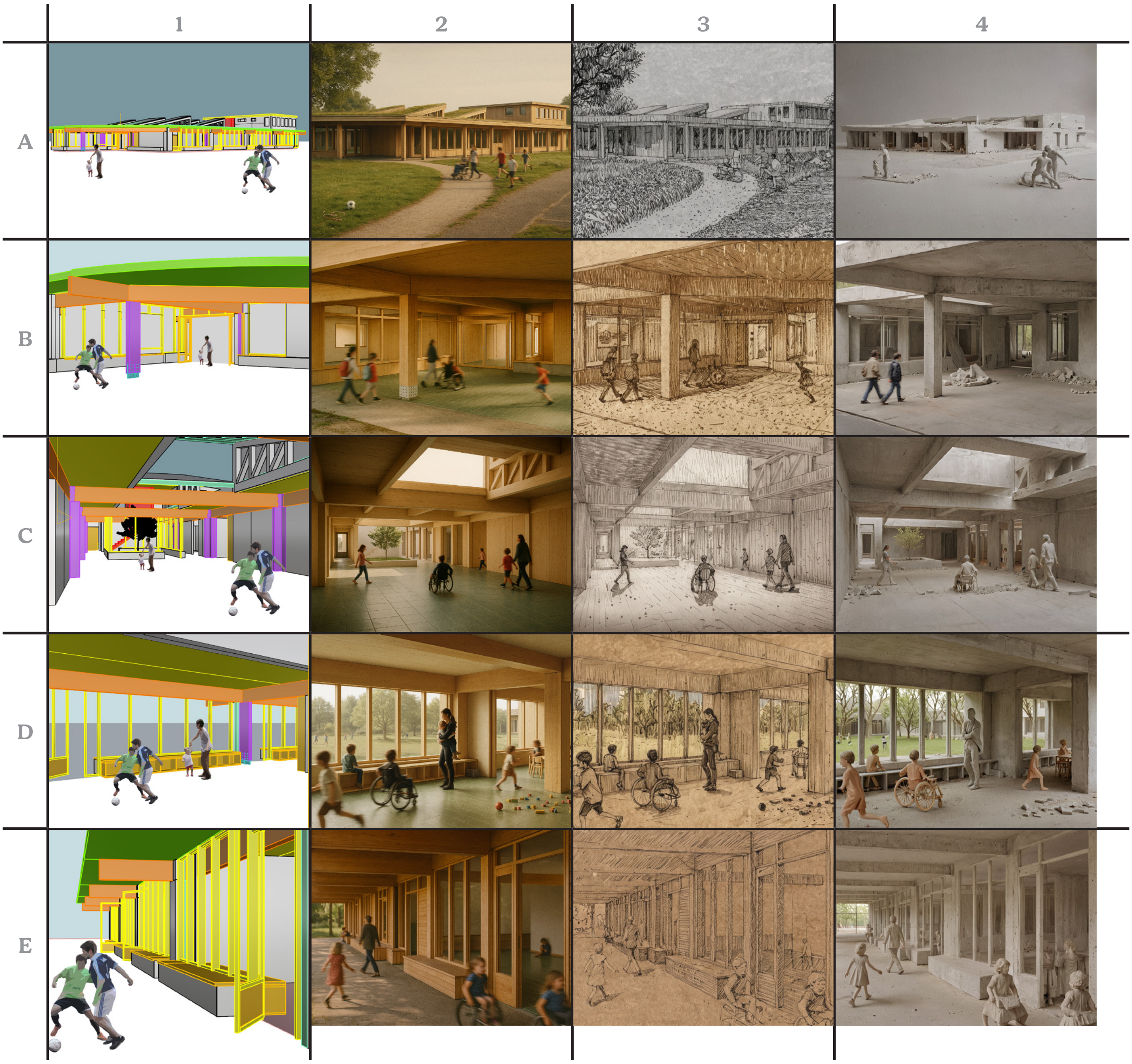


Figure 50. Process of generating more abstract imagery.

5.3.2 External evaluation by competition board

A significant note in Team Vlaams Bouwmeester's evaluation was that some choices in our design are overly detailed for a proposal phase. It is crucial to be able to take a step back. Adjustments and additions are made based on client feedback through ongoing dialogue and communication. A good example here is that the floor plan included detailed placements of sanitary elements such as toilets and sinks, suggesting a level of detail more appropriate for later design stages. Similarly, the classroom windows were not merely positioned but already equipped with integrated benches that combined seating and storage functions. While these decisions reflect our ambition to create a thoughtful and user-oriented design, they also demonstrate a tendency to advance the project beyond the exploratory nature typically expected in a proposal phase. This level of definition can unintentionally limit flexibility and should therefore be approached with caution in early-stage workflows.

The client expresses a preference for fencing and terraces rather than a fully open outdoor space. The upstairs functions raise questions about the footprint, maintenance of a lift and the physical separation of teaching staff. The lack of visualisation of the outdoor areas and traffic circulation is a valuable comment; a mobility plan would be beneficial. For competition designs, focusing on access points, roads, and cycling infrastructure is essential, as different modes of transport play a defining role in the proposal. For the chapel, a passive approach is taken, focusing more on custodianship than on actively maintaining the structure and restoring its function. There is also a need for clear differentiation, for example in colours, to enhance children's orientation and readability of the spaces. Finally, practical aspects such as a tree in the patio, technical systems, budget, and timing are crucial and need to be considered from the early stages of the design.

