

AI into the canonical design process :

diffusion models in architecture
and their impact in the creativity
workflow



Master of Science Program in
ARCHITECTURE FOR SUSTAINABILITY
a.a. 2024/2025
Graduation Session February 2025

DEGREE THESIS

CANDIDATES

MILITO Mattia (s314128)
CHIRAS Vladut Petru (s311626)

SUPERVISOR

SPALLONE Roberta

CO-SUPERVISORS

BARIOGLIO Caterina
PUPI Enrico
BRUNO Edoardo
FORTICH Fredy

The aim of the thesis is not only to demonstrate the potential of these new tools but also to raise awareness of the topic.

For the writing of this thesis, **2327** processes were served using AI. These roughly correspond to **4.39 kWh**. That corresponds to an emission of **1.02 kg Co2**.
These calculations are approximations

Index

1 Introduction

1.1 Preface	
1.1.1 The advent of AI in the architectural field	16
1.2 Thesis Structure and Methodology	18
1.3 Acknowledgements	20

2 Evolution of the relationship between architecture and technology

2.1 Architecture and Technology	
2.1.1 Historical Milestone	28
2.1.2 AI Milestone	34
2.2 Evolution of generative AI technologies	
2.2.1 Evolution of Large Scale Models	36
2.2.2 Large Visual Models (LVMs)	38
2.3 Most relevant AI concepts	
2.3.1 Understanding the machine	40
2.3.2 What AI is able to do ?	44
2.4 Understanding the generative design	
2.4.1 The prompt	46
2.4.2 CIDI method aiming for efficiency	48
2.4.3 Image Generation	50
2.4.4 Style Transfer	52
2.5 Sustainability and efficiency	54

3 The Artificial creativity a controvertial intelligence

3.1 The Role of AI in Architectural Creativity	
3.1.1 What defines creativity ?	60
3.1.2 The concept of GENIUS vs polymath's perceptions and capabilities	62
3.1.3 Strong AI and Lack of General Cognitive Abilities	66
3.1.4 The Use of the Tool: The Duality of Efficiency and Effectivness	70
3.1.5 The Value of a New Way of Expressing	72

4 AI in the process: Different phases

4.1 The Architect Creative Journey	
4.1.1 Project Planning and phases	76
4.1.2 The Initial Stages of Architectural Design	80

5 AI in the process: Different experiences

5.1 The candidates	
5.1.1 MVRDV & UNStudio	84
5.1.2 The Expert's Profile	86
5.2 Learning from the Field	
5.2.1 Methodology	92
5.2.2 The Interview Questions	94
5.2.3 MVRDV Answering	96
5.2.4 UNStudio Answering	98
5.2.5 Pattern Recognition in Interviews, a Reflective Debate	100
5.2.6 Case Studies : Competition	104
5.2.7 Case Study : Living Trails	110

6 AI in the process: Different approaches

6.1 The Canonic Process	
6.1.1 The Dynamic Designing Path	120
6.1.2 The Identities within the Designing Path	128
6.2 AI Integration into the Process	
6.2.1 The Intuitions	132
6.2.2 Helper 01 : The Conceptualizer	134
6.2.3 Helper 02 : The Assistant	138
6.2.4 Shaper 01 : The Inversion	142
6.2.5 Shaper 02 : The Dialogue	146

7 The tools: capabilities and limitations

7.1 ComfyUI	
7.1.1 The tool	152
7.1.2 Interface	154
7.1.3 Main Commands and Nodes	156
7.2 KSampler "The Image Generator"	
7.2.1 The Core of Image Generation	158
7.3 ControlNet "The Guidelines"	
7.3.1 Workflow Explanation	160
7.3.2 Iterations	162
7.4 IPAdapter "The Style"	
7.4.1 General Understanding	166
7.4.2 Workflow Explanation	168
7.4.3 Weight and Weight Type	170
7.5 IPAdapter & ControlNet	
7.5.1 Workflow Explanation	172
7.5.2 Iterations	174
7.6 Segmentation "The Recogniser"	
7.6.1 General Understanding	178
7.6.2 Workflow Explanation	180
7.6.3 Iterations	182
7.7 IPAdapter, ControlNet & Segmentation	
7.7.1 Iterations	184
7.8 Upscaler "The Enhacer"	
7.8.1 Workflow and Iterations	186
7.9 2D to 3D: Transforming Images into Objects	
7.9.1 Tripo AI	188

8 01 - The Last Nuclear Bomb Memorial

8.1 The Competition	
8.1.1 Key dates and overview	192
8.1.2 Lowering of the Curve	194
8.1.3 Brief Analysis	196
8.2 The Design Concept	
8.2.1 The Prompt	198
8.2.2 Following the flow, an uncontrolled brainstom	200
8.2.3 Analyzing the Output	206
8.2.4 The Segmentation, a controlled brainstorm	208
8.2.5 The Concept Conclusions	210

8.3 AI Process - The Dialogue	
8.3.1 The Approach	212
8.3.2 Workflow Explanation	214
8.3.3 The Dialogue Conclusions	228
8.4 AI Process - The Inversion	
8.4.1 The Approach	230
8.4.2 The 3D Model	232
8.4.3 Workflow Explanation	234
8.4.4 The Production	242
8.4.5 The Inversion Conclusions	246

9 02 - Rethinking The Grattacielo della Regione Piemonte

9.1 Understanding the Skyscraper	
9.1.1 Precondition	254
9.1.2 Survey	256
9.1.3 The Site Analysis	258
9.1.4 Programme Analysis	260
9.1.5 Lowering of the Curve	264
9.2 The Design Concept	
9.2.1 Conceptualizer	266
9.3 Fine-Tuning Methods	
9.3.1 Different Types of Fine-Tuning	268
9.3.2 LoRA Training	270
9.3.3 The LoRA's	272
9.3.4 The Used Datasets	274
9.3.5 Image Description	278
9.3.6 Kohya - Dataset Preparation	280
9.3.7 Kohya - Parameters	282
9.3.8 LoRA's Iterations	286
9.4 AI Process - The Dialogue	
9.4.1 Customisation	288
9.4.2 Advanced Customisation	290
9.4.3 Generating the 3D Model	294
9.4.4 The Dialogue Conclusions	296
9.5 AI Process - The Inversion	
9.5.1 Massing Study	298
9.5.2 Step-Story	300
9.5.3 The Final Reference	302
9.5.4 Generating the 3D Model	304
9.5.5 The Production	306
9.5.6 The Inversion Conclusions	318

10 Conclusions

10.1 AI-Driven vs AI-Assisted	
10.1.1 AI in the canonic process ? How ?	326
10.2 Tameing the Creativity Flow with Image Generation	
10.2.1 What is the impact on the crativity process ?	328
10.3 The Role of AI in the Creative Process	
10.3.1 Between Potential and Limitations	330
10.3.2 Integration, Challenges, and Future Prospects	331
10.4 Beyond Image Generation	
10.4.1 Responsible Use and Sustainability	332

Bibliography	334
Sitography	338
Appendix - Interviews with Industry Professionals	340

Abstract

In the world of architecture, technological advancement has always led to major changes in the profession. Until a few decades ago, design was done by hand, whereas today we make use of complex software that supports us throughout all the design phases. Today we are entering the age of AI and our aim with this thesis is to test how this can influence the canonical creative process. The thesis therefore aims to integrate the use of tools based on generative AI, with a specific focus on Diffusion Models. The work therefore interrogates the design process: how it can be shortened to achieve high efficiency or how it can be radically reimagined to become more effective. This was also explored through interviews with professionals from MVRDV and UNStudio that led us to define some key concepts applicable that were then applied to two practical exercises.

Introduction

Bibliography

- // As, Imdat, and Prithwish Basu, eds. The Routledge companion to artificial intelligence in architecture. Abingdon, UK: Routledge, 2021.
- // Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun. Deep Residual Learning for Image Recognition. 2015
- // Brookshear, J. Glenn, and Dennis Brylow. Computer Science : An Overview. 12th edition, Pearson, 2015.

Preface

The advent of AI in the architectural field

The integration of AI into architecture considerably changes several key aspects of how we design, create and solve problems. Once reliant on human intuition, artistic skill and technical knowledge, architecture now integrates machine intelligence and computational design. This transformation fundamentally changes how architects use technology, going beyond simple efficiency and automation.

This thesis explores AI's application in architecture, considering its potential advantages, disadvantages and larger effects. This research investigates how several AI tools are changing the field of architecture. It also focuses on generative design along with the ways leading firms are using this technology. It analyzes different real-world case studies, expert interviews and experimental applications and this analysis seeks to distinguish between AI as an assistive tool and AI as a potential co-creator in the design process.

This work explores the meaningful effect of AI on architectural creativity, an important area of understanding. AI shows important expertise in image generation, iterative design processes and workflow optimization; however, its capacity for conceptual or intuitive design thinking remains considerably undetermined. Will AI ever match the deep contextual understanding and intuitive problem-solving skills architects gain through years of experience? Will it remain a considerably advanced yet deeply constrained tool?

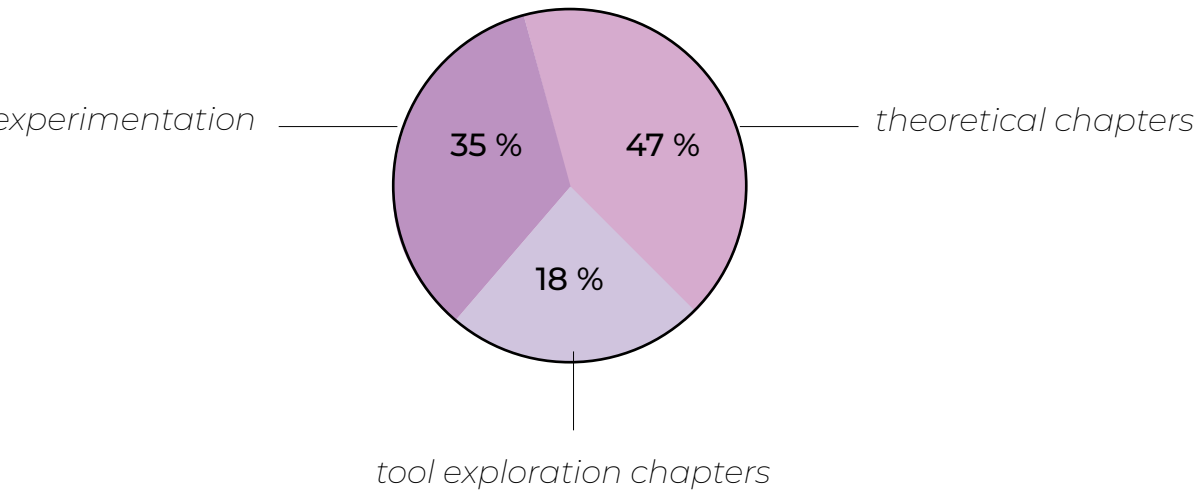
This thesis explicitly acknowledges the natural complexity of these multidimensional questions rather than providing a simplistic or singular answer. Mindfully acknowledging both the meaningful enthusiasm and legitimate concerns within architecture, the statement represents AI not as a finished product, but as an evolving force that will continue to substantially effect the profession. Architects face a technical turning point and they must consider the effective use of AI while preserving the value of human design.

This research explores AI's capability to reshape architecture and how architects can direct AI's development to adjust with their moral and original principles, going beyond simply documenting AI's current capabilities.

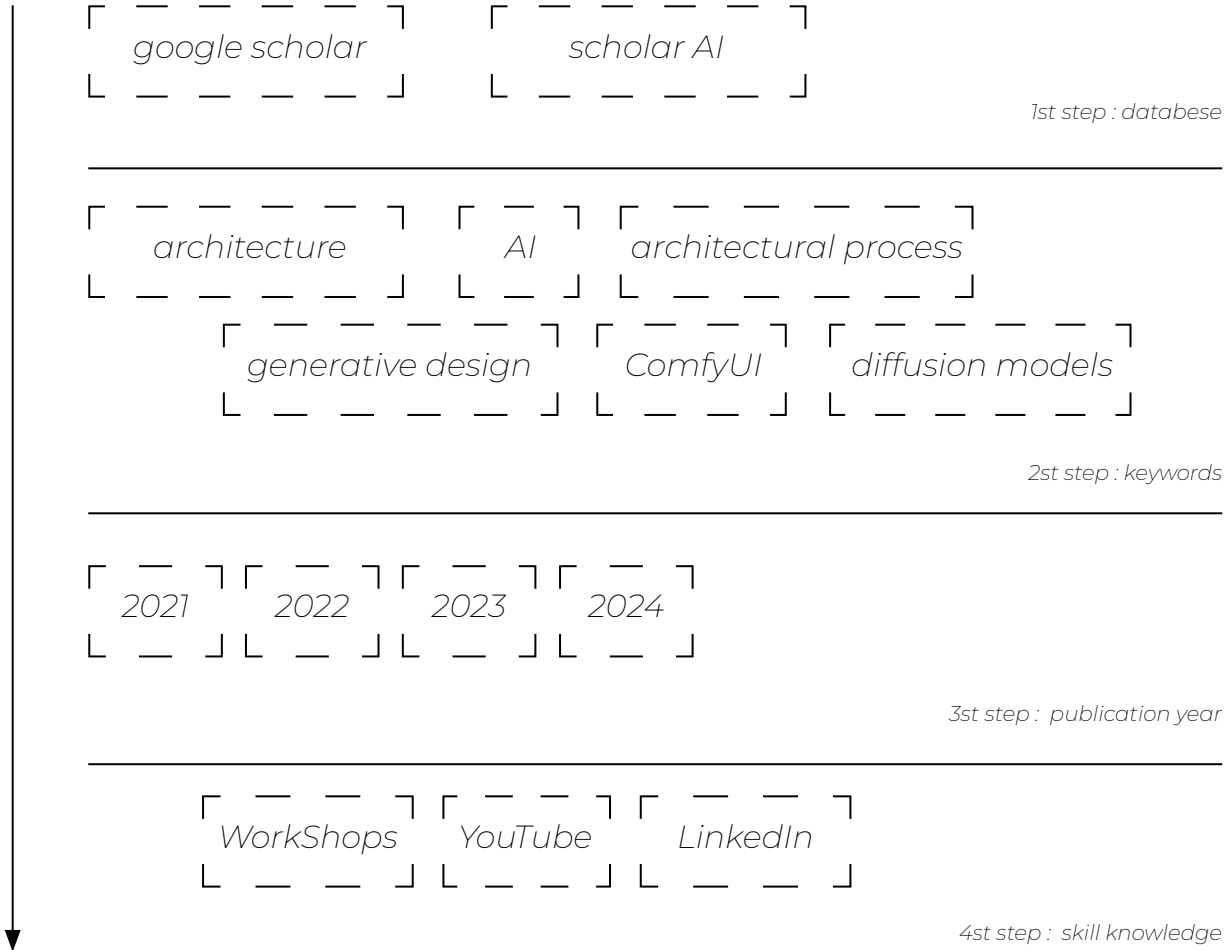
Thesis Structure and Methodology

The main purpose of the thesis is to explore how AI tools can be integrated in the architectural process. This thesis first of all explores how AI came to be today by analyzing the main historical events that have characterized the technological evolution, an in-depth study is done on how AI can be integrated into the architectural process nowadays, all this is also explored within prestigious architectural firms such as MVRDV and UNSTUDIO where we work directly.

It is mainly analyzed how Generative Design approaches the world of architecture and in particular the main tools are analyzed, especially ComfyUI, a powerful tool that allows maximum freedom and customization, it makes use of diffusion models that allow the generation of images through a denoise process, it will also be seen how to create your own models, today one of the big gaps is the lack of models that are optimized for architecture, but this is an incredible opportunity of personalization and learning how to influence AI systems.



01 // conceptual scheme of thesis structure



02 // research and thesis process

Acknowledgements

We're incredibly grateful to everyone who helped make this thesis possible.

Our team of supervisors deserves special thanks for their priceless guidance and they also deserve recognition for their large understanding throughout the entire process.

We further express our deep gratitude to many exceptionally skilled professionals from MVRDV and UNSTUDIO for generously sharing their wide-ranging expertise and perceptive perspectives, providing us with a considerably improved comprehension of AI's role in the architectural process, its natural limitations and its large potential.

We sincerely thank those who generously shared their expertise and perceptive advice, considerably contributing to this research's knowledge base.

A tutti coloro che mi hanno accompagnato nel percorso che mi ha portato a questa tesi.

Ai **miei genitori**, il cui amore, supporto e sacrificio sono stati la mia forza in ogni momento. Grazie per aver creduto in me, per avermi sempre incoraggiato a perseguire i miei sogni e per avermi donato un ambiente di affetto e serenità che mi ha permesso di crescere e affrontare ogni sfida.

A me stesso, per la determinazione e la perseveranza che mi hanno permesso di superare le difficoltà e raggiungere questo traguardo.

A tutti voi.

Vladut Petru Chiras

A chi c'è sempre stato.

Alla mia famiglia, che mi ha supportato e sopportato, rialzandomi quando ne avevo bisogno e sostenendomi in ogni passo di questo cammino.

A chi è rimasto.

Alle amicizie interminabili e vere.

A chi se n'è andato.

Alle amicizie e agli amori che, anche se lontani, mi hanno fatto crescere e mi hanno aiutato a raggiungere questo traguardo.

Grazie

Mattia Milito



The material is used as a reference from a stated source.



The material is a reinterpretation of an existing image from a stated source.



The material is made from human hands.



The material is made through the use of AI tools.

Why are we talking about AI ?
How did we get to AI ?
What AI is able to do ?

Evolution of the relationship between architecture and technology

Sitography

<https://rarehistoricalphotos.com/life-before-autocad-1950-1980/>
<https://ourworldindata.org/brief-history-of-ai>
<https://www.technologyreview.com/2018/12/19/138508/mighty-mouse/>
<https://www.oracle.com/it/artificial-intelligence/what-is-ai/ai-vs-machine-learning/>
<https://www.archdaily.com/785967/the-edge-plp-architecture>
<https://www.ladybug.tools>

Bibliography

- // Mario Carpo. A short but believable history of the digital turn in architecture, 2023.
// Bernstein Phil. Machine Learning: Architecture in the age of Artificial Intelligence , 2022.
// Jiaqi Wang, Zhengliang Liu, Lin Zhao, Zihao Wu, Chong Ma, Sigang Yu, Haixing Dai, Qiusi Yang, Yiheng Liu, Songyao Zhang, Enze Shi, Yi Pan, Tuo Zhang, Dajiang Zhu, Xiang Li, Xi Jiang, Bao Ge, Yixuan Yuan, Dinggang Shen, Tianming Liu, Shu Zhang. Review of large vision models and visual prompt engineering, Meta Radiology, Volume 1, Issue 3, 2023.
// Chaillou, Stanislas. Artificial intelligence and architecture: from research to practice. Birkhäuser, 2022.
// Hetrick, Suzanne H., Roy S. Lilly, and Philip R. Merrifield. "Figural creativity, intelligence, and personality in children." Multivariate Behavioral Research 3.2 (1968): 173-187.
// Brownlee, Jason. Generative adversarial networks with python: deep learning generative models for image synthesis and image translation. Machine Learning Mastery, 2019.
// Li, Pengzhi, Baijuan Li, and Zhiheng Li. "Sketch-to-architecture: Generative ai-aided architectural design." arXiv preprint arXiv:2403.20186 (2024).
// Buonamici, Francesco, et al. "Generative design: an explorative study." Computer-Aided Design and Applications 18.1 (2020): 144-155.
// Almaz, Amira Fawzy, et al. "The Future Role of Artificial Intelligence (AI) Design's Integration into Architectural and Interior Design Education is to Improve Efficiency, Sustainability, and Creativity." Sustainability, and Creativity 3.12 (2024): 1749-1772.
// Jalia, Aftab, Ron Bakker, and Michael Ramage. The edge, Amsterdam: showcasing an exemplary IoT building. Technical report, Centre for Digital Built Britain, University of Cambridge, UK, 2019.
// Caetano, I., & Leitão, A. (2019). Architecture meets computation: an overview of the evolution of computational design approaches in architecture. Architectural Science Review, 63(2), 165–174.
// Desislavov Georgiev, Radosvet. "Analysis of Deep Learning Inference Compute and Energy Consumption Trends." (2021), 11-13.
// Li, Chengyuan, et al. "Generative AI for Architectural Design: A Literature Review." (2024).
// Bernstein, Phil. Machine learning: architecture in the age of artificial intelligence. RIBA Publishing, 2022.
// Tan, L., & Luhrs, M. (2024). Using Generative AI Midjourney to enhance divergent and convergent thinking in an architect's creative design process. The Design Journal, 27(4), 677–699.
// Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

Architecture and Technology

Historical Milestone

Architectural design in human history has undergone many changes, just think of the methods used in antiquity when architects drew by hand on parchment with tools such as compasses, squares and rulers.