Table 16 summarises the derived Communication Generation Workflow, which leverages tools such as Sora or Stable Diffusion to produce rapid visualisations based on textual prompts. This process supports a 70-30 man-machine ratio, prioritising speed over precision to facilitate fluid discussion and timely adjustments. It enables teams to remain responsive to feedback, engage in shared exploration of design intentions, and avoid prematurely defining architectural elements during the proposal stage.

	Communication Generation
Software	Sora, Stable Diffusion
Duration	Short
Use Case	Quickly generating images while in dialogue with others to have clarity in design ideas.
Input Type	Textual prompting
Man-Machine Ratio	70-30
Limitations	Accuracy makes way for speed.

Table 16. Analysis of Communication Generation workflow.

5.3.3 Quantitative Evaluation

A key question that arises in our conversation with engineer and professor Johan Van Rompaey is what the relevance or applicability of these tools might be in the professional field. More specifically, what concrete impact can be expected, and which challenges remain unresolved? In earlier transitions, such as the shift from hand drawing to digital vector-based design, specific tasks were replaced. The current transformation, however, affects not just the tools but also our foundational knowledge and the way we communicate within the design process.

There were tangible shifts with the transition from hand-drawn documents to digital tools, such as a reduction in the number of people needed for specific workflows, from ninety to thirty in some cases, illustrating the potential for increased efficiency. This leads to the broader question of which aspects we would want to see replaced by AI. Maybe these are more profound changes in the way we collaborate, suggesting that the real value of AI may lie in transforming our modes of communication rather than replacing roles. If we aim to make a difference, we must engage with new directions that others may not have the time or resources to explore. In this view, generated imagery is a means of improving clarity and reducing the duration of communication processes, rather than a substitute for the architect's role.

Architects still have a legitimate place in the design process because they can anticipate and guide the transformation of the built environment. The act of designing does not stop at the moment of drawing but continues into construction, where interpretation and adjustment are crucial.

The ability to translate abstract client intentions into spatial proposals remains a core responsibility of the architect. It is often challenging to articulate what the client truly wants, and these tools may help express these latent needs. It can help turn a technical plan into something that feels alive and meaningful.

Despite its potential, AI still struggles with regional disparities in planning, culture, policy and regulation, which means we cannot yet rely on it for autonomous design solutions. More broadly, there is currently no AI tool that integrates all relevant aspects of the architectural design process. Crucial factors, such as building codes, contextual regulations, and local planning, are often overlooked or absent from the generative logic. However, it could play an essential role in supporting the translation of the client's position within a project. This includes identifying and addressing conflicting interests between stakeholders and enhancing communication efficiency, particularly in complex projects involving multiple actors.

Design-driven offices, such as those that participate heavily in competitions, often struggle to operate at the highest level of efficiency. As a result, they are more likely to rely on younger and less expensive staff, including interns. This practice raises broader questions about the sustainability of the current model and the long-term impact on professional development within the field, while also opening doors for the implementation of new tools to optimise the limited time these offices have.

From a technical and financial perspective, reference frameworks such as those from Agion provide useful benchmarks. Subsidies are typically calculated at around €1,800 per square meter, indexed annually; however, in practice, a figure closer to €2,200 per square meter, excluding VAT, is more realistic. Their scope includes the design of all fixed equipment, such as built-in furniture, and calculations encompass the building footprint plus an additional meter around it. Landscape architecture is considered a separate domain, often handled in an independent file, not necessarily assigned to either the architect or the engineer. Particularly on sensitive plots like this, it is advisable to avoid taking responsibility for exterior landscaping, as it is difficult to control.

A common rule of thumb divides the architect's compensation into three phases: approximately 30% for all work up to and including the preliminary design, 40% for work up to the construction file, and the remaining 30% for site supervision. For a standard project, approximately 13% of the total budget, in this case a total budget of €2,880,000 excluding VAT, is generally allocated to design professionals and technical consultants. Of this, 6-7%, in this case €185,000, typically goes to the architect. This implies that approximately €56,000 is available for the initial phase, up to the preliminary design. Based on estimations provided to us by Hedwig, an architectural office would need to stay under roughly 600 billable hours to remain within budget, assuming an average hourly rate of €60 to €90. These calculations are not coincidental: the project budget is designed to reflect such constraints. However, by clearly documenting our hours and comparing them to this structure, we gain insight into whether we remain within these feasible margins.

Assuming the design phase for this project began on February 26th, 2025, and concluded on May 5th, 2025, we can make a reasoned estimation of the total time spent on the design. Over this period, there were 49 working days, in which we averaged about six hours of work per day. As a team of two, and given the full-time nature of our studies, we estimate that approximately three-quarters of our available time was dedicated to the master's thesis, with the remaining quarter allocated to other academic commitments.
Within the time spent on the project, roughly 10% was devoted to experimenting with new tools, refining settings, and testing workflows. While this was essential to remain relevant within the evolving field of design technology, it also reflects a learning curve that, in future projects, would likely require far less time. A further 10% was spent on documentation and critical reflection. This means that approximately 80% of our available time was directly invested in design work.

When converted into hours, this results in an estimated total around 350 hours spent on the design itself. This is significantly less than the benchmark of 600 hours used within professional architectural practice. However, this discrepancy must be understood in its context. Our academic process excluded several time-intensive components that are typically part of a professional workflow. These include regular meetings with the client or developer, coordination with engineers or technical consultants, and the production of physical models. Each of these steps represents not only a time investment but also a deepening of the architectural proposal, bringing it closer to a buildable reality.

At the same time, we must acknowledge that our process was not necessarily more efficient. Lacking the experience of a seasoned design professional, we inevitably spent more time making decisions that, in practice, might be resolved more intuitively. Designing a competitive proposal is a skill in its own right, knowing where to invest time, when to move on, and how to focus effort is something acquired through repetition and experience. The resulting 'buffer' between our documented hours and the 600-hour benchmark, is not a definitive measure of efficiency, but rather a space of possibility. It highlights areas where further development, coordination, and technical refinement would be required for the project to mature into a competitive, buildable proposal. In that sense, this comparison should not be seen as yielding a single conclusion.

5.3.4 Self-Evaluation

To critically assess the outcomes of our design process, we developed Table 17. It compares our original design intentions against the extent to which these were realised. This reflection is structured across the three main categories, each encompassing multiple intentions: surroundings, quantitative parameters, and qualitative parameters. For each design intention, we assigned a score reflecting the degree of achievement: high (H), medium (M), or low (L). In some cases, a score of not available (N/A) was used when evaluation was not feasible within the current scope. This table allows us to formulate a structured and honest appraisal of the project's architectural qualities and limitations, independent of the extent to which AI tools were integrated. Rather than evaluating innovation or experimentation, the aim here is to understand how well the spatial, social, and technical ambitions of the design were addressed in the final design. When comparing the early AI-generated imagery with the final architectural render, a subtle yet undeniable influence emerges. (Figure 48) The first images, produced during the concept phase using generative tools, did not serve as a literal blueprint, but its atmosphere, material palette, and human-centred composition evidently shaped our design trajectory. While our floorplan interpretations developed independently through programmatic and contextual constraints, the welcoming facades, integration of nature, and emphasis on accessibility present in the AI image appear echoed in the built form. The final render, though more grounded in real-world feasibility and functional clarity, retains key spatial gestures first imagined through speculative synthesis. Juxtaposing both visuals reveals the extent to which early atmospheric images can prefigure spatial thinking, subtly steering design decisions not through precision, but through suggestion.

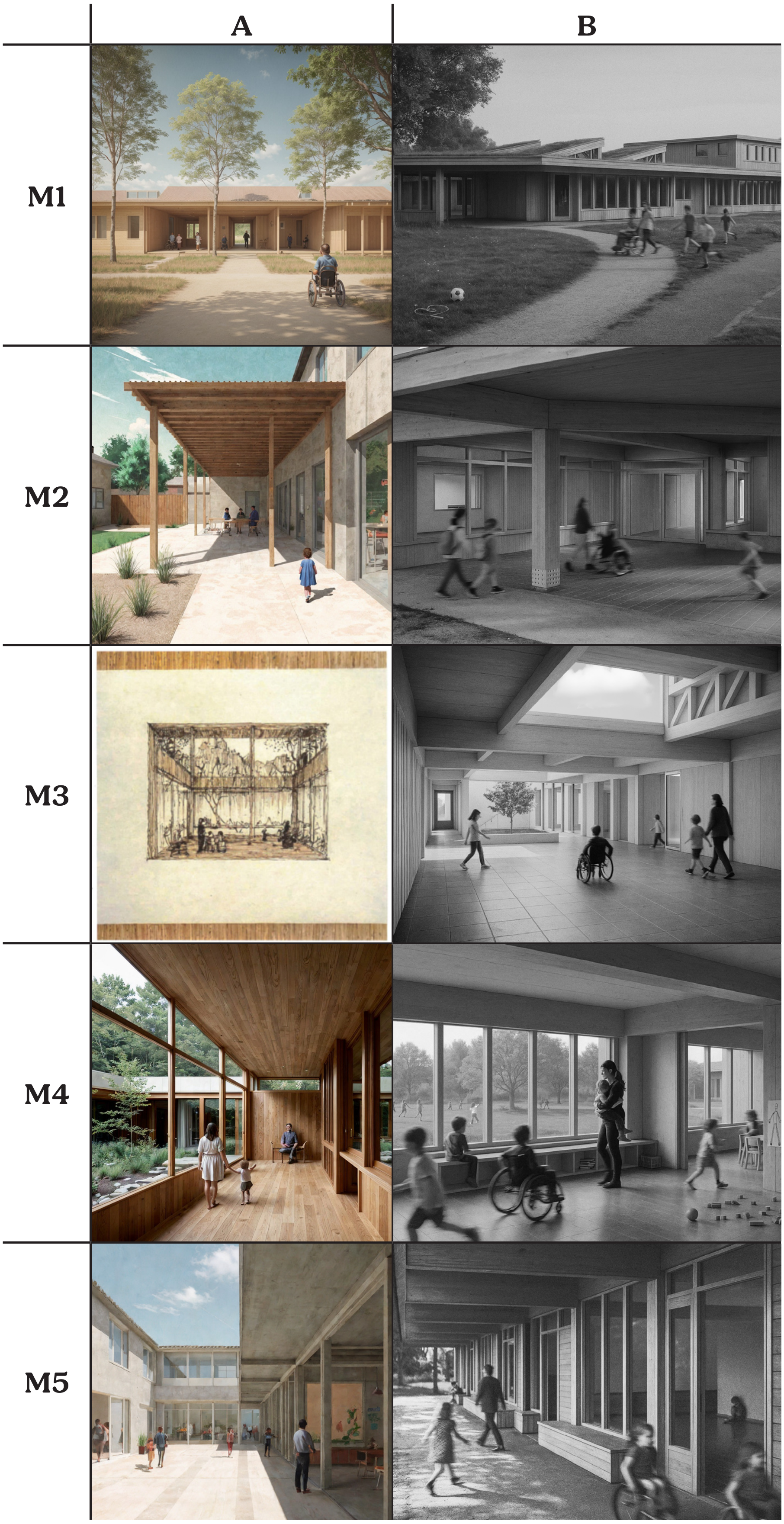


Figure 51. Design elements that reference the original generations.