The Renaissance brought greater attention to proportions and architectural detail, with great artists such as Leonardo da Vinci and Michelangelo combining aesthetics with precision. With industrialisation, the architectural process began to require more standardisation to cope with large-scale requirements. In the 1980s, the first CAD tools such as AutoCad and CATIA were introduced, replacing manual drawing, marking a real revolution in the architectural world. Subsequently, it was thanks to software such as Rhino and Grasshopper (2007) that it was possible to reach a more advanced level of design thanks to the possibility of managing complex geometries and shapes.

The greatest technological developments can be considered within the industrial revolutions: the second industrial revolution introduced a more standardised architecture thanks to mass production and electricity, the third introduced a greater level of automation thanks to calculation systems and electronics, while the fourth integrates data and digital systems within the architectural process.

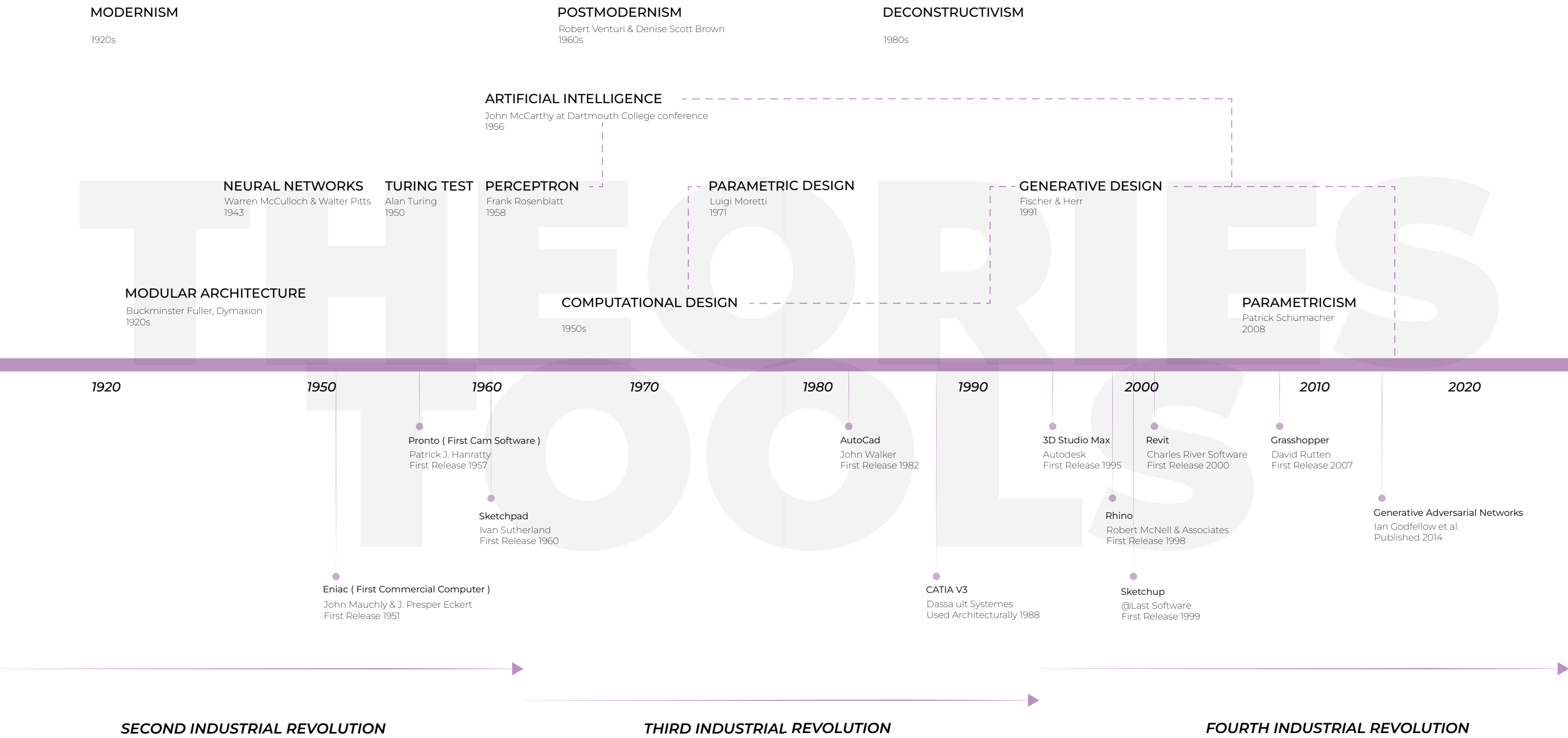
In recent years, the introduction of artificial intelligence has allowed designers to expand their points of view, which is a big change from the past where design was based solely on intuition and past experience.

These events have influenced architecture over the years by changing the way projects are conceived and realised, offering new possibilities.

“Architects and designers will have to learn to master them, if only because they will soon be the cheapest and most effective tools of their trade. If architects decide not to use them, others will.” (Carpo, 2023)¹.

¹Carpo emphasizes the inevitability of architects and designers adopting new technologies, particularly artificial intelligence, as they become the most cost-effective and efficient tools in the field (source : del Campo, Matias. Neural Architecture: Design and Artificial Intelligence. Applied Research and Design Publishing, 2022 : 12-15.)

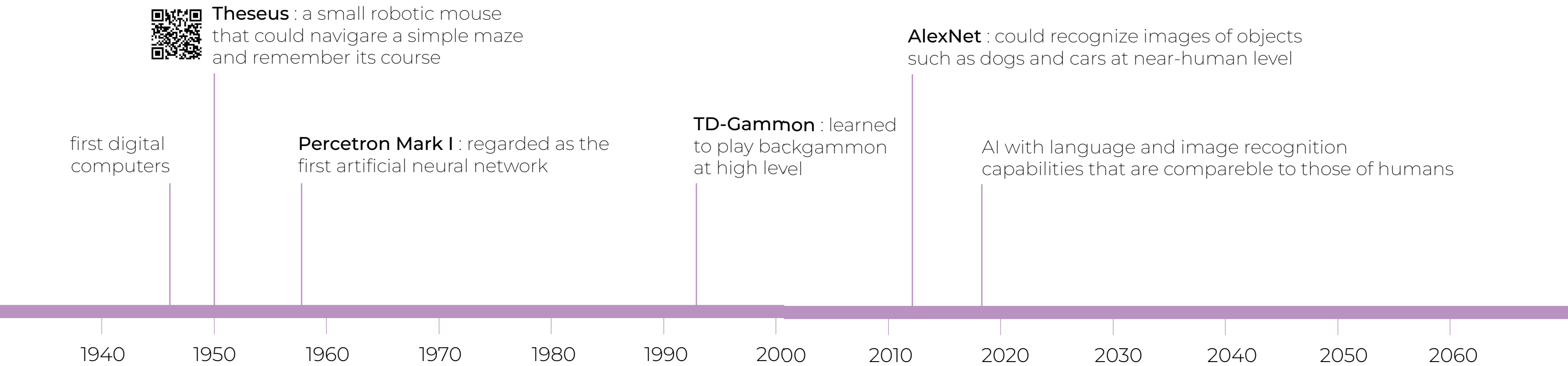
**“The computer is an extension of the design process, not a substitute for it.”
(Zaha Hadid)**



03 // evolution of architectural tools and related theories (source : Caetano, I., & Leitão, A. (2019). Architecture meets computation: an overview of the evolution of computational design approaches in architecture. Architectural Science Review, 63(2), 165-174)



AI Milestone



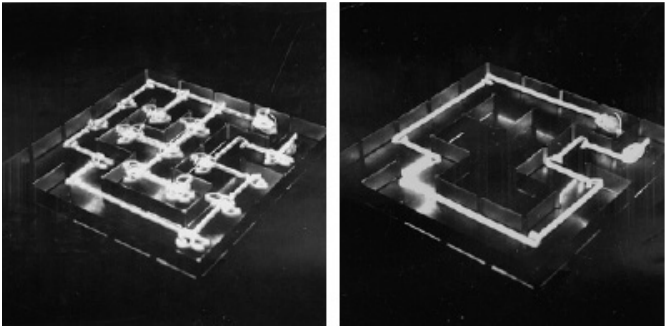
06 // main AI related projects during the years (source : <https://ourworldindata.org/brief-history-of-ai>)

The term artificial intelligence was attributed to computer scientist John McCarthy in 1956, suggesting how machines could somehow mimic the functions of the human brain.

The integration of AI was not an isolated event. Although it seems recent, it has much deeper roots; the first project to use a form of artificial intelligence dates back to 1950 by Claude Shannon, who developed Theseus, a small remote-controlled robot mouse that moves through a maze and then manages to memorise its path and repeat it.

The project was developed through a process of ‘learning’, as Shannon explains, i.e. the ability to analyse a quantity of data and memorise it and then be able to reformulate what has been learnt, the same theory on which even the most advanced processes of today are based.

Technology has changed and continues to change over the years, performing typically human actions at a remarkable speed, and this is thanks to the considerable computing power of computers that continues to grow.



07 // those photos, published in Life magazine in 1952, show the path Theseus took while learning a maze pattern and the direct path taken on its second trip through the same maze (source : <https://www.technologyreview.com/2018/12/19/138508/mighty-mouse/>)

Evolution of generative AI technologies

Evolution of Large Scale Models

Large-scale models today represent the beating heart of the most complex AI systems, e.g. large language models have been developed with huge amounts of data and resources used to train deep neural networks, which will be discussed in more detail later.

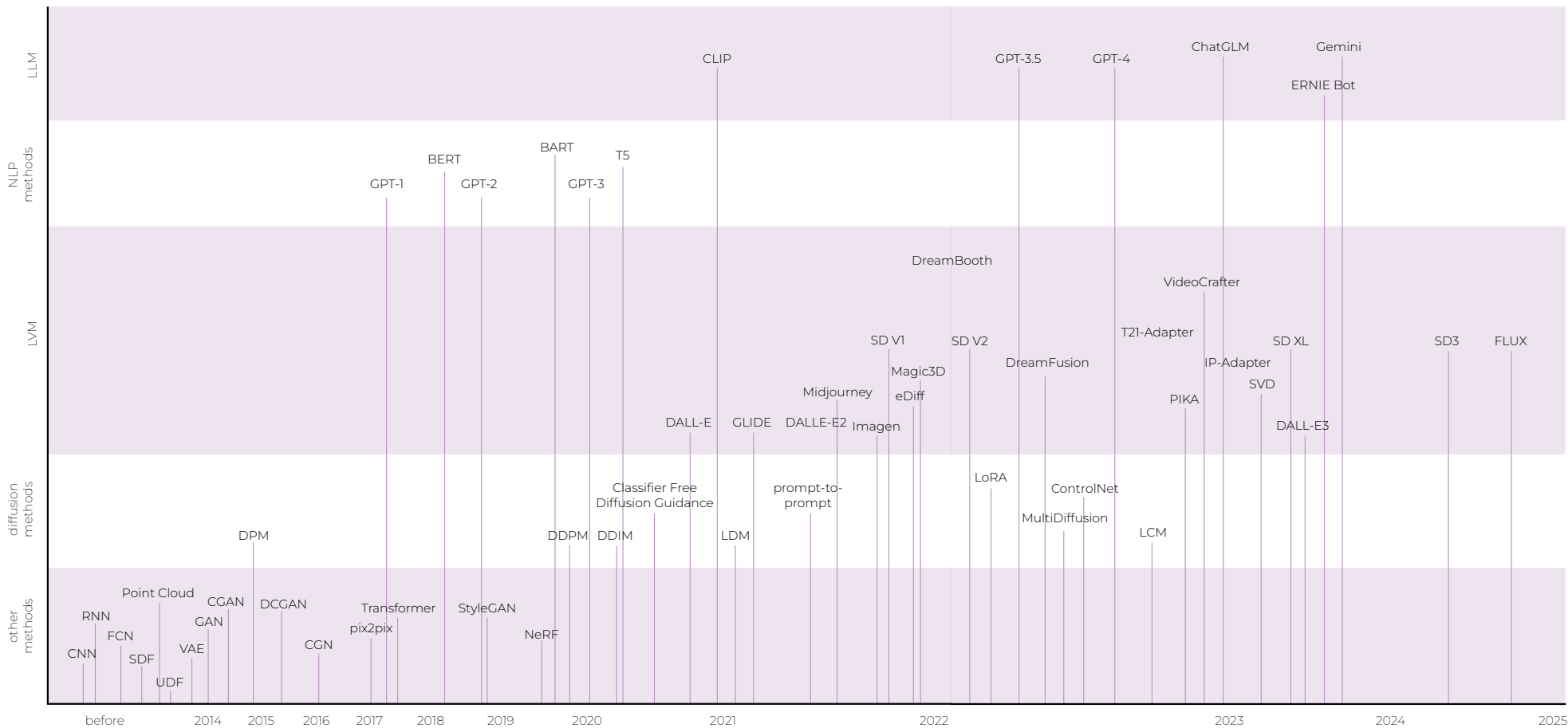
These models belong to various categories, such as natural language processing (NLP), diffusion methods or the large language models mentioned above.

In 2017, the Transformer² model was initially introduced in natural language processing (NLP), the BERT model is based on this technology and has produced outstanding results in the field.

Seeing the great opportunities of the Transformer model, developers started to train larger models such as GPT-1, GPT-2, GPT-3 or GPT-4, these models more advanced than their predecessors have in some cases exceeded human capabilities.

The Vision Transformer (ViT) enabled the integration of the transformer architecture within artificial visualisation, bridging the gap between textual and visual visualisation.

To be more specific, a transformer model uses the token as a unit, which makes it suitable for handling multimodal data as the embedding layers are able to convert any modality into tokens, which led to the development of models such as DALL-E or Stable Diffusion.




The graph opposite shows how artificial intelligence has accelerated its evolution in recent years. The diversification between the types of models also shows how the approach is geared towards developing the most effective algorithm capable of relating to the human world in surprising new ways.

²neural network architecture developed to process sequential data such as text, which is extremely important because it introduces the attention variable, which allows the model to weigh the importance of different parts of the input allowing it to process data differently than RNN models that process data in order Transformer models consist of the encoder, which processes the input data, and the decoder, which generates the output based on the input information

Large Visual Models (LVMs)

AI expert Mark Greaves³ argues that AI systems can be divided into four categories: evaluation, simulation, generation and understanding:

capability	explanation
understanding	being able to find, index and deploy data
evaluation	understanding the implications of data based on statistical analysis of very large data sets
simulation	projecting future states or conditions based on characteristics of past situations or constellations of data
generation	creating entirely new ideas or concepts based on knowledge and understanding of a given context

 09 // modified greave's model of AI capabilities (source : Bernstein, Phil. Machine learning: architecture in the age of artificial intelligence. RIBA Publishing, 2022, 18.)

Large Visual Models (LVMs) are able to interpret, process and generate visual content in a way that was unthinkable a few years ago. Their ability to analyse huge amounts of data and interpret its content has enabled their application in a variety of sectors in addition to architecture.

As mentioned earlier, this type of model is based on Transformer models capable of balancing the weight of incoming information. The CLIP (Contrastive Language Image Pre-Training) model allows the exceptional combination of text and images. These capabilities have enabled the development of more complex and powerful models such as the recent FLUX, which instead uses T5 prompting, allows the creation of high quality images that reflect the textual input.

Despite ongoing advancements, one of the biggest challenges AI faces is its ability to generalize to new and unpredictable contexts. When confronted with unfamiliar scenarios that deviate from its training data, AI tends to produce less accurate or unreliable results.

Additionally, the transparency of AI models, how and why they generate specific outputs, remains an active area of research. Understanding the internal decision-making process of these models is crucial to ensuring their reliability and trustworthiness, particularly in fields where accuracy and accountability are essential. One way to improve this is to adapt the model through a process of "Prompt Tuning". This method allows new learnable parameters to be added to the model without changing the parameters of the starting model, one of the most successful methods being VPT (Visual Prompt Tuning). Unlike a LoRA, which focuses on a more structural change of the model's weights by introducing low-rank matrices into the Transformer model that allow the acquisition of new information.

Although different, these methods have the same goal of improving the output generated by the model or filling certain data gaps.

In 2023, Meta AI introduced SAM (Segment Anything Model), an image segmentation model based on prompt engineering that enabled greater image comprehension and higher quality in the final output.

**“Artificial Intelligence however, should have the ability to come up with ideas or artefacts that are new, surprising and valuable.”
(Boden, 2004)⁴**

³a senior technical leader in artificial intelligence and data analytics at Pacific Northwest National Laboratory (PNNL), has contributed significantly to the field of AI
⁴source : Boden, Margaret A. The creative mind: Myths and mechanisms. Routledge, 2004 : 108

Most relevant AI concepts

Understanding the machine

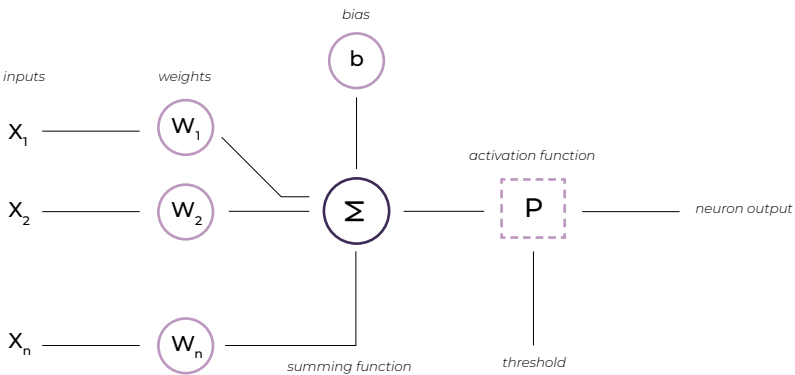
“Machine learning, a branch of artificial intelligence, is the study of how machines can be taught to recognize data, categorize data, predict data and even generate data ‘without being explicitly programmed”
(Simon, 2015)⁵.

Machine Learning relies on a large amount of data to define a mathematical representation of the relationships within the data, which is then used to predict new data.

These models are used in many different fields, such as Facebook’s algorithms that use a similar system to suggest personalised content to users, or the example of credit cards where an AI algorithm can be trained with numerous if/then rules in order to detect fraud, and based on the results it can develop a pattern such as purchases at unusual times or from unusual geographical areas.

One of the most interesting aspects of machine learning is its ability to improve over time as the amount of input data increases and computers become more powerful.

This model works through a system of neural networks that attempt to simulate the functioning of the human brain, creating a structure of artificial neurons organised in layers in which each neuron transmits information to the next, creating an interconnected network of information.

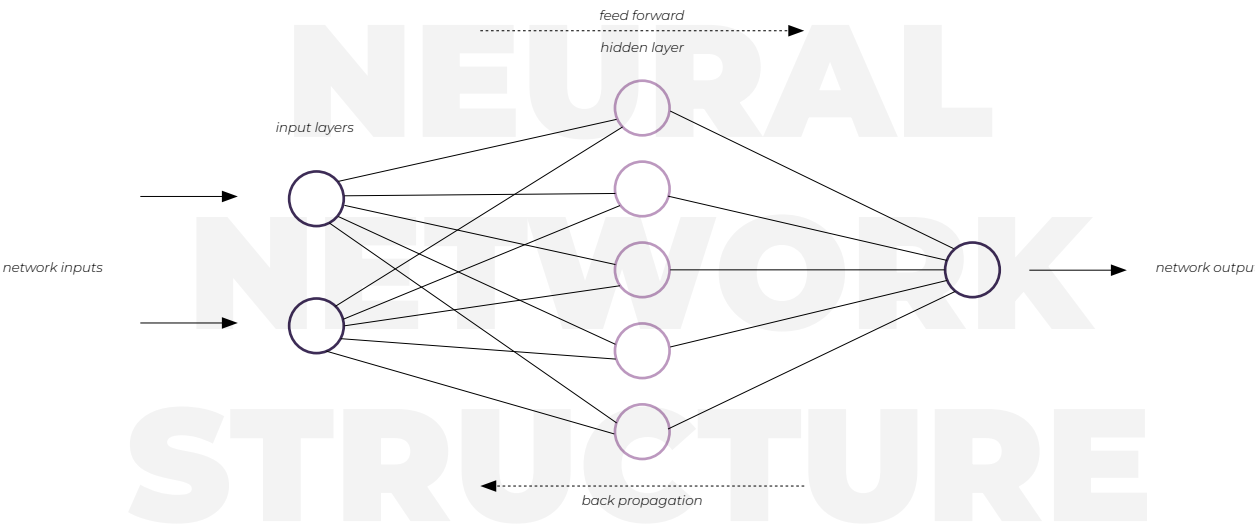


10 // artificial neuron structure

⁵source : Simon, Phil. Too Big to Ignore: The Business Case for Big Data. Vol. 89. Wiley, 2013

Each of these neurons represents a mathematical function that takes on a specific weight and bias, the inputs are summed together and then multiplied by a specific weight value at each connection between neurons, after which a bias is added to adjust the neuron’s output.