	Design intention	Score	Explanation
A.	Surroundings		
A1	Thoughtful integration of a new access road, ensuring smooth flow and a safe drop-off zone	L	No detailed mobility plan produced, despite neighbourhood concerns about traffic; intention remained implicit
A2	Possible integration of the adjacent city-owned chapel	H	Chapel repurposed as therapy space during school hours and community venue after hours, reducing the newly built area by ≈ 25 m²; feasibility depends on municipal subsidies.
A3	Continued availability of the site as a public play area outside school hours	H	The design maintains unrestricted public use, with no anticipated physical or managerial conflicts.
A4	Inclusion of outdoor green spaces and playgrounds that support both relaxation and education	H	The design maintains unrestricted public use, with no anticipated physical or managerial conflicts.
B.	Quantitative Parameters of the Building		
B1	All need-to-have programme spaces are included in the design	H	Creative stacking of support on +1 keeps the ground-floor footprint compact while meeting the full brief.
B2	All nice-to-have programme spaces are included in the design	M	All optional spaces are accommodated, except for dedicated staff bike storage.
B3	Delivery of the project within a budget of € 2,880,000 (excl. VAT and fees)	N/A	Cost accuracy could not be validated within the thesis scope (see § 5.3.3)
B4	Entire learning environment on ground floor, support functions on +1	H	Ground-level learning is achieved; support is stacked above to minimise footprint, though lift cost and maintenance are a trade-off.
B5	Integration of generous play and sports facilities to foster interaction and collaboration	M	The central polyvalent hall doubles as a play and sports zone and can be subdivided, but its versatility is not fully illustrated.
B6	Aim for a minimal ecological footprint through sustainable and energy-efficient design.	H	Compact plan, CLT/timber structure and 8×8m module support low embodied carbon and future adaptability
B7	Canopy creates a transition zone between inside and outside	H	Continuous overhangs limit solar gain and provide a sheltered edge.
B8	Entrance canopy stretches out like an open arm, generous yet unpretentious	H	The canopy clearly articulates the entry and mediates the transition from indoor to outdoor spaces.
B9	Low gable green roof, sloping northward to bring natural light into each space	H	Roof geometry admits diffuse northern light to learning spaces effectively.
B10	Central patio-courtyard that evokes light and nature into the building	H	The courtyard delivers daylight and biophilic elements; however, maintenance, planting trees, and sunlight still need to be addressed.
B11	Floor plan organised on a clear 8×8m structural grid	H	Regular timber grid ensures structural clarity and long-term programme flexibility.
B12	Classrooms can be opened up to form one larger classroom	H	Operable partitions permit team-teaching and large-group activities

B13	Each classroom has direct access to the outdoors	H	Every class has a dedicated door or terrace to an external play space
B14	Ceilings vary in sectional view; all classrooms capture daylight through north-oriented skylights.	H	North-sloping trusses with 8×2m (classrooms) and 8×4m (hall) skylights prevent glare and overheating.
C.	Qualitative Parameters of the Building		
C1	Flexibility to accommodate diverse pedagogical methods	H	Open grid and movable walls enable multiple teaching formats and future re-programming
C2	Architectural expression aligned with the values of care, inclusivity and educational excellence	M	Clear modular order aids legibility; however, the overall character tends toward generic and lacks intimate niches for overstimulated pupils.
C3	Use of warm, inviting materials and colours to create a homely and secure atmosphere	H	Timber surfaces and acoustic buffering between classrooms and hall cultivate warmth and privacy.
C4	Design relates to the human scale	M	Classrooms are finely tuned, but the polyvalent hall remains oversized and under-programmed
C5	Strong legibility and visible boundaries for optimal orientation	H	Single-direction circulation, daylight cues and window-bench thresholds facilitate intuitive wayfinding.
C6	Sequence of spaces balancing structure, openness and shelter with clear, recognisable routes	M	Logical progression achieved, but insufficient colour/texture contrast to reinforce spatial identity
C7	Spatial transitions as purposeful moments for orientation, reflection or sensory engagement	L	Logical progression achieved, but insufficient colour or texture contrast to reinforce spatial identity.
C8	Accessibility without stigma, universal design, ensuring all spaces and routes are usable by everyone.	H	One principal entrance, level routes, and adjacent lift-stair pairing ensure equitable access for all users.

Table 17. Complete evaluation of design intentions.

6. Conclusion

6.1 Findings

This research aimed to investigate how diffusion model-powered workflows can transform architectural design processes. Through the development and implementation of multiple workflows, we were able to test not only the technical capacities of these generative tools but, more importantly, their potential when embedded within different stages of the design process. Table 18 provides a structured overview of these workflows, comparing aspects such as input type, required software, duration, human-machine ratio, limitations, and, most critically, the specific use case and corresponding design phase. This overview enables direct comparison and informed conclusions about when and how each workflow may be most effective.
Generative AI, in the form of diffusion models, proves particularly valuable in the early stages of design, when ideas are still vague and malleable. In this phase, where freedom and exploration matter most, AI can open a broader frame of reference, supporting rather than restricting creativity. It does not have to threaten authorship or intuition; instead, it can act as a partner, extending the architect's reach. Workflows such as Canvas Storm and Communication Generation exemplify this potential. They offer rapid image generation from minimal or suggestive input, enabling the designer to test atmospheres and geometries in minutes rather than hours. The time saved in these workflows is not necessarily reinvested in finalisation, but rather in expanding the conceptual field before decisions are made.

Another important application lies in the visualisation of design intentions and variations. Workflows such as Communication Styling, Retouche Technique, and Storytelling Technique were used to illustrate design options or clarify narratives. Although these images are less precise than those produced through traditional rendering software, they were generated significantly faster and with enough quality to support communication and decision-making. Their role is not to replace professional renders, but to serve as an agile, low-barrier alternative that enables reflection and dialogue.

Workflows such as 3D-to-Presentation and Correction Technique, on the other hand, align more closely with later design stages. These rely on a pre-existing 3D model and require more apparent architectural intent. They are not inherently faster than traditional methods. Still, their advantage lies in flexibility: changes to images can be made through prompting or partial regeneration rather than restarting entire render sequences. These workflows are especially useful when time is limited and atmosphere remains more important than precision. Still, the danger of overcommitting on design decisions too early remains a consideration.

The comparative matrix, visualised in Table 18, shows that no single workflow is universally superior. Each method offers specific benefits and trade-offs depending on the design phase, available input, and desired output. For example, Canvas Storm and Communication Generation provide substantial advantages in ideation speed, enabling rapid conceptual development. Conversely, Floorplan Interpretation or Collage Technique demand more from the designer, particularly in curating or interpreting qualitative content, and serve better as complementary tools rather than as standalone solutions.

These findings confirm that diffusion model-powered workflows do not necessarily replace architectural design, but can meaningfully reshape parts of it. They redistribute creative labour across phases, shifting time and effort away from technical production towards conceptual decision-making and communication. The choice between traditional and generative methods is not binary; it depends on the project's intent and timing within the broader scope of the project. If refinement and accuracy are essential, traditional workflows remain the standard. If speed, atmosphere and iteration are key, generative tools offer a compelling and adaptable alternative. Importantly, once a visual identity has been established, subsequent generations become progressively faster, making these processes especially effective in developing a coherent series of images with minimal manual adjustment.

In conclusion, the integration of diffusion model-powered workflows has not only allowed us to visualise design proposals more efficiently, but also to question and refine how we think through and communicate architecture. These tools, when applied critically and strategically, open a broader architectural field, one that prioritises dialogue, exploration and responsiveness over fixed representation. In a professional context defined by growing constraints on time and resources, this capacity for adaptive design support is not just valuable, but possibly necessary.

	Phase Styling	Collage Technique	Canvas Storm	Retouche Technique	Communication Styling
Software	Stable Diffusion	Stable Diffusion, Photoshop Generative Fill	Stable Diffusion	Photoshop Generative Fill, Stable Diffusion	Stable Diffusion
Duration	Short	Medium	Short	Medium	Short
Use Case	Converting to sketches in sketch phase, illustrations for concepts or storytelling, realistic images for proposals	Collaging qualitative parts of images and reprocessing them into a singular image	Exploring different design options and curating specific desired images	Refining specific elements in exploratory phases to communicate ideas or concepts	Refining output before meeting with different stakeholders in a project, based on what material they would want to see
Input Type	Images to be styleswapped	Generated images, reference images	Textual prompts, images	Generated images	Images to be styleswapped
Man-Machine Ratio	10-90	70-30	50-50	50-50	10-90
Limitations	Dependent on available presets or LoRA, preset development takes time	Needs large amount to imagine qualitative collages	Technical knowledge is required to produce at least some desired results within a larger batch	Additional step in the design process	Dependent on available presets or LoRA, preset development takes time

Table 18. Comparative summary of all developed and identified workflows.

Floorplan Extraction	Storytelling Technique	3D-to-Presentation	Correction Technique	Communication Generation
Stable Diffusion	ChatGPT	Stable Diffusion, Sora	Photoshop Generative Fill, Stable Diffusion	Sora, Stable Diffusion
Medium	Medium	Medium	Short	Short
Analysing qualities visible or suggested in a generated image and translating this into usable floor plans	Using textual descriptions together with generated images to convince clients	Quickly generating images for final presentations based on scenes exported from an earlier developed 3D model	Refining specific elements in the final phases to visualise the project	Quickly generating images while in dialogue with others to have clarity in design ideas
Prompting images with a clear goal	Text, sketches, web links, images	Textual prompting	Generated images	Textual prompting
70-30	70-30	60-40	50-50	70-30
Needs an experienced and educated reviewer to interpret images	Based on qualitative and extensive input	Requires a 3D model or a portion of it	Additional step in finalising the design process	Accuracy makes way for speed

6.2 Diffusion Models In Practice

6.2.1 Architectural Office

In late November, we transitioned briefly from exploration mode to real-world testing, delivering a one-day workshop at a well-established Ghent-based architecture office whose partners had expressed interest in generative AI. The agenda centred around Stable Diffusion: how to install it, select checkpoints, and navigate the vast number of parameters. Although the firm's culture is notably open-minded, with even some partners and project architects in attendance, the session revealed a substantial gap between curiosity and operational readiness. Stable Diffusion's interface, command-line technicalities and GPU dependencies felt alien to their habitual BIM-to-Enscape workflow.

A follow-up visit some months later expanded on the first workshop. One architect requested assistance in visualising a precise construction detail extracted from a detailed 3D model. Two obstacles surfaced immediately: first, the task itself: high-resolution, deterministic rendering. This is not the best use case of Stable Diffusion's expertise in atmospheric image synthesis. Second, running the model on an untuned office workstation produced mediocre results, and there was no time to replicate the custom settings that had taken us four months to refine. That is why we gave them the pragmatic advice: "Render it in Enscape; reserve AI for early ideation." This underscored a vital consideration: steep learning curves collide with the time-poverty of practice. Even the architect who left the workshop with the most enthusiasm later admitted he had not found the time to explore further.