When a training input is provided to the neural network, the output is compared with the expected output and the weights are used to decrease the discrepancy through a process of ‘backpropagation’.



11 // simply fully connected neural network

It is difficult to fully understand how a neural network arrives at a certain conclusion; we know the input given and the output generated, but not the internal processes that take place between the various interconnected neurons.

A human being is able to follow a path of logical thought that led him to a certain observation or conclusion, but it is problematic to have the same development for neural networks.

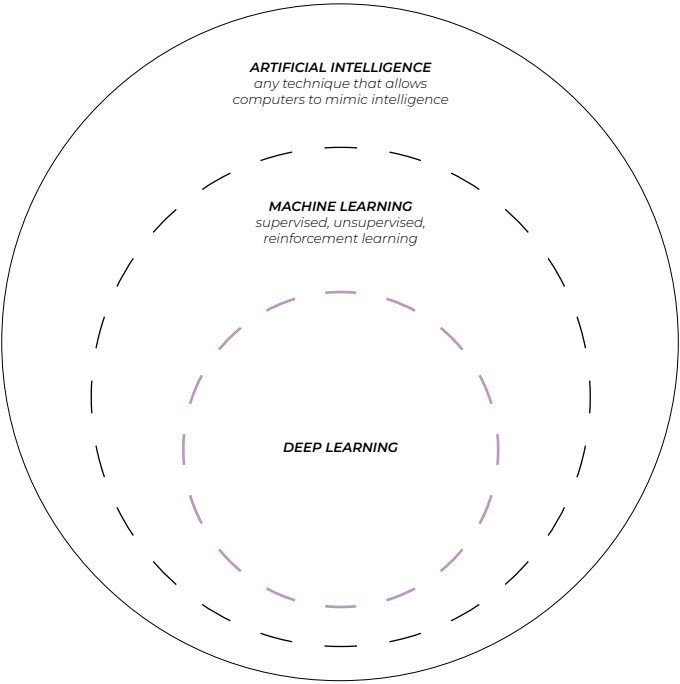
For some sectors this may not be a problem, but for environments such as medicine where understanding reasoning and everything that leads to a given output is crucial.

Because of this lack of transparency, this aspect of neural networks is referred to as the ‘black box’, to emphasise the difficulty in understanding their decision-making processes.



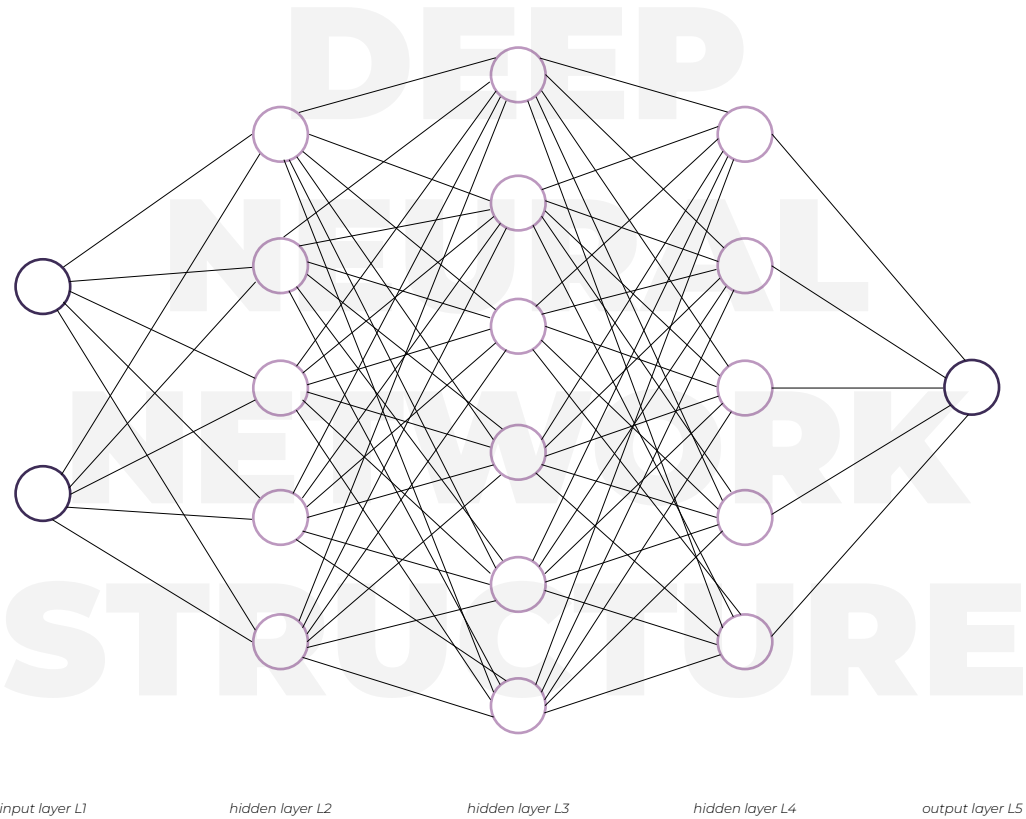
12 // black box

Furthermore, machine learning algorithms can be divided into several categories: in supervised learning, the algorithm is trained with labelled data, i.e. data known to the developer; in unsupervised learning, the data are not labelled and the algorithm tries to identify structures in the data; and in reinforcement learning, the algorithm learns through interaction, receiving positive or negative feedback based on its actions.



13 // artificial intelligence and his branches

Of the various techniques, Deep Learning differs in the use of deep neural networks, i.e. neural networks with multiple hidden layers, which allows the algorithm to operate with more complex data. In this type of algorithm, learning takes place automatically without having to intervene in the definition of parameters or characteristics; this ability to learn from raw data makes it extremely powerful in image recognition. Deep Learning models naturally require a greater amount of data as well as greater computational power.



14 // deep neural network (DNN)

What AI is able to do ?

AI appears to be the next big thing in architecture, the integration of which is redefining design processes, giving way to new visions and perspectives.

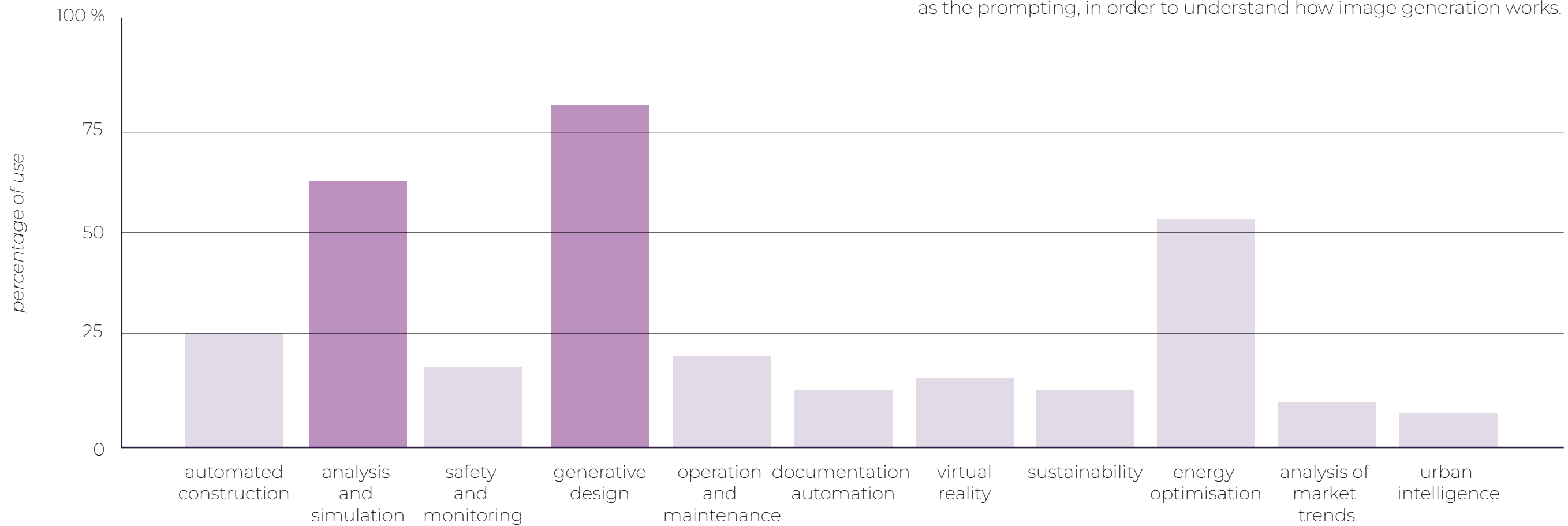
Of course, AI can be used for different topics as shown in the graph below, but we will mainly focus on generative design, where it is most used at the moment.

It is also used to make energy assessments and simulations that can help designers in the interpretation of design also from a climate point of view.

The Generative Design approach has become increasingly applied in many sectors besides architecture. It consists of the visual elaboration of plausible solutions to a problem or design idea.

Thanks to the enormous computing power and easy access through commercial apps available today, it has become increasingly popular in recent years and has become part of today's workflows, inspiring many architects and designers in defining their concepts.

GenAI encompasses various techniques, processes and methodologies, and it is important to delve into some of its fundamental elements such as the prompting, in order to understand how image generation works.



15 // main architectural fields where AI can be integrated (source : Yildirim, Erdem. "Text-to-image generation AI in architecture." Art and architecture: theory, practice and experience, 2022, 97-101.)

Understanding the generative design

The prompt

Prompt engineering is the process of structuring an instruction that can be interpreted and understood by a generative artificial intelligence model. A prompt is a natural language format that is then encoded into token or vector that then AI can uses to interpret and generate a most probable output based on this "instruction".

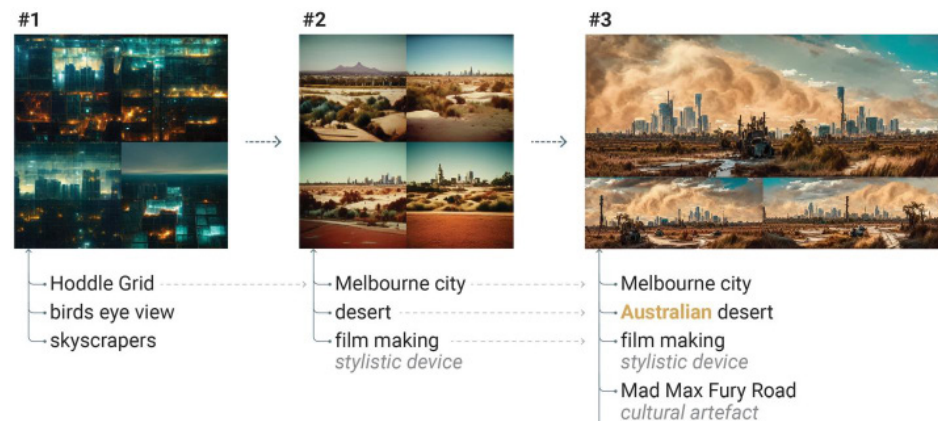
To understand how AI can be used to enhance or expand an architect's creativity, one must understand the difference between visual creativity and creative cognition.

Visual creativity (Hetrick, Lilly, and Merrifield 1968)⁶ is defined as the creation of new visual forms such as drawings and paintings, while creative thinking can be understood as the unconscious process of processing information to obtain new insights and ideas.

The research conducted by Linus Tan & Max Luhrs (13 May 2024) focuses on studying how GenAI influences architects' creative cognition and how it is useful in enhancing their creative process.

Several GenAI softwares are available on the market such as: Midjourney, Dall-E, Stable Diffusion and ComfyUI, the study below is based on the use of Midjourney.

The study attempts to create a vision of a future Melbourne through keywords or additional imagery for reference, words such as ‘Bird’s eye view’ ‘Melbourne city’ ‘Desert’ were used.



¹⁶ // influence of prompt on the output (source : Tan, L., & Luhrs, M. (2024). Using Generative AI Midjourney to enhance divergent and convergent thinking in an architect's creative design process. The Design Journal, 27(4), 677–699.)

⁶source : Hetrick, Suzanne H., Roy S. Lilly, and Philip R. Merrifield. "Figural creativity, intelligence, and personality in children." Multivariate Behavioral Research 3.2 (1968): 173-187.

The use of certain words implies a certain style or direction. In the following table, the words are divided into topics, indicating what you can control by using them and what part of the image they affect.

SUBJECT	ENVIRONMENT	LIGHTNING	COLOR	MOOD	COMPOSITION	STYLE
building floor plan sculpture structure landmark urban residential historical sustainable innovative	natural artificial chaotic tranquil vibrant pollused rural arid eclectic welcoming	soft harsh natural artificial dramatic directional warm cool diffused intense	vibrant bold subdued faded warm cool bright natural moody monochromatic	romantic joyful nostalgic serene mysterious melancholic dramatic tranquil vibrant surreal	balanced dynamic symmetrical asymmetrical harmonious strong creative abstract framed traditional	architect name modern classical minimalist brutalist gothic victorian futuristic contemporary ancient
FORM				FUNCTION		
scale	walls	openings	floors	lights	furniture	materials
large small big little huge tiny massive miniature enormous minuscule	solid study thin thick decorative separating transparent dividing interior exterior	windows doors arches pathway skylight double height balconies ventilation atrium court	two levels floor open plan divided spaces	romantic joyful nostalgic serene mysterious dramatic dreamy vibrant intimate peaceful	elegant modern rustic sleek vintage minimalist functional versatile contemporary	epoxy flooring rug and carpet vinyl porcelain natural stone concrete hardwood anti static floor tiles marble

¹⁷ // group of words and their category for prompting (source : Almaz, Amira Fawzy, et al. "The Future Role of Artificial Intelligence (AI) Design's Integration into Architectural and Interior Design Education is to Improve Efficiency, Sustainability, and Creativity." Sustainability, and Creativity 3.12 (2024): 1755.)

With the introduction of SD3 and the more recent Flux, more complex parameters and nodes have been included, and this also affects prompting.

Basically, there are two different methodologies for ‘talking’ to the model: CLIP (Contrastive Language-Image-Pre-Training) with a limit of 77 tokens, the same text encoder used in earlier models such as SD1.5 or SDXL, and T5 (Text-to-Text Transfer Transformer), a text encoder that can process up to 512 tokens, allowing for more detailed prompts.

The main difference between the two is that T5 allows the use of more descriptively structured sentences and phrases, whereas the classic CLIP makes use of concise words, all separated by commas.

CIDI method aiming for efficiency

The use of the CIDI⁷ method for building up a correct input can help a lot on building something that is modular and repetible.

This system aims to boost the efficiency of the work providing a clear base to follow and help to aim the efforts in the right direction.

This prompt engineering system is highly modular and scalable, offering a clear structure that colleagues and team members can easily follow.

Its power derives from archiving several different prompt inputs. It also reuses these inputs in many situations. A structured, layered approach to understanding, provided by context as well as instructions, the first two elements, guarantees flexibility in addition to adaptability.

⁷the CIDI framework, developed by Gianluca Mauro, as is a structured approach to prompt engineering based on four key elements: Context, Instructions, Details, and Input. This method helps optimize interactions with AI, ensuring more relevant and accurate responses (source : Wang, Simon, et al. "Seven principles for effectively partnering with Generative AI for teaching and learning." Available at SSRN 4950507, 2024.)

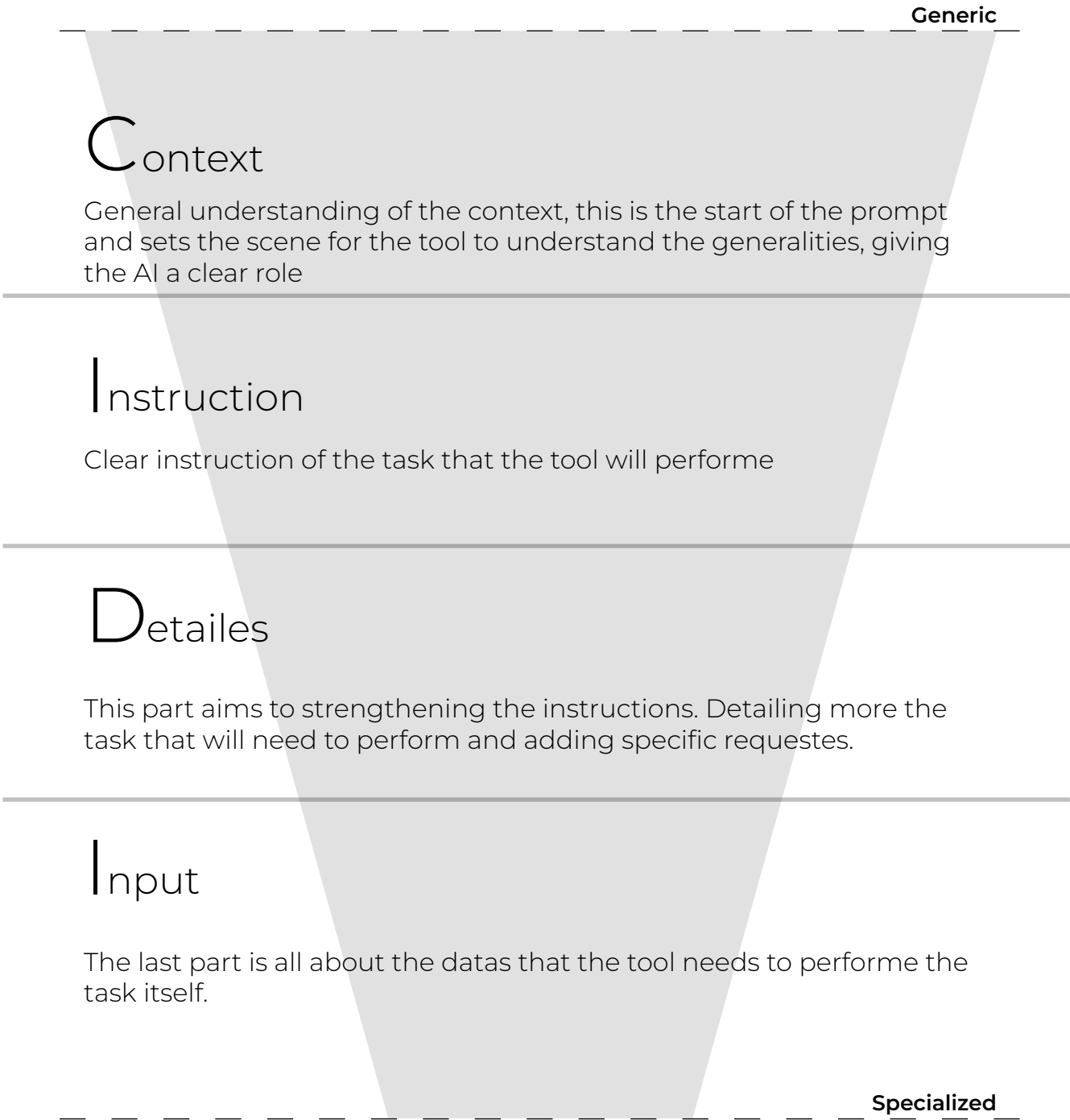


Image Generation

When we talk about image generation we enter a very broad field and we need to understand the differences between some of the main types.

First of all, what is image generation?

Image generation creates visual content using AI, often starting with random noise. Models like diffusion networks refine this noise step by step, revealing patterns and structures until a coherent image emerges. There are various ways of doing this: Text - to - Image, Image - to - Image and Inpainting.