To illustrate a more tangible application, we later collaborated with the same firm on a live project; this time not in a workshop setting, but in direct service of design communication. The project involved a large-scale residential development combining new-build volumes with the partial renovation of an existing site, including a former chapel and the sports hall beneath it. The property developer, experienced in new construction but unfamiliar with heritage renovation, was unsure how to approach these spaces. To support the architectural team in their discussions with the client, we generated a series of AI-driven visualisations using Stable Diffusion's image-to-image pipeline (Figure 49). The process began with cleaned base photographs (columns A and B), followed by successive rounds of generation that reimagined the interiors as playful, atmospheric, and programmatically diverse spaces (columns C–E). These speculative images, though not meant as final design proposals, proved instrumental as a low-threshold communication tool, helping the client quickly grasp the architectural potential of rooms they had otherwise overlooked. proposals, proved instrumental as a low-threshold communication tool, helping the client quickly grasp the architectural potential of rooms they had otherwise overlooked.

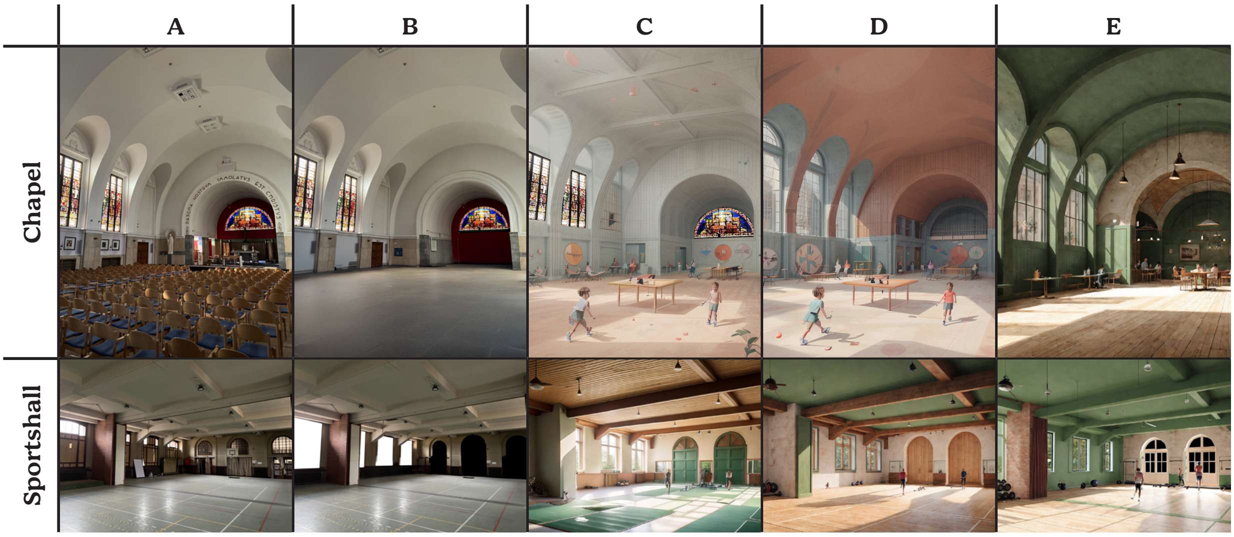


Figure 52. Process of generating visuals for an architectural office.

6.2.2 Estate Agency

As part of our exploration into a practical implementation, we also collaborated with a local estate agent to test the use of generative image techniques for visualising an unfinished apartment renovation. The goal was to offer potential buyers a clearer sense of the spatial and atmospheric qualities of the future interior. By moving from the delivered 2D plans to a 3D model with camera exports, we generated photorealistic interior images using ControlNet combined with LoRAs and prompting. This approach allowed us to deliver tailored visual content, before finalising the physical renovation. The quality of the output proved highly dependent on the level of detail and coherence in the 3D model, underlining the importance of accurate representation in these stages. Although the experiment was not further developed, the estate agent's positive response confirmed the potential of this method as a communication tool in this context.

Figure 53. Images created in dialogue with estate agent.

6.3 Discussion

Generative AI applications are expected to continue growing in prominence within the field of architecture, especially as new tools are developed daily to streamline and enhance design workflows. For instance, the emergence of models capable of converting multiple images into detailed 3D representations illustrates the rapid technological progression that supports increasingly efficient visualisation. This evolution aligns with the broader trajectory discussed earlier, in which digital design tools increasingly shape contemporary architectural practice.

At the same time, the need for ongoing research into the integration of generative AI in architectural design remains, and it arguably becomes even more relevant. As these tools advance, designers must navigate the balance between the time invested in learning new systems and the actual value they bring to the design process. A recurring challenge lies in managing constant updates in both software and hardware. Should one continually adapt to the newest tools or instead select a set of valuable applications and master those fully?

Reflecting on current explorations, it is also vital to acknowledge missed opportunities, both in specific design outcomes and in the broader implementation of generative AI technologies. Where could these tools have made a more meaningful impact, and where were they underutilised? Identifying gaps in the current market may further guide the development and refinement of these technologies, ensuring that their integration moves beyond novelty towards real-world architectural relevance.

7. Personal reflection

7.1 S. van Drie

Several months after first outlining my expectations and ambitions regarding the use of AI in architectural design, this reflection revisits the topic with new insights shaped by direct engagement. The core belief remains: AI should serve as a tool, not a replacement, within a creative process led by the architect. Still present is the idea that the actual value of design lies in conceptual thinking and authorship, not just production. However, my initial assumptions about ease of implementation and efficiency have now given way to a more critical understanding of how these tools influence processes, accuracy and authorship.