"Text-to-Image Generation refers to computer approaches that can convert human-written textual descriptions, such as keywords or phrases, into visually portrayed concepts with the same semantic meaning as the text." (Erdem Yildirim, 2022)⁸

Text-to-image allows the creation of new images simply by using a textual description called a 'prompt' while image-to-image relies on a reference image to create a similar one.

Inpainting allows new images to be created or through the use of masks to change only parts of the image through drawings, sketches or colour combinations made directly by the user, for example the user might sketch a shape and through AI add details based on the input provided.

The image generation is an area that is in continuous development, thanks to the continuous increase in computer power, every day tools and techniques are studied and developed that allow greater control and greater quality of the image as well as the coherence of the image itself, at the same time as we speak new models and processes are being tested and are taking hold in the world of architecture by professionals in the sector, for this reason we must not stop at the first discovery but continue the research and study.

19 // on the right : different images created with AI

⁸Source : Yildirim, Erdem. "Text-to-image generation AI in architecture." Art and architecture: theory, practice and experience 97, 2022 : 7.



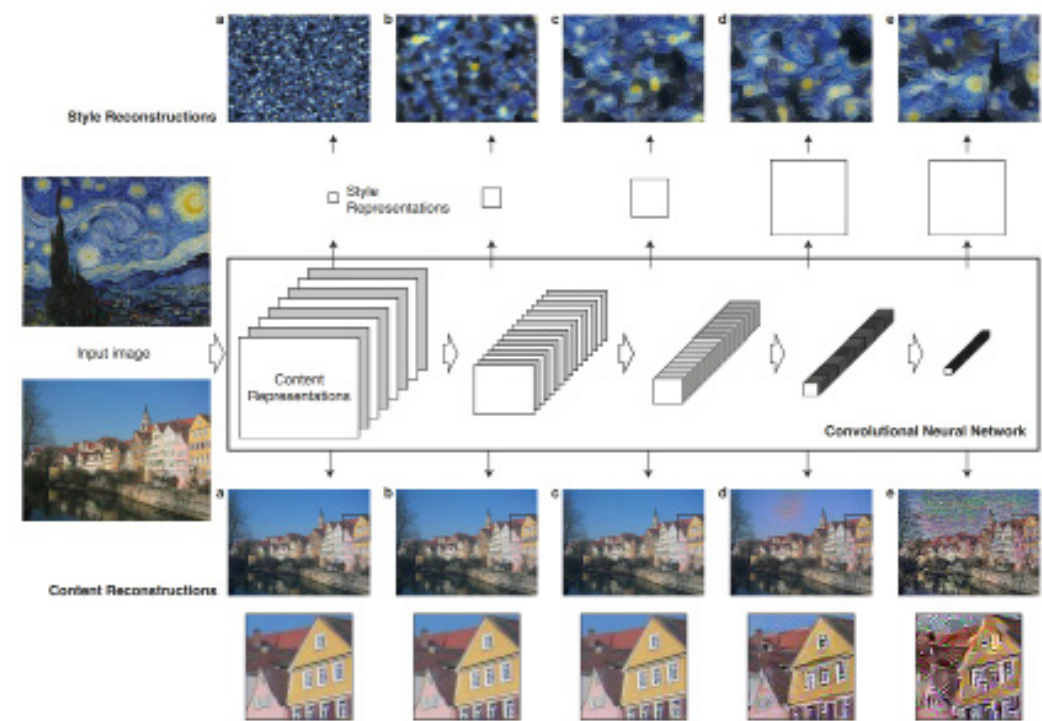
model : RealVisXL V4.0 Lightning
prompt : parametric architecture in a dense forest
steps : 5, cfg : 2, dpmpp_sde, karras

Style transfer

The work of Gatys et al. demonstrates how convolutional neural networks (CNNs) are able to create images by separating and recombining image style and content. This process is called Neural Style Transfer (NST), and the study conducted analyses the main advances and algorithms used in the NST world.

The paper proposed an algorithm capable of exploiting NST to reformulate the content of an image and the style of famous works of art.

The example below represents a style transfer of a famous painting onto a street photography.



To better understand how an NST works, we need to understand how to extract the image style data we want to use.

To do this we rely on Visual Texture Modelling, which is divided into two different approaches : Parametric Texture Modelling and Non-Parametric Texture Modelling.

The first method makes use of mathematical models or statistics to describe the texture, such as the difference in colours or shapes in a given area.

On the other hand, the second method is not based on mathematical models but makes direct use of pieces of texture from the original image, adapting them to the final image, thus making the transfer of style more detailed and varied.

Then the new information must be applied to the final output by iteratively modifying the new image to minimise the loss of content and loss of style between the new image and the reference image, this comparison can be calculated using the Gram matrix above for example.

During this process, the information extracted from the texture gradually replaces the pixels of the new image; it is a step-by-step process that allows the original content to be maintained with a certain degree of precision.

In the following chapters we will see processes and algorithms that allow us to work with style transfer in more detail.

20 // style transfer process (source : Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016 : 2.)

Sustainability and efficiency

The construction industry consumes around 40 % of global consumption and is responsible for 30 % of annual greenhouse gas emissions.

The buildings of the future must minimise energy consumption and carbon footprint, AI algorithms could make use of historical data to recognise patterns within the data and be able to make predictions for the future.

The population is also estimated to double by 2050, so it is crucial to design by evaluating the climate and sustainable aspects of a project.

Generative design can help the designer find new forms and ideas that can maximise natural lighting and thus reduce the thermal load, or support management by controlling energy consumption in real time. One example is the Edge Building in Amsterdam, which adopts AI systems for energy management, using 28 000 sensors to transmit data into a 'data lake' used to better manage the building.

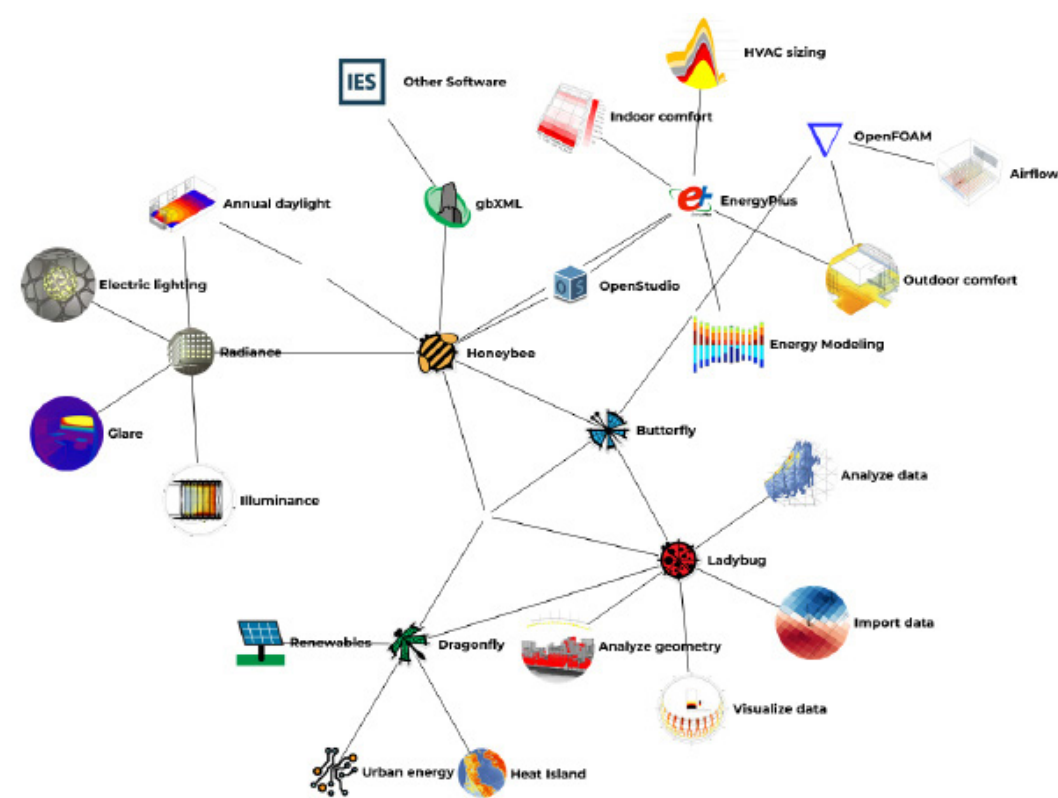
**"A computer with a roof".
(Randall, 2015)⁹**

⁹source : <https://www.enconassociates.com/news/breeam-smartest-building/>



One of the fields in which AI has the greatest influence is that of energy simulations.

Honeybee¹⁰, Ladybug¹¹ and Galapagos¹² (extensions of Grasshopper in Rhino) are tools that allow buildings to be evaluated from a thermal and climatic point of view under different conditions using advanced calculation algorithms that allow designers to intervene in their design from a climatic point of view.



22 // Ladybug, Honeybee and Grasshopper Ecosystem for building performance simulation (source : www.ladybug.tools)

¹⁰supports detailed daylighting and thermodynamic modeling that tends to be most relevant during mid and later stages of design

¹¹imports standard EnergyPlus Weather files (EPW) into Grasshopper. It provides a variety of 2D and 3D interactive climate graphics that support decision-making during the early stages of design

¹²a tool that searches a parameter space to find designs that optimize a “fitness” function

Can artificial creativity surpass human intelligence in fields like art and design, or is it inherently limited ?
How does artificial creativity differ from human creativity in terms of originality and authenticity?

The Artificial creativity a controversial intelligence

Bibliography

// Hanna, Robert. "The Non-Conceptuality of Creative Imagination and The Impossibility of Strong AI." (2016).
// López de Mántaras, Ramon. "Artificial intelligence and the arts: Toward computational creativity." (2016).
// Hui, Yuk. "On the limit of artificial intelligence." Philosophy Today 65.2 (2021): 339-357.
// Frantz, Roger. "Herbert Simon. Artificial intelligence as a framework for understanding intuition." Journal of Economic Psychology 24.2 (2003): 265-277.
// Leach, Neil. "In the mirror of AI: what is creativity?." Architectural Intelligence 11.1 (2022): 15.
// Gero, John S., and Mary Lou Maher. Modeling creativity and knowledge-based creative design. Psychology Press, 2013.
// Casakin, Hernan, and Andrew Wodehouse. "A systematic review of design creativity in the architectural design studio." Buildings 11.1 (2021): 31.
// Milošević, Jelena, et al. "Automated compositions: artificial intelligence aided conceptual design explorations in architecture." Proceedings: Geometry, graphics and design in the digital age, The 9th International Scientific Conference on Geometry and Graphics MoNGeometrija 2023. Faculty of Technical Sciences, University of Novi Sad, 2023.
// Hetrick, S. H., R. S. Lilly, and P. R. Merrifield. 1968. "Figural Creativity, Intelligence, and Personality in Children." Multivariate Behavioral Research 3 (2): 173-187

The Role of AI in Architectural Creativity

What defines creativity ?

Creativity is the ability to produce or use original and unusual ideas¹³. In architecture, creativity is the foundation of the process. The architect combines aesthetics with functionality, giving birth to spaces that respond to human and environmental needs. It is very important to highlight that this kind of creativity doesn't manifest just for aesthetics but also for complex problem-solving solutions.¹⁴

This chapter aims to address the multilayered complexity of architecture. The profession requires the skill to comprehend deeply and holistically the multitude of components of both the constructed and natural systems. This aspect is impossible to replicate by artificial intelligence right now.

Creativity in architecture is expressed in different ways and styles, defining the identity of the architect, which evolves with experience and knowledge gained throughout the years, project by project. This means that every architect brings unique experiences, cultural influences, and their own vision into the project¹⁵. In this way, the project absorbs those different influences and becomes rich and original.

The holistic capacity or skills are fundamental in the designing process: the originality of the method used to approach a project allows designers to face challenges in different and innovative ways.

With this focus, we wanted to express how important creativity is in the process and how it goes far beyond aesthetics or formal beauty. It is manifested in every single step of the process, through resolving complex problems, designing livable spaces, and promoting a good dialogue between the building and its context.

Through our reading and study of creativity and its effects, we firmly believe that AI possesses creativity and can achieve the same level of creative quality as humans. This is made possible by the way AI is designed, enabling it to mimic human thought processes.

¹³ Definition given from the Cambridge dictionary (<https://dictionary.cambridge.org/dictionary/english/creativity>)

¹⁴ in *Artificial Intelligence and the Arts: Toward Computational Creativity*, R. López de Mántaras, 2016 "... we believe that we must aim at more ambitious relations between computers and creativity. Rather than just seeing the computer as a tool to help human creators, we could see it as a creative entity in its own right..."

¹⁵ From the same document "...Since nothing can arise from emptiness, we must understand that every creative work or creative idea is always preceded by a historical-culture scheme; it is a fruit of the cultural inheritance and the lived experiences. As Margaret Boden states in her book *Artificial intelligence and Natural Man* (Boden, 1987)

The concept of GENIUS vs polymath's perceptions and capabilities



23 // view of an architectural office

In the paper “Christopher Alexander and the Inadequacy of Genius in the Architecture of the Coming Age,” it is discussed how the concept of Genius, in which the architect is figured as a “Creative Genius” capable of infusing his own artistic vision with a profound comprehension of the local culture and environment¹⁶.

In this way, the architect needs to know how to integrate unique instances into the project so that it can have its own importance in the context. However, Alexander shows that this canonical model risks promoting generic and universal solutions, which neglect the originality and singularity of that specific context.

During the academic journey to becoming an architect, one learns how to design following these principles. But doing so tends to standardize projects. On the other hand, Alexander tells us that the architect should not limit himself to theories and specific methods. Instead, he should have a critical eye and adapt to different situations so that the method becomes something to learn from rather than something to rigidly follow, allowing for the exploration of the variety of contexts. Looking at the Genius theory from this new perspective, we can see that this method risks not fully valuing human experience, thereby reducing the quality of the work.

In contrast to the Genius model, Alexander proposes the concept of the Polymath, a figure with knowledge and competence in a multidisciplinary environment. This kind of approach is much more flexible and capable of adapting based on the complex situations that can occur several times during a design proposal. The complexity that projects can reach, especially in contemporary times, requires a more dynamic and flexible approach, seamlessly introducing different figures and subjects into the project.

The architect no longer plays the singular authoritative role; instead, the aim of the profession is to help and guide the process, requiring the skills to adapt and collaborate with different expertise.

The importance of this concept is relatable to the role of AI and how it behaves in the design process. In fact, AI is a tool capable of collecting data from all different and existing disciplines, bringing the possibility of managing this complexity with very valuable assistance.

However, AI still lacks rationality and intuition. This “humanity” plays a crucial role in the process, and this is one of the reasons why, even though excellent tools exist, they are still not fully implemented in the design workflow.

¹⁶source : Dickinson, Duo. "Christopher Alexander and the Inadequacy of Genius in the Architecture of the Coming Age." Urban Science 4.2 (2020): 17.



24 // view of an architectural office

AI can mimic creativity with a quality comparable to human ingenuity; however, it remains far from achieving the intrinsically human trait of intuition.

Strong AI and Lack of General Cognitive Abilities

Artificial intelligence has advanced considerably in the last decade, but it still can't fully replicate human intelligence.

The widely known "Strong AI" is a model ambitiously attempting to create a flawlessly thinking and reasoning machine. Natural limitations greatly obstruct its achievement.

Artificial intelligence is fundamentally based on complex algorithms and rigidly predefined schemes, and therefore it cannot develop a subtle and genuinely perceptive comprehension of the world like humans. It concurs with R. Hanna's Paradosso del Limite and The Non-Conceptuality of Original Imagination and The Impossibility of Strong AI¹⁷, arguing that while AI can improve in speed and efficiency, it will always lack the complexity and depth of human thought.

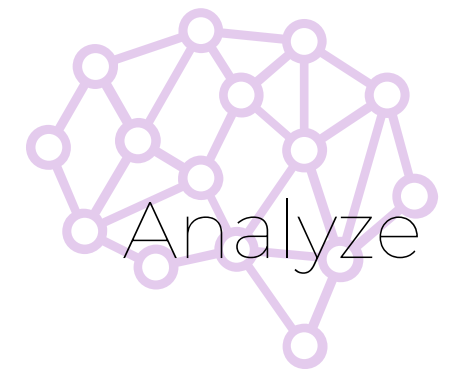
This paper divides all creativity into two specific typologies, and these typologies are human creativity and artificial creativity. The latter attempts to replicate the former. It will always lack some intuition along with conceptual capacities.

The original process has natural limitations that digital enhancements cannot overcome; these limitations extend beyond fundamental comprehension, importantly in relation to the original methods, intuitive comprehension of context, human emotions, and experiences.

¹⁷source : Hanna, Robert. "The Non-Conceptuality of Creative Imagination and The Impossibility of Strong AI."

Weak AI

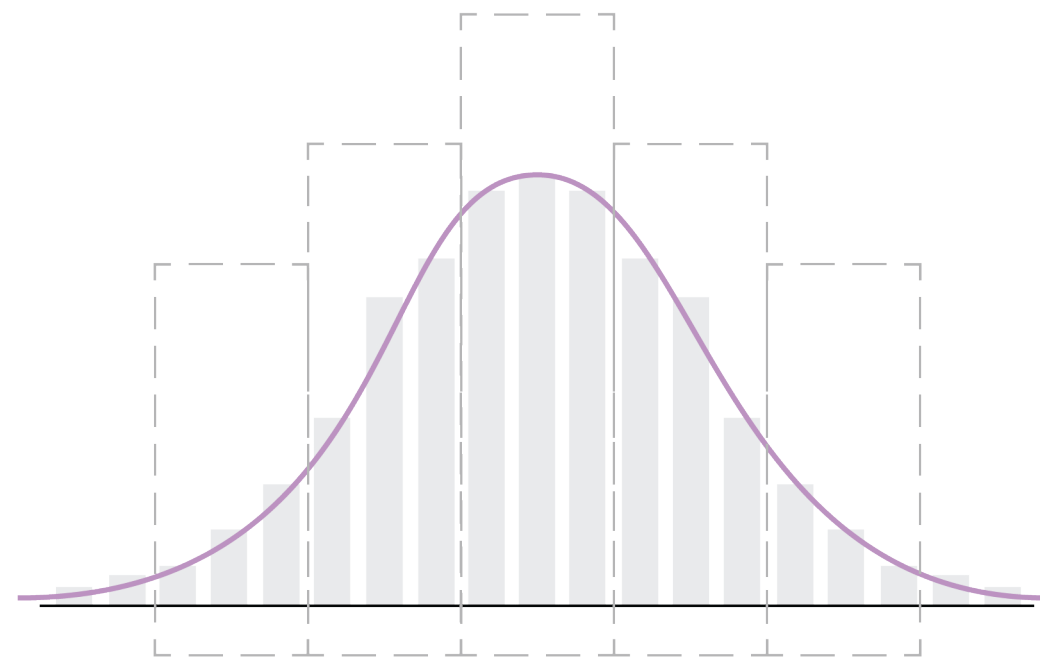
- Lack of intentionality and intuition
- Can't adapt in situation outside his knowledge
- No coprehension of the context



Strong AI

- Replicate the cognitive abilities of a human beeing
- Theoretical model, no existent
- Autonomous thinking and self-awareness





 25 // representation of the gaussian graph

This is representative of behavior of the artificial intelligence, the decisions made by the artificial intelligence are all based on probability of matching the requirements given by the human.

This is representative of the behavior of artificial intelligence; the decisions made by artificial intelligence are all based on the probability of matching the requirements given by the human.

To reinforce this notion of the limitation of artificial intelligence, we can refer to John Searle's Chinese Room¹⁸ Thought, where he highlights another fundamental limitation of AI: the lack of intentionality. Although an AI system can manipulate symbols and respond correctly to specific inputs, it does not really understand what it is doing. This experiment demonstrates how artificial intelligence lacks deep and meaningful understanding, merely following predefined rules without awareness.

However, despite the very fast progression of artificial intelligence in society and at the same time its boom in capabilities and digital progression, becoming faster, more reliable, and consistent in the outputs, it will never replicate exactly the human cognitive capacity. This kind of "intelligence" lacks intuition, intentionality, and deep contextual comprehension, basic elements for the way of approaching design for the human mind, which is completely out of the capabilities of machines right now. As in the subchapter before, we understood how the polymath aims to use this deep and holistic contextual comprehension at the base of the architect's thoughts and design.

In fact, artificial intelligence applied to the designing process, like other creative processes, should aim not to substitute the human hand and its presence but to help and assist during the work. Essentially, the human should have full control of the process, keeping it monitored and knowing what directions to give.