Beyond the first experience with a competition format itself, the entire project felt particularly refreshing. Engaging with generative tools opened new ways of thinking and creating, although the final result does not fully capture the richness of that development. This is, in part, intrinsic to the nature of such tools: they allow for rapid visualisation, endless iteration, and the ability to create custom references that respond to the designer's intent. However, this strength also presents a risk. Without clear boundaries, it becomes difficult to determine when enough has been generated and when the process starts to overshadow the design itself. There is a tendency to become absorbed in the production of images and ideas, potentially losing sight of the design's direction or spending excessive time without apparent progress. This highlights the need for critical engagement with the tools, ensuring they remain in service of the design, not the other way around. In our project, diffusion models played a significant role in early ideation, yet much of the design was ultimately shaped by hand. A human designer remained central for guiding the process, making decisions and refining outcomes. However, the tools affected our way of thinking more than the outcome itself, in a way that is subtle and difficult to represent.

Nevertheless, I think the value of AI in design is undeniable. While it seems that, at this moment, its use is entirely optional, it presents a worthwhile opportunity for designers willing to experiment, reflect and grow. Our project revealed both the potential and the current limitations of AI tools to me. This suggests that generative models may not yet be well-suited to the complex and nonlinear nature of architectural design, which often employs an intuitive approach. AI excels at linear, goal-driven tasks; yet, architecture usually relies on ambiguous thinking, gradual synthesis, and the nuances of human judgment. Moreover, while these tools can be integrated relatively easily in an academic context, their use in professional settings involves higher standards, greater responsibilities, and more constraints.

I think the research would have benefited from a longer timeframe, allowing continued experimentation and deeper reflection on the role of these tools. Still, within the scope of this project, the impact was significant. The integration of diffusion models prompted us to evaluate our work in new ways, with the images serving as moments of pause, reflection or even renewed inspiration. While I was not surprised that most AI-generated content required adaptation or reinterpretation, I was neither surprised nor disappointed. During my five years of architectural education in Antwerp, I developed a strong foundation in both conceptual thinking and practical design, gained through studio work, research, and hands-on projects. I have always been drawn to methods that combine creativity with reflection. I consider myself a context-aware and detail-oriented designer, curious about new approaches but grounded in architectural reasoning. The integration of generative AI into this project aligns with that mindset: it is not a replacement for design, but a prompt to think differently. I look forward to continuing this exploration in future work.

I want to take this opportunity to thank my project partner, Thomas, whose perspective challenged my assumptions and broadened my understanding. His contributions played a vital role in shaping the direction and outcome of this research.

7.2 T. Nelissen

My year with generative AI taught me as much about myself as it did about any software. The most valuable outcome is a clearer sense of how I want to learn, collaborate, and position my work, rather than the specifics of diffusion models, more so the habits and relationships that will support lifelong practice.

I discovered I am both energised and overwhelmed by steep learning curves. Diving into new tools lights a fire under my curiosity, yet I can easily slip into “tunnel-vision tinkering,” losing sight of bigger questions. Recognising that tendency will shape how I structure future projects: shorter technical sprints, deliberate checkpoints, and a bias toward low-friction platforms that allow design thinking to stay front and centre. For context, my creative approach has been an integral part of my identity for almost my entire life. Starting from the time I was 8 years old, I attended an art academy, which introduced me to architectural pre-education when I was 14 years old. I continued until graduating as a secondary school student at the age of 18. In my spare time, I enjoy reading philosophy and other fields where rational and creative thinking intersect. This attitude fuelled my motivation for this studio on generative AI. I wanted to explore the new and trending topic of AI. Reflecting on it now, I’m still delighted with this decision.

As critical as I am, I still find some elements to regret about this process. The single regret that keeps resurfacing is the feeling of partly having "operated in a vacuum." If I could rewind the last months, my first move would be to initiate an early partnership with an R&D-oriented office, such as MVRDV NEXT, or any peer group already deeply involved in architectural AI. Sitting down with them at the outset would have: revealed which research topics truly matter in the field; prevented dead-end detours on questions others had already answered; given me valuable feedback on translating technical novelty into architectural relevance. This is less about borrowing prestige and more about embedding myself in a live knowledge network. In the future, I will make "find the right collaborators fast" a non-negotiable project milestone.

Another personal takeaway is the tension between my analytical drive and my need for tactile, intuitive making. Long stretches of prompt-writing and parameter-tweaking left my creative instincts underfed. I now see that I thrive when digital exploration alternates with analogue reflection, whether that is sketching by hand or simply stepping away from the screen to think. Future workflows will schedule those breathing spaces by design, not by accident.

I have become more deliberate in how I communicate, process, and seek guidance. Early hesitancy, born of not wanting to "bother" mentors, cost me valuable feedback windows. I now accept that timely questions are a professional courtesy, not an imposition. This mindset shift will enable me to be a more proactive collaborator and, ultimately, a more reliable project partner.

First instinct in the next project: reach out to practitioners, researchers, and even sceptics before the brief feels frozen.
Personal metric of success: not the quantity of experiments, but the clarity with which each informs a design decision.
Ongoing discipline: protect time for both exploratory play and rigorous critique, recognising that I need both to stay creatively whole.

In conclusion, this research has sharpened my technical awareness, but more importantly, it has clarified how I learn, when to seek allies, and where I must guard space for intuition. Those lessons will guide every project I undertake from here.

I would like to thank my project partner, Sebastiaan, for helping me identify my blind spots. I do recognise that our collaboration was a significant factor in the success of our research.

8. Supplementary Materials

8.1 Acknowledgements

We would like to express our sincere gratitude to our promotors, Elien Vissers-Similon and Tine Segers, for their invaluable guidance, constructive feedback, and unwavering support throughout this research. Their expertise and commitment have played a crucial role in shaping both the content and direction of our research.
We are also thankful to all those who contributed to the realisation of this work, whether through academic input, technical support or by providing space for reflection and dialogue. This process has been both challenging and rewarding, and we sincerely appreciate everyone who helped us bring it to a successful end.