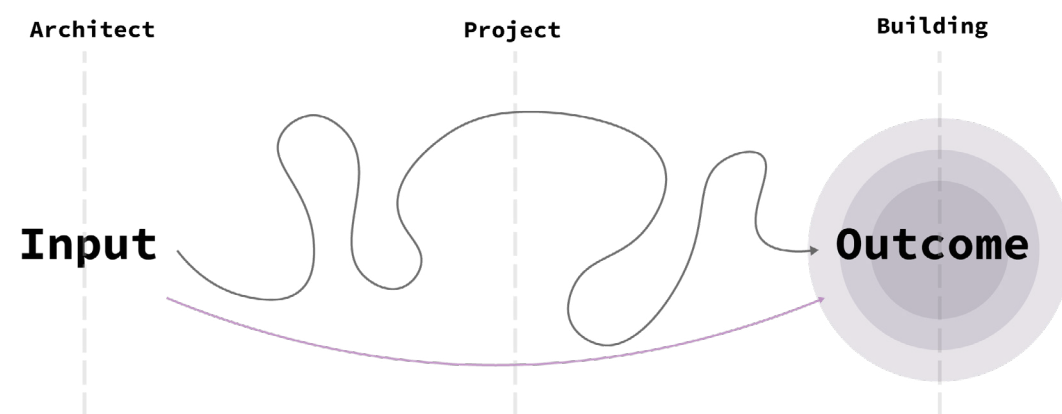
¹⁸source : López de Mántaras, Ramon. "Artificial intelligence and the arts: Toward computational creativity", 2016 :.

The Use of the Tool: The Duality of Efficiency and Effectivness

In fact, in the next future, the AI agents and Assisted Design will be essential tools to compete against other firms and stay on the market. These crucial tools will help to explore very fast different design solutions. The Assisted Design will aim to help [out] on the generation and development of ideas, accelerating the process and allowing the architect to concentrate on the creative and innovative aspects of the work.

Efficiency

The use of AI allows to make the design process much faster if used correctly. Thanks to the quality of elaborating vast amounts of data in a very short time, artificial intelligence would be a perfect assistant for elaborating fast concepts, tests, and final documentation that, summed up, would use a lot of the designer's time just in production. This means less time spent thinking about the idea, meaning also loss of quality in the work.



26 // schematic representation of efficiency and effectiveness in the process

Effectivness

Anyway, it's very important that the designer keeps the creative control of the process. This can be achieved utilizing the AI as an assistant rather than a ready-made solution.

In fact, the worst situation that can occur comes when, utilizing too much the tool without thinking about where to use it and what its aim is, can lead to leaving the control of everything to the tool and then losing the human quality in the project, leading also to unsustainable and unfeasible projects.

This happened also with the birth of new tools like CAD, BIM, and Grasshopper. These kinds of tools are very powerful, but if not understood and mastered correctly, they can completely change how we think and approach the project, creating limitations in the designer's mind, forcing the designer to think in a determined way. For example, if we use Grasshopper for developing a design, we often get stuck in a very pragmatic way of approaching the development of the geometry, using very strict and logical steps, when maybe the problem, using another perspective, can be resolved and approached very easily.

It's very important to highlight how the prompting is very important to master and is a completely different way of controlling a tool. It allows us to understand the machine in depth but also what we ourselves seek from the machine.



27 // final goals of efficiency and effectiveness

The Value of a New Way of Expressing

Artificial intelligence can bring significant value to the creative process, not only by improving efficiency, but also by enriching the expressive language of design.

Artificial intelligence offers new expressive possibilities, allowing designers to explore innovative and unexpected solutions. It can open new avenues in visual representation, allowing them to overcome the limitations of traditional techniques and experiment with new styles and forms¹⁹. By nature, artificial intelligence tries to connect the directives given through the written prompt with what it is aware of, thus its LLM. This allows, with large amounts of data, to put together pieces of different designs, images, and types of visualization to create something new and unique in its appearance. It cannot be described as new and original or creative but still works like the human brain by making connections and associations, putting together knowledge and creating something new.

A crucial aspect to consider is ethics in the use of artificial intelligence. AI should not only be used to speed up the process at the expense of quality, but rather to enrich the creative process. Its integration should aim to improve the quality of the final work, giving the designer a powerful tool to better express his or her ideas and concepts²⁰. Just as artificial intelligence can be considered a deceptive tool, it is frowned upon by society as a tool that, obviously if used with certain intentions, can lead to a poor-quality result, attributing the result not to the author per se but to the machine because it “did all the work.” Instead, also through this thesis, it is intended to show how it is a faithful assistant in developing material and how it can be a very powerful tool not only for “copying” but for experimenting, developing, and learning, creating something unique. Another concern could be that using AI doesn't facilitate learning in the same way, as it reduces the friction typically involved in the process. However, this doesn't change the fact that the person in charge must still review the material and maintain full control over the development. Relying entirely on AI for decisions would compromise the quality of the work.

¹⁹source : López de Mántaras, Ramon. "Artificial intelligence and the arts: Toward computational creativity.", 2016 : 2.

²⁰source : Simon, 2019



How is the desining process divided ?
How can we manage the process ?
What makes it efficent ?

AI in the process: Different phases

Bibliography

- // Beyan, Eleonora Vilgia Putri, and Anastasya Gisela Cinintya Rossy. "A review of AI image generator: influences, challenges, and future prospects for architectural field." *Journal of Artificial Intelligence in Architecture* 2.1 (2023): 53-65.
- // Jaruga-Rozdolska, Anna. "Artificial intelligence as part of future practices in the architect's work: MidJourney generative tool as part of a process of creating an architectural form." *Architectus* 3 (71 (2022): 95-104.
- // Agkathidis, Asterios, Yang Song, and Symeonidou Ioanna. "AI-Assisted Design: Utilising artificial intelligence as a generative form-finding tool in architectural design studio teaching." (2024).
- // Almaz, Amira Fawzy, et al. "The Future Role of Artificial Intelligence (AI) Design's Integration into Architectural and Interior Design Education is to Improve Efficiency, Sustainability, and Creativity." *Sustainability, and Creativity* 3.12 (2024): 1749-1772.
- // Fitriawijaya, Adam, and Taysheng Jeng. "Integrating multimodal generative AI and blockchain for enhancing generative design in the early phase of architectural design process." *Buildings* 14.8 (2024): 2533.
- // Tan, Linus, and Max Luhrs. "Using Generative AI Midjourney to Enhance Divergent and Convergent Thinking in an Architect's Creative Design Process." *The Design Journal* (2024): 1-23.

The Architect Creative Journey

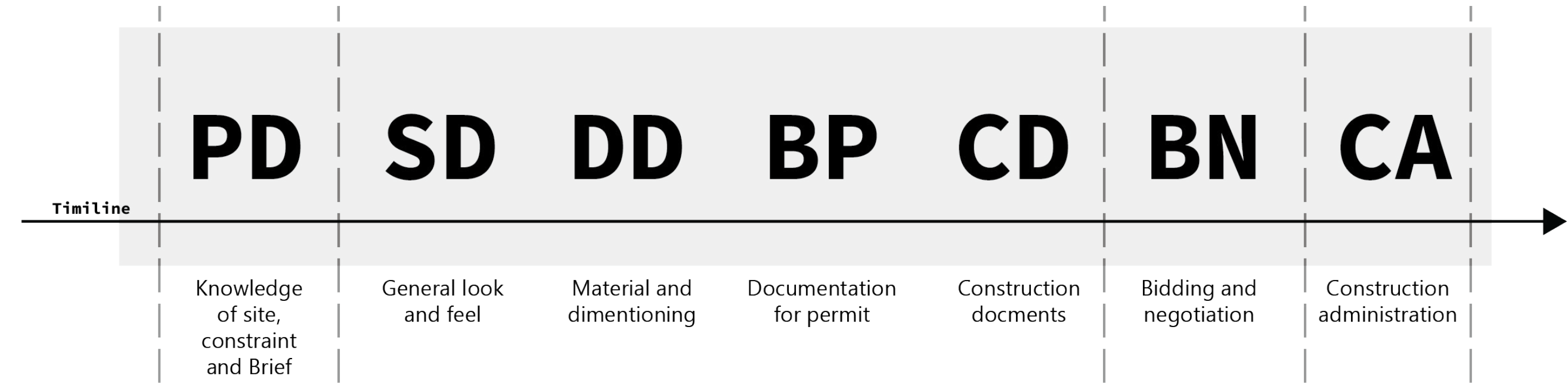
Project Planning and phases

The standard, in the international offices, is to subdivide the different phases of the project. However, this changes based on the specific needs of an office or dedicated laws in the country where the office is working.

This division in phasing is the plinth of the office structure; this helps to maintain coherence with the timing and the management of the process, focusing to have a constant effort and not peaks. Understanding, with a deep dive, this process and how it is structured helps us to understand where the use of diffusion models would be more powerful and interesting to study.

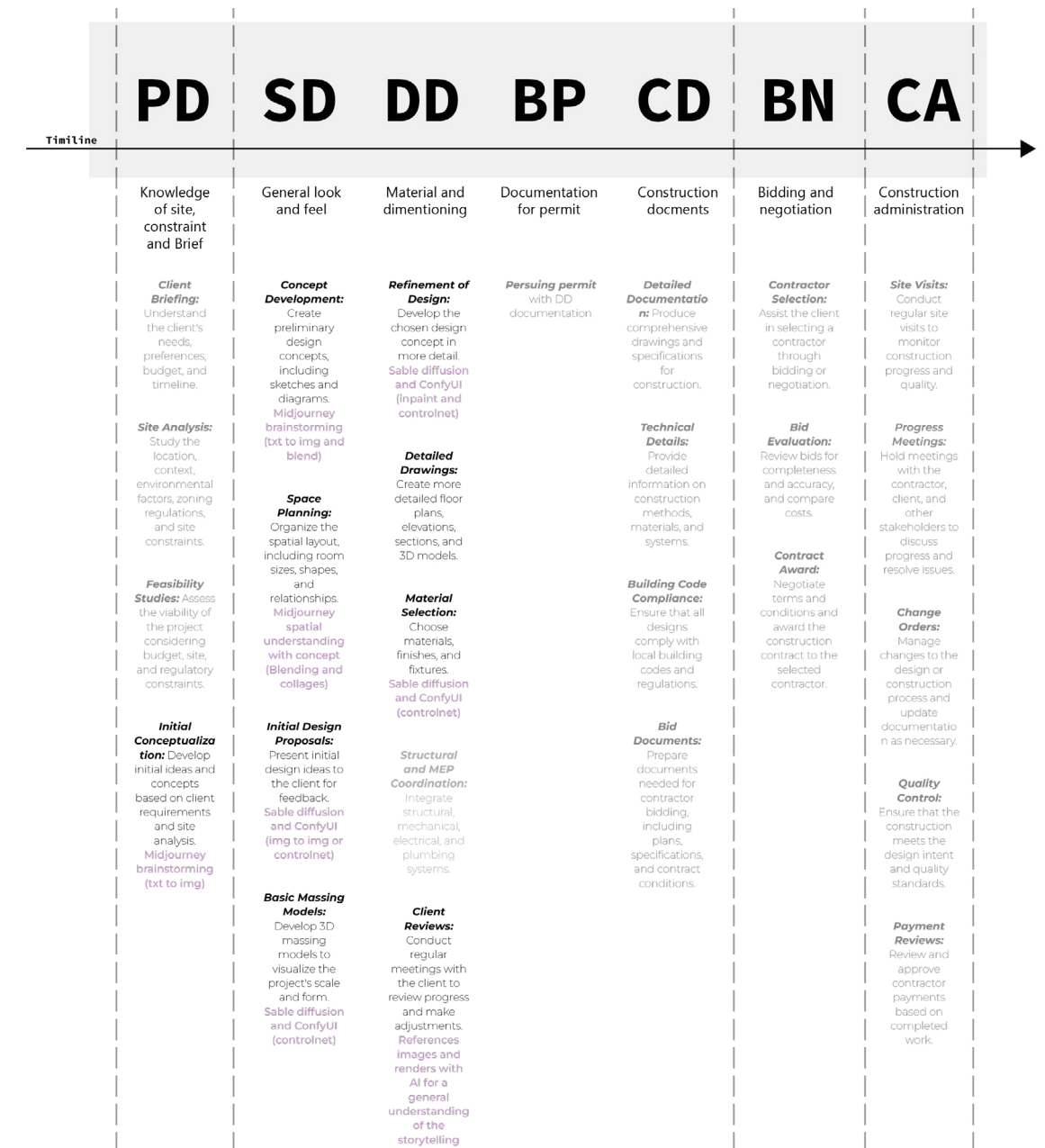
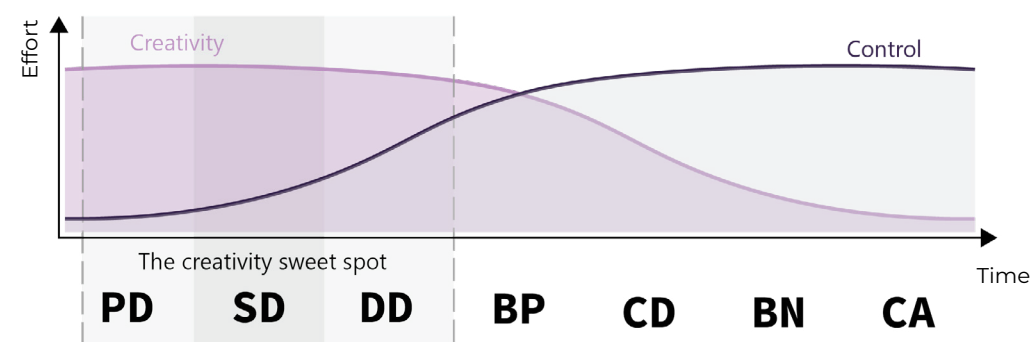
We can distinguish the process in 4 macro stages:

- 1. Research
- 2. Design and documentation
- 3. Pricing
- 4. Construction



The last four categories are the least interesting for the aim that this thesis has, this is because they are very technical, where the control of the drawings is key. This doesn't mean that the use of AI is useless in this part of the process, but it focuses on rather something different, where also the use of a tool that is able to manage large amounts of data can also be useful and efficient. Indeed, it is a world that has been under research and observation for several years now, and research still continues, especially by large software giants and distributors.

This thesis, however, instead wants to address how a specific branch of artificial intelligence, that of Diffusion Models, can influence these different phases by coming to the conclusion that it is appropriate to test and research only in the most creative and most dynamic phases of the project. So we will focus on the areas of Research, in part, and in Design and documentation for the most part.



The Initial Stages of Architectural Design

These macrocategories include within them:

PD, the preliminary phase where there is purely research on the site and gathering information on possible limitations and bureaucratic information, thus getting to know the site as much as possible. This part is not of interest to us though, as in this phase already we start experimenting and researching what will be the basis of the overall design concept.

SD, the beginning of the designing phase where there is a first real phase of study and research of the concept that will then be extensively explored in the subsequent phases.

DD, this phase enters into the details of the project when the general concept has been defined, the volumes and spaces. Here you want to give a more detailed representation and plausibility of the project.

Once the different phases of the process were identified, we proceeded with a further, more in-depth analysis, through also what were the personal experiences we had within these large international firms. We drew up a list for each phase of the main actions to be carried out within it. Obviously, this list is not meant to be qualitative, but rather it is meant to explore in a generic way the process. This premise is made because the process varies along with the variation of the project, from the type of client and the types of contracts.

Once this list was developed, an attempt was made to understand what and if there was a connection to artificial intelligence. By connection, we mean help or assistance from AI through tools or workflows that could be developed, thus trying to find the best methodology, new ideas of workflows, and the possibilities that these software can give.

PD Preliminary Design

Define the project's scope, feasibility, and initial concept.

Key Steps:

- Conduct site analysis (topography, zoning, context)
- Outline client needs and program requirements (spaces, functions).
- Develop rough massing studies or conceptual sketches.
- Prepare initial cost estimates and timeline projections.

SD Schematic Design

Purpose: Translate concepts into visual/spatial solutions.

Key Steps:

- Create floor plans, elevations, and 3D massing models.
- Establish spatial relationships and circulation flow.
- Select primary materials and systems (structural, mechanical).
- Refine budget and align with client feedback.

DD Design Development

Purpose: Refine the design with technical precision.

Key Steps:

- Finalize dimensions, materials, and construction details.
- Integrate engineering systems (HVAC, electrical, plumbing).
- Develop detailed drawings for key components (e.g., walls, windows).
- Update cost estimates and specifications for contractors.

How is artificial intelligence integrated by MVRDV and UNStudio ?
How can these applications and experiences be useful to us ?

AI in the process: Different experiences

Sitography

<https://space10.com/projects/regenerative-futures>
<https://www.unstudio.com/team>
<https://www.mvrdv.com/about/team>
<https://arch-hive.com/competitions>
<https://www.youngarchitectscompetitions.com/past-competitions/weihai-shandong-art-museum>

The candidates

MVRDV & UNStudio

After an interview research on how the design process works in international offices, in order to be able to deal with exercises and understand how artificial intelligence can best be incorporated into the process, we decided to investigate the experiences of professionals. In large offices, it is important to note that there are people who are highly specialised in certain aspects, therefore in contact with one phase rather than another. Moreover, these different roles have distinct responsibilities within the same phase, based on their professional expertise.

This introduction is therefore necessary in order to be able to set the correct interpretation keys for the interviews that will be presented and analysed in the following pages. In fact, without any preamble they may appear to be mere professional chats when in fact they target specific interests.

The interviews are aimed at analysing and collecting the experiences of professionals in highly competent offices such as those of mvrDV and unstudio. They want to reveal the potentials and limitations of these tools through the eyes of very competent people. Despite these competences, not all figures within these offices have decided to approach the use of artificial intelligence. Many people do, but the majority are amateurs and with very casual uses, so the choice of figures was of fundamental importance in order to have profitable and interesting discussions.

It is important to emphasise that the presence of these interviews within the thesis is intended to strengthen knowledge of the topic and to support the development of four different insights. These insights will be at the heart of the thesis, they are derived from the process in its entirety and from the extensive experience developed from the internship period and beyond, carried out in the respective studies.

The logo for MVRDV, consisting of the letters 'MVRDV' in a bold, black, sans-serif typeface.The logo for UNStudio, featuring the letters 'UNS' in a large, bold, black, sans-serif font, with the word 'UNSTUDIO' in a smaller, all-caps, black, sans-serif font directly beneath it.

The experts profiles



 32 // Lorenzo Mattozzi

Nationality and cultural background

Italian, with extensive international experience, having worked in Italy, Denmark, Brazil, and the Netherlands.

Profession

Architect with expertise in architectural design and the coordination of complex projects from concept to detail.

Work experience

20 years of professional experience across offices of various sizes and profiles. Over the past 7.5 years at MVRDV, he has taken on a leadership role in the projects he is involved in.

Specialisation or reputation

Lorenzo has extensive experience in the Chinese and German markets, ensuring effective communication with clients and stakeholders while streamlining the design and delivery process. Many of his recent projects focus on refurbishment and adaptive reuse. Additionally, for the past five years, he has been a regular lecturer internationally and a tutor in the two-week summer course Venice Studio at the University of Melbourne.



 33 // Cosimo Scotucci

Nationality and cultural background

Norwegian, with a strong influence from Scandinavian culture, known for its focus on sustainability, functional design and technological innovation.

Profession

Senior engineer, specialising in the management of complex infrastructure projects.

Work experience

He has worked as an engineer for more than 20 years, with 12 years of experience at his current firm, where he holds a leadership role in planning and technical coordination.

Specialisation or reputation

He is known for his expertise in sustainable infrastructure design, particularly in the optimisation of energy systems and the construction of bridges and structures resistant to the harsh climates of Northern Europe. He has participated in numerous high-profile international projects.



34 // Fredy Fortich

Nationality and cultural background

Colombian, with a rich cultural background influenced by his international work experiences, which include collaborations with European and South American firms.

Profession

BIM Manager, a senior figure specialised in managing processes and workflows related to Building Information Modeling (BIM).

Work experience

He has worked in the BIM sector for more than 10 years and has held the position of BIM Manager for 4 years at his current firm, where he is responsible for digital integration in design and construction processes. In addition, he is one of the leading figures responsible for experimenting with artificial intelligence within the professional field, as part of the MVRDV NEXT cluster.

Specialisation or reputation

He specialises in the management of complex projects with a focus on advanced 3D modelling, interdisciplinary coordination and the implementation of BIM standards to optimise efficiency and accuracy. He plays a central role in the introduction of technological innovations within the firm.



35 // Olaf Olden

Nationality and cultural background

Born and raised in the Netherlands, I graduated in 2021 from TU Delft with a Master's degree in Architecture. My academic journey at one of the world's leading institutions for design and technology shaped my passion for innovative and sustainable architecture. Currently residing in Rotterdam, I am deeply inspired by the city's dynamic urban environment and its forward-thinking approach to design.

Profession

As a Sustainability Specialist at UNStudio, I focus on creating tools that enhance architects' workflows and decision-making processes. My role revolves around facilitating awareness of sustainable practices and promoting a data-driven approach to design. By integrating advanced tools and methodologies, I aim to bridge the gap between creativity and performance, ensuring that sustainability becomes an inherent part of every project.

Work experience

Prior to joining UNStudio, I gained valuable experience as a Sustainability Consultant at MVRDV, where I contributed to embedding sustainable strategies into architectural projects. At OMRT, I further specialized in computational design, working as an expert and project manager to deliver innovative solutions that combined technology and architecture. These roles allowed me to develop a strong foundation in managing multidisciplinary teams and delivering complex, high-impact projects.

Specialisation or reputation

My specialization lies in computational design and sustainability simulations at both urban and building scales. I am passionate about developing tools that empower architects to make informed, data-driven decisions, ensuring that their designs are not only innovative but also environmentally responsible. By combining technology, creativity, and a commitment to sustainability, I strive to push the boundaries of what architecture can achieve.



36 // Enes Sever

Nationality and cultural background

Born in Turkey and raised in North Macedonia with Albanian ethnicity, I grew up in a diverse environment that shaped my perspective on architecture. I was always drawn to Brutalist architecture in the Balkans. Its raw materiality, strong forms, and functional approach influenced how I see design.

Profession

I pursued my passion for architecture at Politecnico di Milano, where I earned my Master’s degree and gained exposure to an international design culture. During my studies, I participated in multiple competitions and won several awards, which strengthened my creative and problem-solving skills.

Work experience

As an Architectural Designer, I worked at UNStudio, contributing to projects at different stages with a focus on innovation and design. Before that, I interned at Mecanoo, where I worked on a range of design challenges that improved my technical and conceptual skills. I also represented North Macedonia at the Venice Biennale, which gave me experience in presenting work on an international stage.

Specialisation or reputation

My expertise is in architectural innovation and design, with a growing interest in artificial intelligence and its role in architecture. I am exploring how AI can improve design processes, increase efficiency, and bring new possibilities to architecture. By combining technology and design, I aim to create practical and forward-thinking solutions.



37 // Cristobal Burgos

Nationality and cultural background

Chilean, with a professional journey shaped by a blend of Latin American creativity and global technological innovation. My experiences in Chile, Finland, and the Netherlands have enriched my perspective on design and technology.

Profession

Tech Lead and architect, with expertise in integrating cutting-edge technology into design workflows to enhance creativity and efficiency.

Work experience

I have worked in architecture and technology for over 10 years, collaborating with renowned firms like Cecilia Puga and UNStudio. As the Tech Lead at UNStudio.Tools, I lead the development of innovative tools such as The Carbon Builder, implement AI strategies, and oversee the deployment of globally integrated web-based systems. Previously, I founded MALAForma, a design studio specializing in parametric design, data visualization, and BIM strategies, and led the creation of DATA Lab at MOBIL Arquitectos.

Specialisation or reputation

I specialize in bridging the gap between high-quality design and technological innovation. My expertise lies in developing tools and strategies that optimize design workflows, enhance data-driven decision-making, and push the boundaries of what is possible in architecture and technology.

Learning from the field

Methodology

The interviews were conducted using a semi-structured methodology, sufficiently balancing structure with flexibility. The respondents' suggestions enabled adjustment of the conversation. This guaranteed the conversation was relevant to their input. The professionals received the questions before the interview. They could think quite deeply and offer a perspective that was very well researched and consistently mindful because of this.

The main goal was to guarantee the respondents were always comfortable by providing a friendly and relaxed setting. This method improved all postures of all involved parties. The goal was to achieve a relaxed discussion of the issues.

Permission was gained to audiotape all conversations for accurate transcripts and trustworthy documentation.

After the interviews were done, the next step was transcription. The interview was not transcribed word-for-word. Instead, while preserving the meaning, we constructed a remarkably coherent and surprisingly fluid dialogue. We did this to make the document much easier to read and understand all while keeping it whole. The professionals checked and approved the transcriptions and their profiles to confirm accuracy and grant publication permission.

The transcript analysis at this stage looked at the main themes and different opinions. This activity helped build a very solid foundation, as well as understand all of the respondents' answers. This analysis fully looked at varied views, noting repeated concepts and distinctions that supported the development of fresh hypotheses and guidelines.

This method let us discover new ways to use the toolsets. We also learned about their function in the process. This method, along with these guided conversations, stressed the quality of all the collected information because it was quite trustworthy as well as very helpful for creating new standards and methods to be followed in professional practice.

The Interview Questions

CONTEXTUAL UNDERSTANDING

- How do you perceive the current landscape of AI in architecture? Is it truly transformative, or is it being overhyped in some respects?
- What do you see as the most critical research areas or questions regarding Generative AI in architecture today?

AI'S ROLE IN DESIGN

- Do you see AI as a continuation of technological evolution in architecture, or does it represent a completely new paradigm?
- What are the potential and limitations of AI in generating design ideas? Can AI ever be a true partner in the creative process?
- How do you see AI being integrated across different project phases such as Preliminary Design (PD), Concept Design (CD), and Schematic Design (SD)?
- Do you think AI will primarily serve to enhance efficiency in architecture?

TECHNOLOGICAL EVOLUTION

- What AI tools are you currently utilizing, and how do you assess their capabilities and limitations in practical terms?
- What are your thoughts on tools like ComfyUI and Midjourney? Have they changed the way you approach architecture?

LOOKING FORWARD

- What do you believe will be the most significant paradigm shifts in architecture due to AI over the coming decade? Do you foresee the emergence of new roles or a restructuring of the architectural process as a result?
- Looking ahead, do you believe AI will enable more personalized and responsive architectural designs, or will it standardize and limit creative expression? Could it lead to new methods of communication between architects, clients, and the public?

MVRDV Answering

Lorenzo discusses how AI tools like MidJourney, Stable Diffusion, and ComfyUI are enhancing architectural processes, offering real-time modifications and new creative possibilities. While he values these tools for inspiration, he emphasizes that AI won't replace the core aspects of architecture, such as sociological, regulatory, and ethical considerations. AI will undoubtedly increase efficiency, but human expertise in materials, structure, and design details will remain crucial.

Cosimo recognizes AI's role in streamlining workflows, aiding research, and supporting creative exploration. He values tools like ChatGPT and Stable Diffusion for idea generation but notes that they fall short in creativity and abstract thinking. Cosimo believes human input is vital in interpreting and applying AI's capabilities.

Fredy highlights AI's potential to enhance efficiency in architecture, boosting productivity through quick renders and data-driven designs. While acknowledging AI's capabilities, he raises concerns about its effects on creativity and the environment. He believes that architects' critical thinking and human touch remain irreplaceable.

Lorenzo Mattozzi

Architect | Associate



Cosimo Scotucci

Senior Project Leader |
Architect CNAPPC



Fredy Fortich

Architect | Technical
specialism



UNStudio Answering

Olaf acknowledges AI's potential to enhance efficiency in architecture, especially for quick renders and project optimization, but raises concerns about the data it relies on and its impact on creativity. He believes that while AI could dramatically increase productivity, the essential human role of the architect will remain crucial.

Enes sees AI as beneficial for improving efficiency and creativity in the early stages of architectural projects, where conceptual work is more flexible. However, he recognizes AI's limitations in tasks that require precision, suggesting that architecture must adapt to embrace AI's potential, though it's not yet at the point where AI can fully replace human input.

Cristobal believes AI is currently overhyped, often used for simple tasks instead of its full potential. While tools like MidJourney lead to repetitive results, ComfyUI offers greater customization and learning. He foresees the emergence of the role of prompt engineer, though he considers it more a skill than a distinct profession.

Olaf Olden

Sustainability Technologist
at UnStudio



Enes Sever

Architect at UnStudio



Cristobal Burgos

Senior Digital Developer



Pattern Recognition in Interviews, a Reflective Debate

THE CONCEPTUALIZER

AI is developing into a influential tool to help generate ideas. Many interviewees wanted to use this tool for better brainstorming, but some contradictions and perplexities still exist.

Lorenzo, for example, acknowledges AI as a particularly strong tool for idea generation, but argues that architecture is simply not about visual representations, downplaying the importance of images in the process. Images are more than just illustrations. Like good references, they are tools that can explore new opportunities, increase creativity and reinforce ideas. This dualism between AI's usefulness and its limited use in architecture is a recurring theme.

In his interview, Olaf emphasize the architect is a all-including figure; thus, even though image is important, it is not capable of handing all the tasks in the process. However, he introduced one important concept: AI is similar to hand sketching, a useful tool for experimenting with, and studying many ideas. This comparison makes sense and is solid because they both let you quickly check out multiple design ideas.

Enes provided a description of creativity. He stated it results from plentiful connections between experiences as well as knowledge. He stated that AI can strengthen and accelerate this process with the use of Diffusion Models, whose main trait involves generating many images with variable control and large speed. Enes acknowledges that control is still an important limitation, but he sees AI as a way to explore truly new and original routes.

AN ADDITION TO THE TEAM

AI is widely viewed as a useful addition to architects' toolkit, not as a complete replacement for them. While interviewees think AI can help architects and not take their place, they also say its job shifts depending on the project stage and what tools are used.

Lorenzo described how these tools can be added to the future design process, allowing tasks to be done faster as well as better. The main goal is not to reinvent the process, but to considerably optimise it by importantly accelerating the most time-consuming and highly repetitive phases. A number of people share this approach; Cosimo, for instance, pictures a future where varied AI 'agents' can help at all project stages. Cosimo concentrated on using LLMs (Large Language Models) rather than Diffuse Models, allowing us to greatly expand how we see AI. He also agreed that these tools will be more integrated throughout the process. This is for increased control and efficiency.

Fredy described the different ways AI can be used throughout the process, from the start to the finish and pointed out specific workflows for each stage. He emphasised that Diffuse Models can be integrated into brainstorming, materiality study, modelling and final rendering. Fredy considers AI to be a transversal element. It can make each project stage more efficient. He also recognised the suitability of LLMs like ChatGPT for general tasks. These LLMs are better suited for many thorough tasks, including certain data analyses and integrations with BIM softwares.

THE CUSTOMIZATION

One of the most discussed topics is making AI tools fit what architects need. Customising AI tools is key for getting the most from them. This helps on fighting the lack of originality and loss of quality in the work done.

Lorenzo stated that AI is an efficient toolkit for idea generation; however, he also noted that because these models frequently use identical datasets, one reuses previously encountered ideas. He suggests to make Fine Tuners or Models trained in the style of a large studio. This would allow outputs to be customized for the project's specific needs, avoid being repetitive and create more original results.

Cosimo addressed a related problem, noting the generated images are still too general and unmanaged. Best practices hinge on accuracy and originality. This is particularly problematic. Cosimo anticipates finding a way to address this entire problem. He states that the problem comes when the big offices are addressed to design an iconic and monumental architecture. He also said that these models, especially if trained correctly, can mix different references in ways that we would not be able to do. One strength of AI is its ability to combine different parts in new ways.

Fredy emphasised the importance of certain dedicated models for architecture because such models can speak the same language as architects. These models would be used specifically for each of design intention, style, or materiality. Several limitations must be addressed to achieve superior control of the outcome. These steps should make AI far more useful as well as precise. Olaf as well as Enes both particularly emphasised that the current tools remain too generic. They require more specific models for architecture. However, Cristobal noted that AI, which has been already around for quite some time, is being rediscovered and used as a new tool, especially by architects. Many more specialized branches are developing from generic models.

DRIVEN ARCHITECTURE

The topic of AI-driven architecture receives large discussion. AI could become an important driver in architectural processes, a future that may be nearer than many believe. The majority of Interviewees are optimistic about AI's potential to play a large role in the coming years. For now, AI is a tool that helps optimize and accelerate the process. Complete human control remains necessary, especially in all advanced project stages.

Cosimo stated AI could eventually show the path, yet acknowledged we are not currently prepared. Today's tools do not really allow for dealing with the large complexity of architecture, but many future possibilities will exist. An AI-driven process is worth considering when AI can be precisely controlled as well as when it is more specific to architecture.

Discussing the importance of prompting and the difficulty of using AI later in a project, Olaf spoke on the same topic. At these stages, precision counts. That level of control is not yet achievable by AI. Olaf emphasised that because many decisions have been made already, extreme precision is necessary. Enes particularly emphasised thorough control over AI tools as especially important in the project's later stages. He explained that precision matters because many things have already been decided and planned in these phases. He also believes AI can be used well if each step of its work is fully controlled.

Case Studies : Competitions

The first step in understanding how AI works in the architectural process is to analyse and study concrete examples of projects in which AI is used.

Today, it is difficult to find actual projects in which the use of artificial intelligence is openly declared, but one can find multiple design competitions in which the use of AI alone is requested. It is interesting to investigate and study in order to understand the raw output that AI can give to today's designers, these are images and projects where no other tool is used to modify or improve the design or the quality of the image, having said that we can also dwell on the limits that AI has and understand how to implement its correction in a new architectural process.

The first projects we are going to analyse are some of the winners of the contest organised by Space10²¹.

Analysing some of the projects submitted for the contest we can immediately see how it is the work of AI, the abundance of curved shapes and structures that are difficult to implement makes its use stand out, but beyond this it is interesting to highlight the details and common errors that are recurrent in the use of AI.

Deepening the image we notice how there is a lack of definition and correct human anatomy and in some parts of the building, these are minor but common errors in most images, AI is constantly expanding and developers are constantly working on solving these types of problems, at the end of the case study analysis an example will be given on how these types of errors can be integrated into a workflow and fixed for a correct image.

Apart from the consistency or otherwise of certain architectural parts, many images realised with AI always present a smooth and scattered appearance. This is common and normal, especially when an upscaler²² process is used to improve image qualities.

²¹a research and design lab on a mission to create a better everyday life for people and the planet

²²process of increasing the resolution of an image, usually by enlarging the pixel size or increasing the size of the image itself





39 // image part of the SpaceIQ contest with highlighted common mistakes made by AI



40 // image part of the SpaceIQ contest with highlighted common mistakes made by AI

To remedy simple visualisation errors such as those listed above, one can use photoshop or artificial intelligence itself in a second pass once the image is confirmed, in case we directly need the images produced by the AI as final output, in case it is for purely design and subsequent modelling purposes there is no need to stress or correct these errors.

This makes it possible to create a new process that includes the AI among the main characters.

Of course, there are many online competitions in which only the use of AI is allowed. Generally, AI allows the creation of out-of-the-ordinary shapes, situations and architecture, which is another reason why we can immediately tell if they are images created with the help of artificial intelligence, as in the case of some of the projects presented in the ArchHive²³ contest.

Continuous research and developments in the field are trying to completely eliminate graphic and image errors due to the machine to create an output that reflects reality as best as possible.

These competitions do not have any kind of constraints from a design point of view, unlike a real project where there are constraints such as the location of the project or the massing or otherwise parts of the project building.

This requires more control on the part of the AI and with slightly more complex processes one can achieve the desired control over the starting image or idea.

²³platform designed for architects to create and showcase professional portfolios easily (source : <https://arch-hive.com/competitions>)



AI in the process : Different experiences

41 // image part of an ArchHive AI contest (credits : Opal Naomi Markus Ilan)



42 // image part of an ArchHive AI contest (credits : Opal Naomi Markus Ilan)

Page 108



43 // image part of an ArchHive AI contest (credits : Vlad Vulturari)



44 // image part of an ArchHive AI contest (credits : Anda-Cristina Popescu)

Page 109

Case Studies: Living Trails



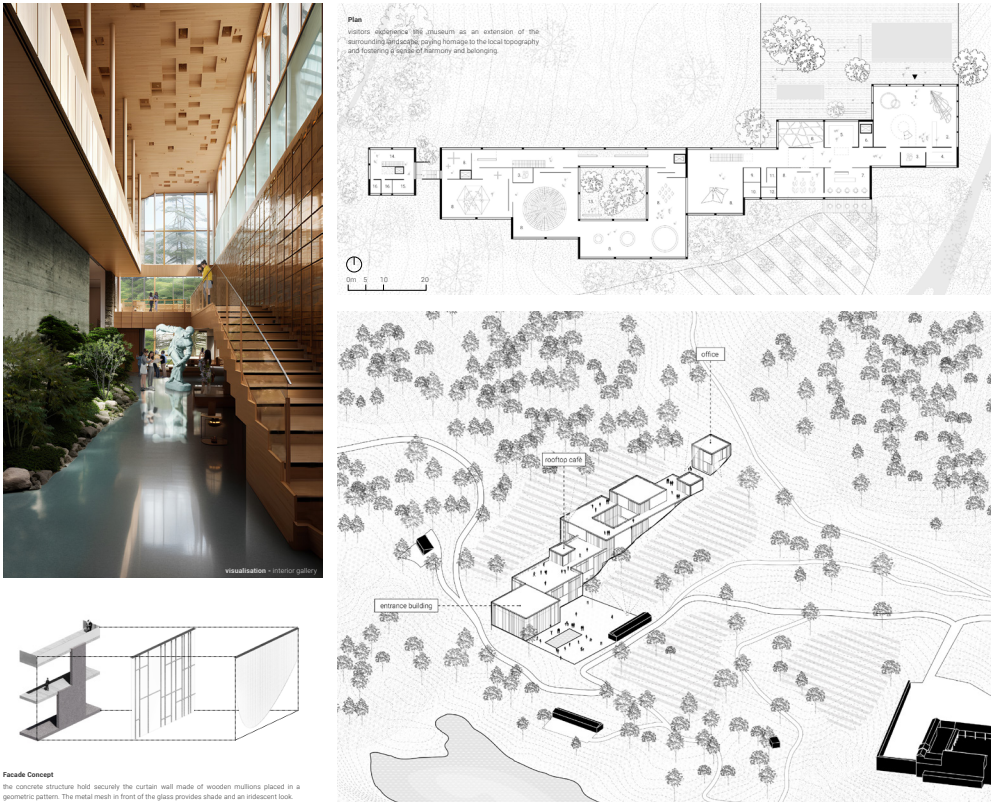
"This competition aims to create an art/museum hub that will become a socio-economic driver for the city of Rushan. It poses challenges on multiple levels for all architects. Firstly, the exhibition theme has always been a particularly sensitive design issue: the success of a museum or cultural center does not only depend on the quality of the collection displayed there, but also - if not especially - on the quality of the visitor route. It is movement, discovery, and scenic and theatrical devices that make a cultural center an attractive place, providing visitors with a rewarding and cathartic experience. Secondly, the client's goal is to create a museum hub that is connected to its context while involving and enhancing the surrounding landscape. Therefore, architects must design an exhibition route that is intertwined, connected, and enhanced by virtues of its relationship with the natural environment. Furthermore, the project will enhance and materialize local identity: by visiting the museum, visitors will experience the most authentic essence of rural China, of an ancient and fascinating region. It will be a pilot project, which can serve as an example and inspire further development of neighboring areas, such as Weihai, and thus turn the region into the cultural heart of Shandong province."²⁴

Living Trail is the project realized by Marco Papagni and Chiara Catani for a new art museum in the city of Rushan, China. The project was carried out with the support of AI. In this case, Midjourney was used to generate the images and the initial concept. Then, these images were used as the basis for the construction of the 3D model and were retouched in Photoshop to make them completely follow the direction of the project and the idea.