8.2 Bibliography

- (2024, February 26). Concrete SDXL & SD1.5 [LoRA model]. Civitai. https://civitai.com/models/322633/concrete-sdxl-and-sd15
- Adobe Inc. (n.d.). Adobe Photoshop. https://www.adobe.com/products/photoshop.html
- Agion. (n.d.). Praktische informatie voor scholen. Agentschap voor Infrastructuur in het Onderwijs. Retrieved June 5, 2025, from https://www.agion.be/praktische-informatie-voor-scholen
- Artesis Hogeschool Antwerpen. (n.d.). School in de stad, stad in de school. UPA–University Press Antwerp.
- artificialguybr. (2023, October 7). AnalogRedmondV2: Analog-style photography LoRA for SDXL [LoRA model]. Civitai. https://civitai.com/models/129689/analogredmond-analog-style-photography-lora-for-sd-xl
- baoanhnguyenkts784. (2023, July 18). Bartlett School [LoRA model]. Civitai. https://civitai.com/models/111396?modelVersionId=120162
- Cases. (2024, August 17). Archmat [LoRA model]. Civitai. https://civitai.com/models/656784/archmat
- De Vocht, S., & Van Sande, H. (Eds.). (2021). a+u architecture and urbanism: Marie-José Van Hee architecten (Vol. 613). A+U Publishing Co., Ltd.
- Designinput. (2023, September 11). LAN_Architecture_Visualization [LoRA model]. Civitai. https://civitai.com/models/143926?modelVersionId=159829
- epinikion. (2023, October 31). epiC35mm [LoRA model]. Civitai. https://civitai.com/models/184260?modelVersionId=206796
- Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems, 33, 6840–6851. https://arxiv.org/abs/2006.11239
- Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, L., Wang, W., Chen, W., Rajbhandari, S., Xu, Y., Zhang, T., & He, Y. (2021). LoRA: Low-rank adaptation of large language models. arXiv. https://arxiv.org/abs/2106.09685
- JerryQAQ. (2023, April 6). Architecture_illustrate [LoRA model]. Civitai. https://civitai.com/models/31673?modelVersionId=38113
- JZCG. (2023, August 11). Jzcg014-Sketch V1 [LoRA model]. Civitai. https://civitai.com/models/126830/jzcg014-sketch-v1
- klaabu. (2025, March 8). add_detail: Detail slider LoRA [LoRA model]. Civitai. https://civitai.com/models/1333749/add-detail-slider
- Laizhende. (2023, April 21). Zarch illustration [LoRA model]. Civitai. https://civitai.com/models/31479?modelVersionId=51144
- Mnemic. (2024, February 3). Cardboard Style [LoRA model]. Civitai. https://civitai.com/models/277115/cardboard-style-make-anything-out-of-cardboard-sd15-sdxl
- Mnemic. (2024, October 16). Paper Draft SDXL [LoRA model]. Civitai. https://civitai.com/models/814244?modelVersionId=910507
- OpenAI. (2022). ChatGPT [Large language model]. https://chat.openai.com/
- OpenAI. (2024). Sora. https://openai.com/sora
- Pallasmaa, J. (2012). The eyes of the skin: Architecture and the senses (3rd ed.).

John Wiley & Sons.
- Pezo, M., & von Ellrichshausen, S. (2022). El Croquis 214: Pezo von Ellrichshausen 2005–2022: From A to B. El Croquis.search.worldcat.org+29petermiller.com+29de.wikipedia.org+29rizzolibookstore.com+8ribabooks.com+8aabookshop.net+8
- ReplicantArmy. (2024, April 11). Paris1947 [LoRA model]. Civitai. https://civitai.com/models/395798?modelVersionId=441427
- Seleit, I. (2024). Handsketcher-XL [LoRA model]. Form-Finder. https://form-finder.squarespace.com/download-models/p/handsketcher-xl
- Seleit, I. (2024). Modelmaker-XL [LoRA model]. Form-Finder. https://form-finder.squarespace.com/download-models/p/modelmaker-xl
- Seleit, I. (2024). Petals-XL [LoRA model]. Form-Finder. https://form-finder.squarespace.com/download-models/p/du12omdou5bar0g7svbeyr5x55naxj
- Seleit, I. (2024, May 18). Nunu-XL [LoRA model]. Civitai. https://civitai.com/models/461798/nunu-xl
- Stability AI. (n.d.). Stable Diffusion. https://stability.ai/stable-diffusion
- wuzhebushihaoren501. (2023, May 17). CWG_archi_sketch_v1 [LoRA model]. Civitai. https://civitai.com/models/68286/cwgarchisketchv1
- XSarchitectural. (2023, March 28). XSarchitectural-7Modern interior [LoRA model]. Civitai. https://civitai.com/models/25383/xsarchitectural-7modern-interior

8.3 Appendix

To ensure easy access and version control of all relevant supplementary materials, we have chosen to provide a publicly accessible GitHub repository instead of including a traditional appendix in this document. This repository contains all figures, references, tools, models, and other resources that would typically be consulted in the appendix. The repository can be accessed via the following link:
https://github.com/S-van-3/Drawing-from-Diffusion.

8.4 List of Figures

8.5 List of Tables

Drawing from Diffusion

The role of image generation amid the rise of AI in Architecture

Thomas Nelissen
Sebastiaan van Drie

Promotors:
Elien Vissers-Similon, Tine Segers