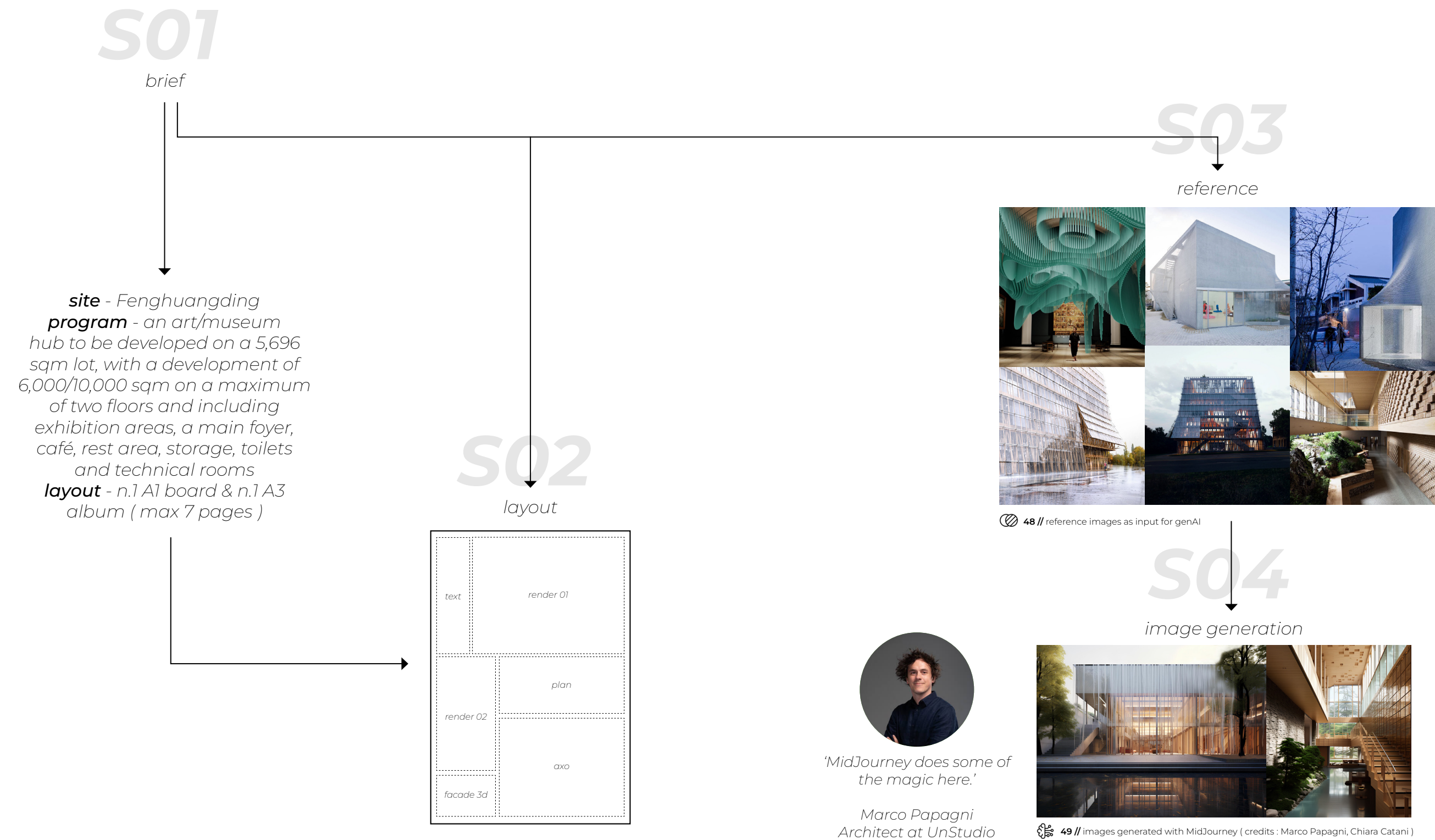
This is an excellent practical example of how AI was seamlessly integrated into the architectural process as an added value, with the process used being the reverse, starting with the creation of the images and then moving on to the 3D modelling and subsequent production of the required deliverables.

45 // on the right : submitted panel of the project Living Trails (credits : Marco Papagni, Chiara Catani)

²⁴source : from competition brief, <https://www.youngarchitectscompetitions.com/open-competitions/weihai-shandong-art-museum>

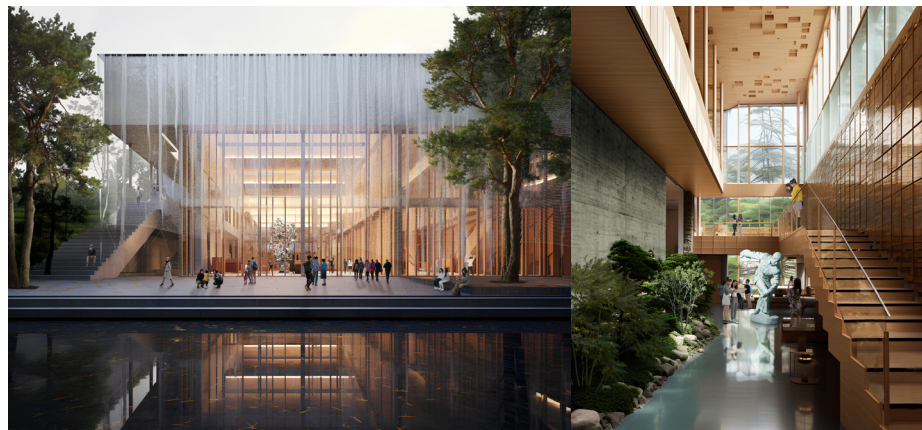


<i>S01</i>	<i>brief</i>	the first step consists of analysing the brief, in this case it was done manually but as we will see later it can be analysed through an LLM to speed up the process and get the key parts
<i>S02</i>	<i>layout</i>	once the brief has been analysed and it is understood what is required in terms of deliverables, the layout has been set up so as to understand the spaces that each one occupies on the board
<i>S03</i>	<i>reference</i>	prior to image generation, a few references are chosen to serve as a basis and reference point, in this case the competition already offers references as a starting point, but personal references were chosen
<i>S04</i>	<i>generation</i>	the key step, the generation of images, based on the given references and the prompt, the software (in this case Midjourney) generates new images that will drive the project
<i>S05</i>	<i>3d modeling</i>	unlike the traditional process, this step comes later, modelling takes place on the basis of the generated images, modelling is used to produce technical drawings and diagrams, at the moment we cannot have this precision through AI
<i>S06</i>	<i>post production</i>	post production on Photoshop of created images, adding small details or correcting typical AI errors



S06

image post production (Photoshop)



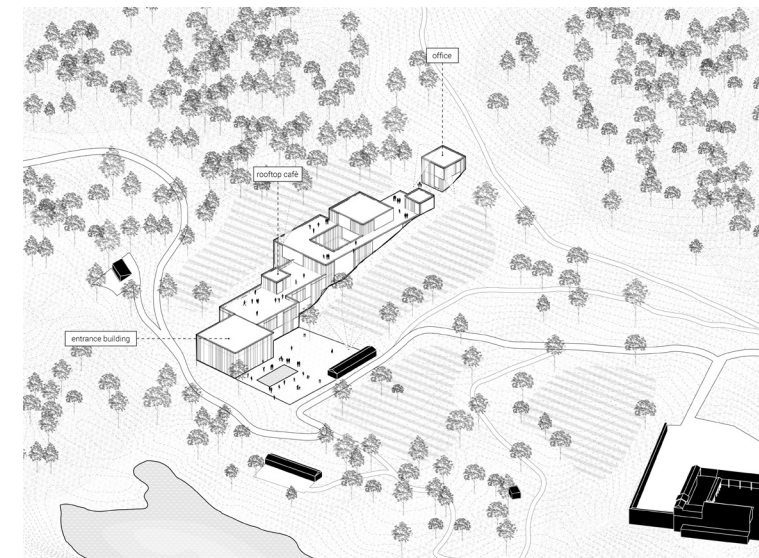
53 // final images after photoshop retouch

Thanks to the case study analysed above, we can say that artificial intelligence has been integrated into the architectural process by changing the order in which the deliverables are produced.

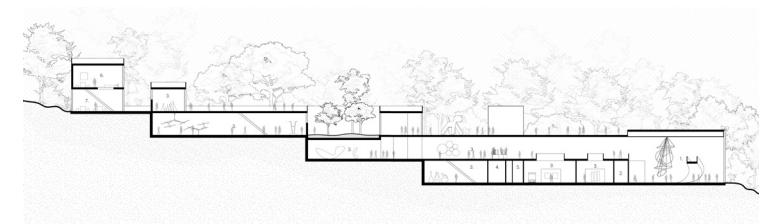
This is an excellent example where the final image is created first and then the 3D model for the production of technical drawings where greater precision is required.

S05

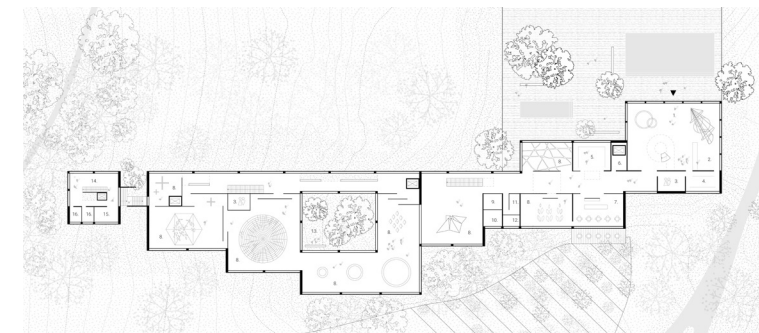
3d modeling



50 // axonometric view of the project



51 // section



52 // plan

What is the traditional architectural process in the early stages of the project ?
Can AI be integrated into the architectural process ?
How does the process change?

AI in the process: Different approaches

Sitography:

// <https://medium.com/@shahramseifcar/how-to-use-ai-in-architecture-and-managing-the-design-process-phases-1116286190ac>
// <https://wc-studio.com/journal/2019/7/5/working-with-an-architect-understanding-phases-of-design-construction>

Bibliography

- // Enjellina, Eleonora Vilgia Putri Beyan, and Anastasya Gisela Cinintya Rossy. "A Review of AI Image Generator: Influences, Challenges, and Future Prospects for Architectural Field." Department of Architecture, Atma Jaya Yogyakarta University.
- // Jaruga-Rozdolska, Anna. "Artificial intelligence as part of future practices in the architect's work: MidJourney generative tool as part of a process of creating an architectural form." ResearchGate.
- // Agkathidis, Asterios, Yang Song, and Symeonidou Ioanna. "AI-Assisted Design: Utilising artificial intelligence as a generative form-finding tool in architectural design studio teaching." ResearchGate.
- // Almaz, Amira Fawzy, et al. "The Future Role of Artificial Intelligence (AI) Design's Integration into Architectural and Interior Design Education is to Improve Efficiency, Sustainability, and Creativity."
- // Fitriawijaya, Adam, and Jeng Taysheng. "Integrating Multimodal Generative AI and Blockchain for Enhancing Generative Design in the Early Phase of Architectural Design Process." National Cheng Kung University.
- // Tan, Linus, and Max Luhrs. "Using Generative AI Midjourney to Enhance Divergent and Convergent Thinking in an Architect's Creative Design Process." *The Design Journal*, 2024.

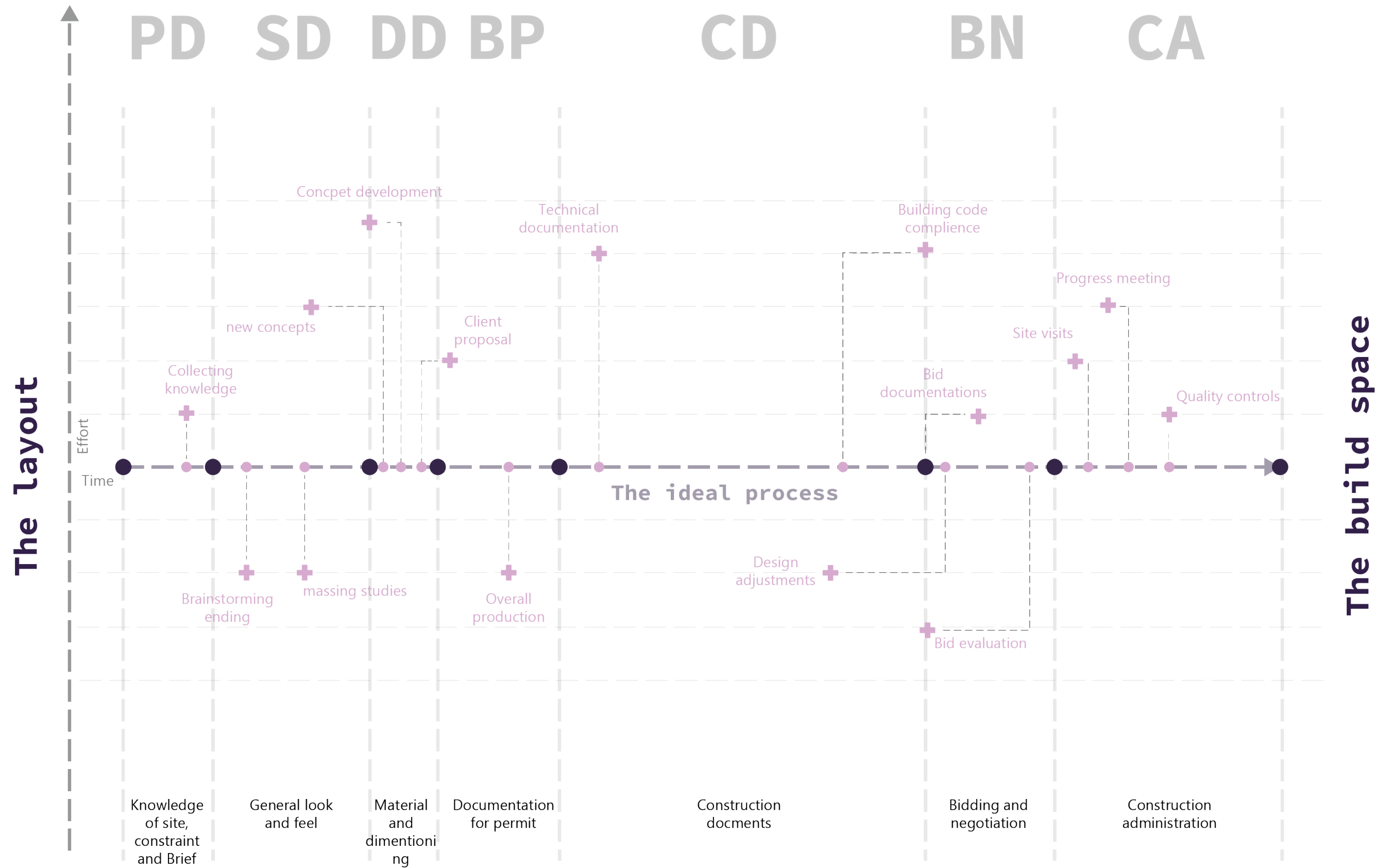
The canonic process

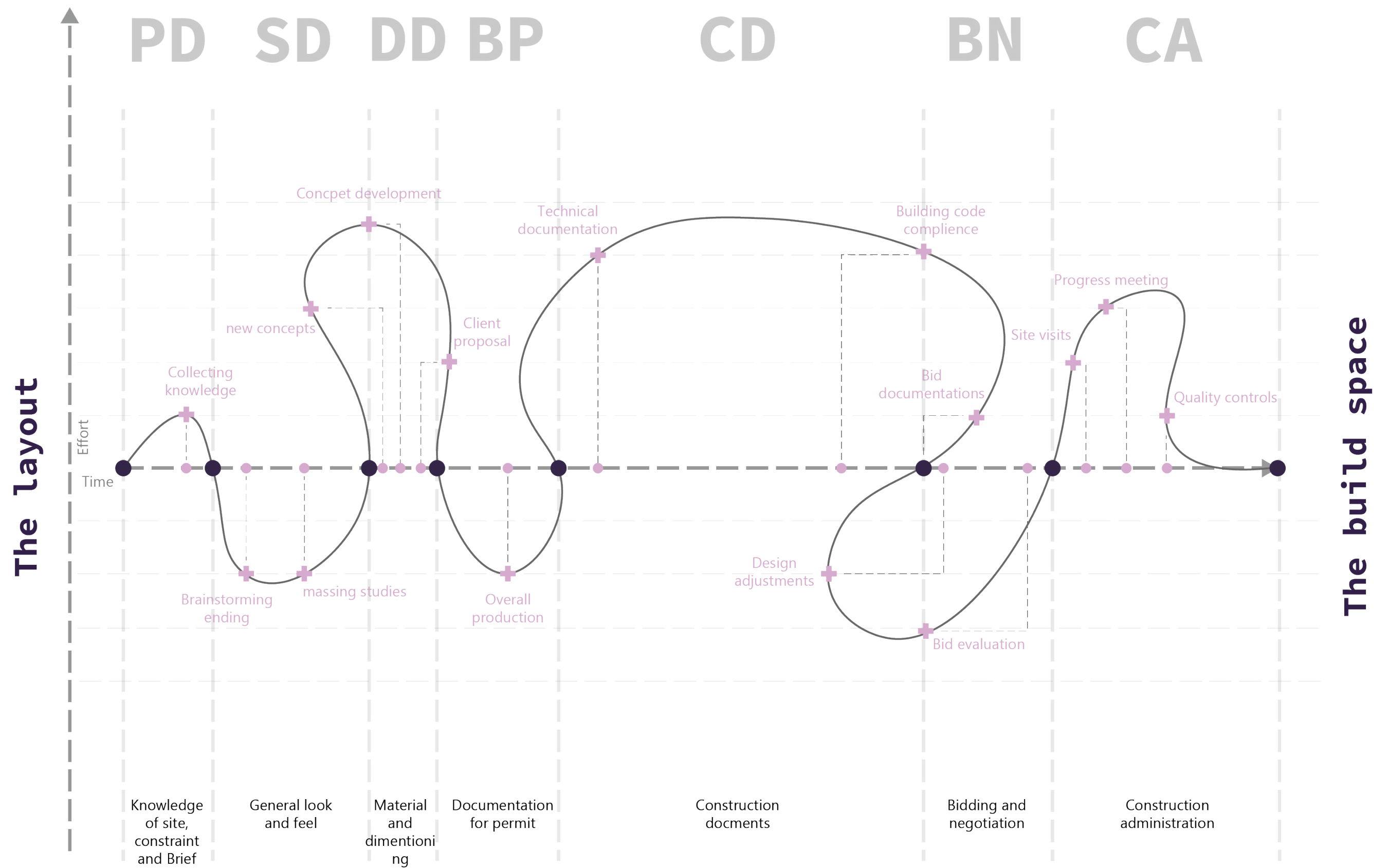
The Dynamic Designing Path

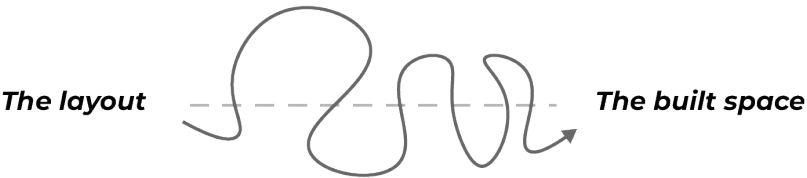
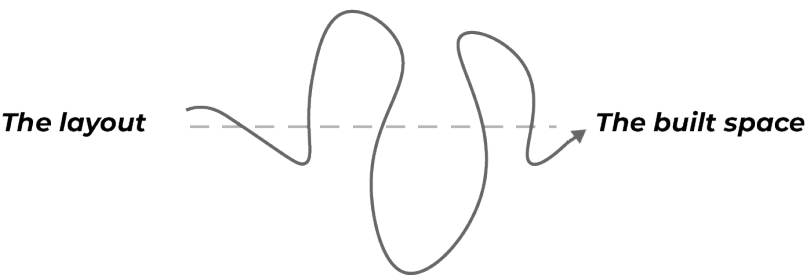
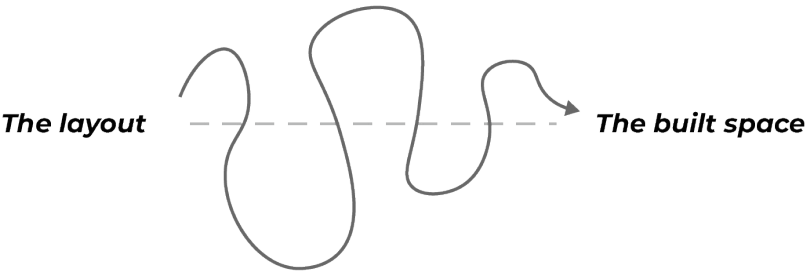
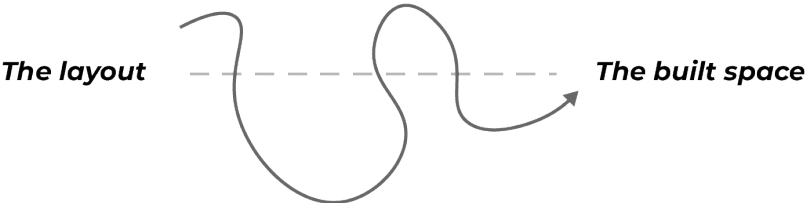
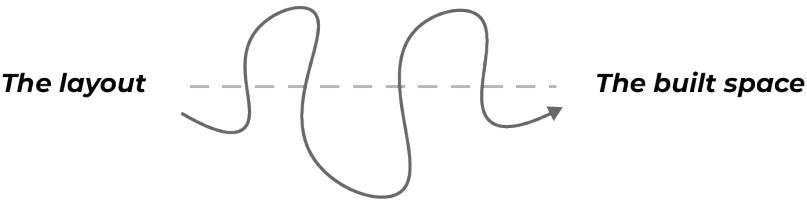
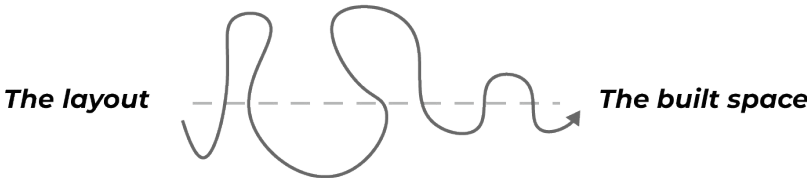
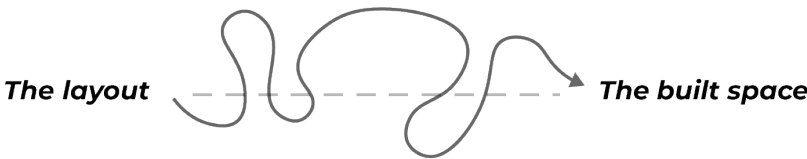
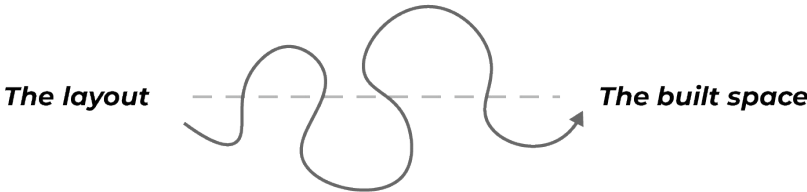
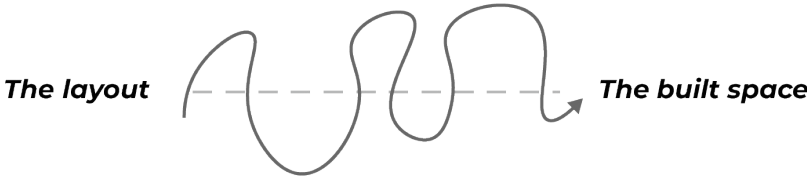
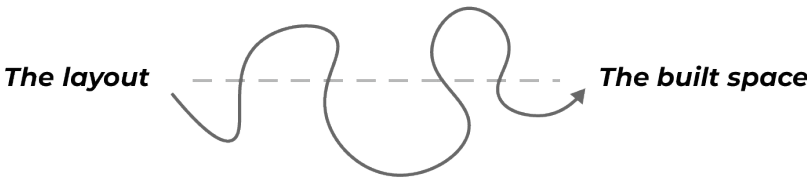
The research explored the process through an experiment in the professional field, examining the multiple guidelines, several functional methods and multiple effective actions of major international offices. Both the initial and final assumptions represent key points in the timeline of this process. Despite a few differences, every project must achieve particular, important milestones or stages. We have already explored these stages defined within the project's multiple phases. These cover the most original in addition to the most executive phases. The diagram offers a simplified view of these importantly complex processes, concentrating only on actions. The organizational analysis fully addresses no stakeholder relationships, budgets, or political considerations.

On the next page, we will examine more in detail how, after defining an ideal linear process, practical realities cause the process to evolve, expanding and stretching over time. This leads to a shift in the planned actions, which are then spread across different phases than originally anticipated.









The Identities within the Designing Path

It is crucial not only to observe the process in its entirety, but also to understand the complexity and specific characteristic of each phase. In fact, each phase has an intrinsic identity that, once understood, can help us develop an appropriate strategy to integrate artificial intelligence within the design process.

Analysing the various phases of the project:

Initial Phase (Preliminary Research)

It is characterised by a linear and canonical process, with a stable methodology and no significant variations regardless of the project.

Phase SD (Schematic Design)

This is the most dynamic and creative phase, where ideas move frantically and, depending on the project, frequent changes in the frequency and fluctuation of workflows can be observed.

Phase DD (Design Development)

In this phase, the process returns to being predominantly linear, but changes in relation to the scalarity of the project, responding to its specific needs.

BP Phase (Building Permit)

This phase is mainly dedicated to the production of technical documentation, and is therefore very linear and rational. Its efficiency depends to a large extent on planning.

Phase CD (Construction Documents)

Here the process is again non-linear and highly customised according to the project. This phase, together with the seventh, is among the most dynamic and adaptable.

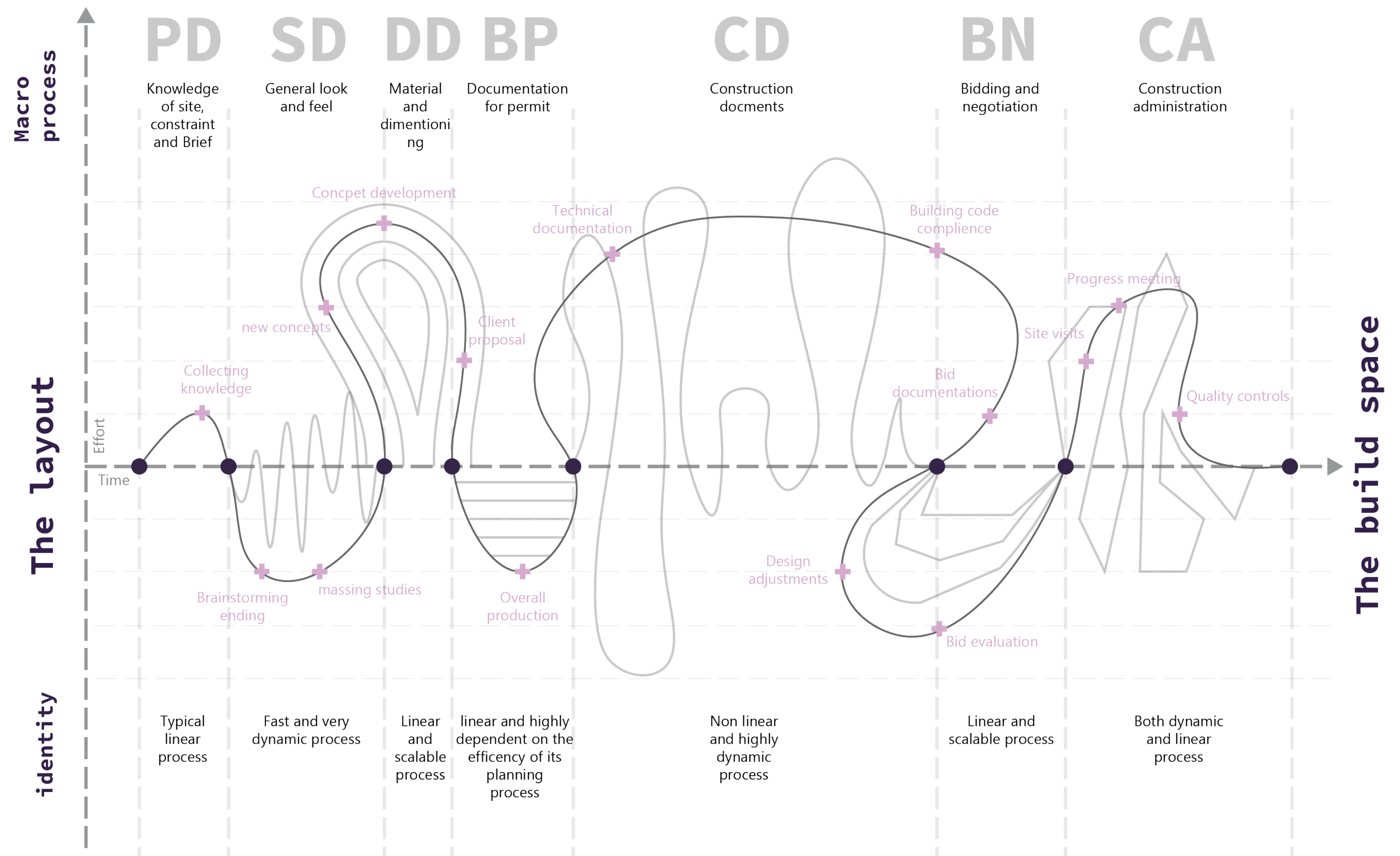
Phase BN (Bidding & Negotiation)

The process becomes linear again, but maintains a certain scalability, as it can be directly influenced by the previous phases.

Phase CA (Construction Administration)

This is a hybrid phase, with half dynamic characteristics, linked to the needs of the construction site, and half linear characteristics, linked to bureaucratic processes and a well-defined schedule.

This analysis highlights how the different phases of the process require customised approaches to integrate artificial intelligence. Understanding the peculiarities of each phase allows us to identify the most appropriate strategies.



AI Integration into the Process

The intuitions

After intensive research and practical experience gained through internships in the offices of MVRDV and UNstudio, as well as through in-depth analysis of interviews conducted, we were able to identify four distinct approaches to integrating artificial intelligence into the canonical design process.

These insights are based on a general guideline in the design approach, in which AI can intervene in the traditional path in various ways: through addition, subtraction, reversal or even total modification of the conventional process.

The insights we present are theoretical; the aim is to explore them in the following chapters through small pilot projects that demonstrate their potential. These projects will serve to highlight the freedom of creating new types of processes, changing the relationship between client and author, communication methods, design approach, idea generation, and speed of execution, without compromising effectiveness.

As already mentioned, the proposed strategies explore two main directions: a radical revolution of the process, capable of overturning it or even developing the design through a simple dialogue with the AI, and a more targeted intervention, aimed at improving the traditional process, emphasising the creative identity of the early design phases.

Therefore, insights can be grouped into two macro-categories:

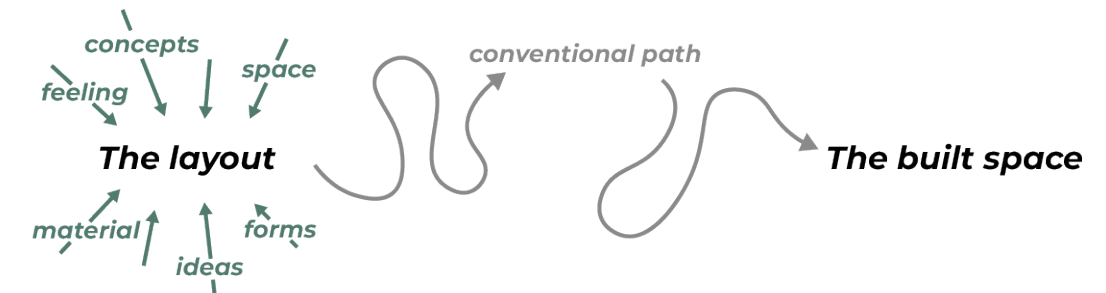
THE HELPERS, punctual actions in the process

These focus on supporting and enhancing the initial phases, where the creative and dynamic approach plays a crucial role.

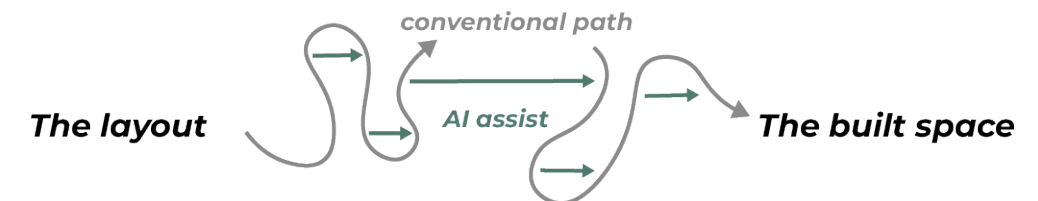
THE SHAPERS, strategies for the entire process

These have a global approach, aiming to redesign the process in its entirety without limiting it to the specific identities of individual phases.

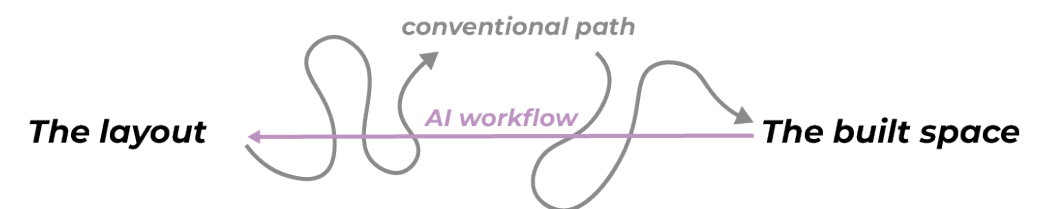
HELPER 01: The conceptualizer



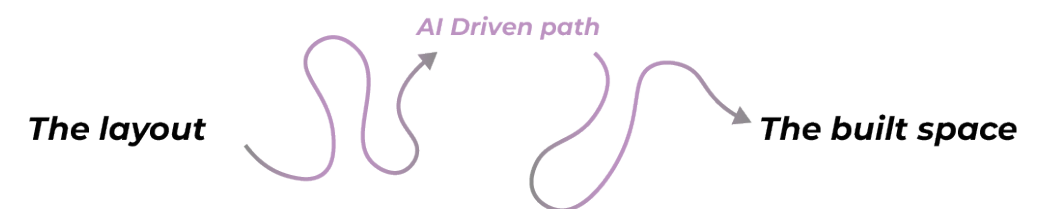
HELPER 02: The assistant



SHAPER 01: The inversion



SHAPER 02: The dialogue

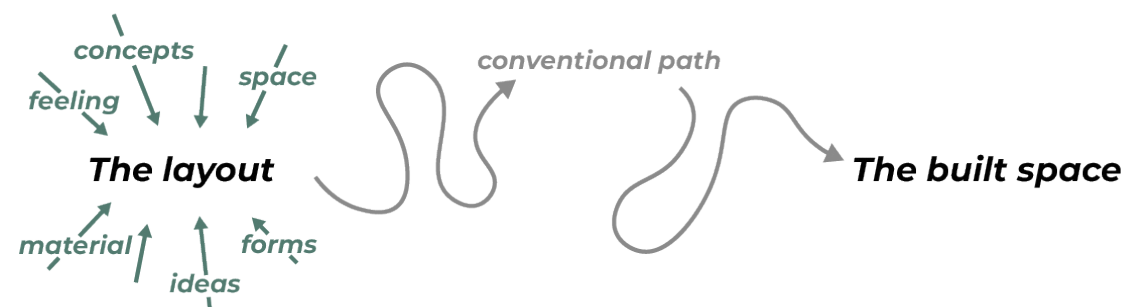


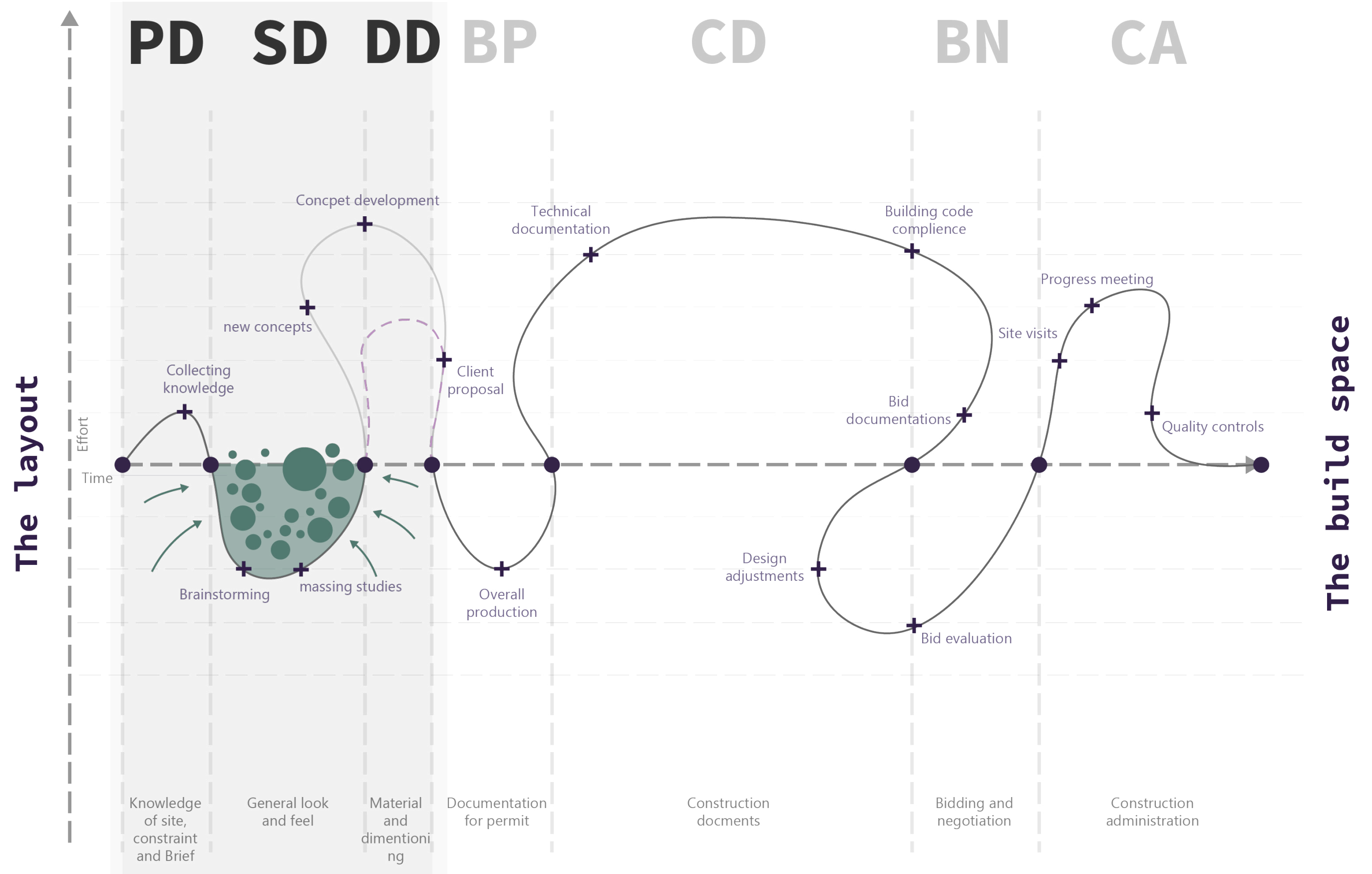
HELPER 01: The conceptualizer

This process aims to use the current software as an additional engine only in the conceptualisation phase, helping to conceive architecture through new perspectives rather than using the software merely as a tool to produce what we are already capable of producing ourselves.

Similar to the inversion, this one serves as a reference for a great number of images, each of them having specific characteristics different from the others, but all together constructing a unique complex in the mind of the architect. So, control of the end result and space is fully in the hands of the architect, who decides by arbitrary choice and also based on classic studies on what can work and what cannot.

Rather, this process tends to boost and assist the creativity of the architect himself in getting new perspectives toward arriving at original and innovative solutions.



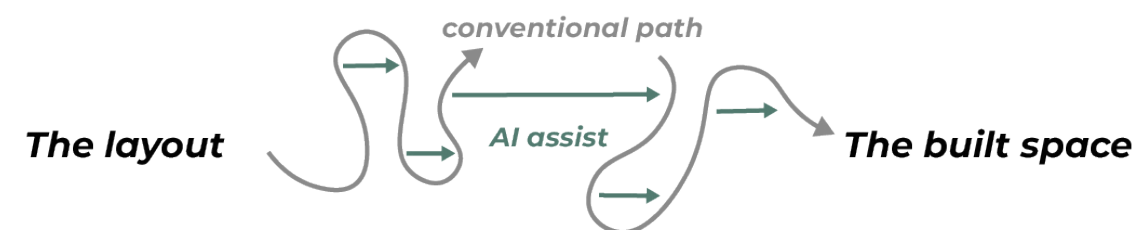


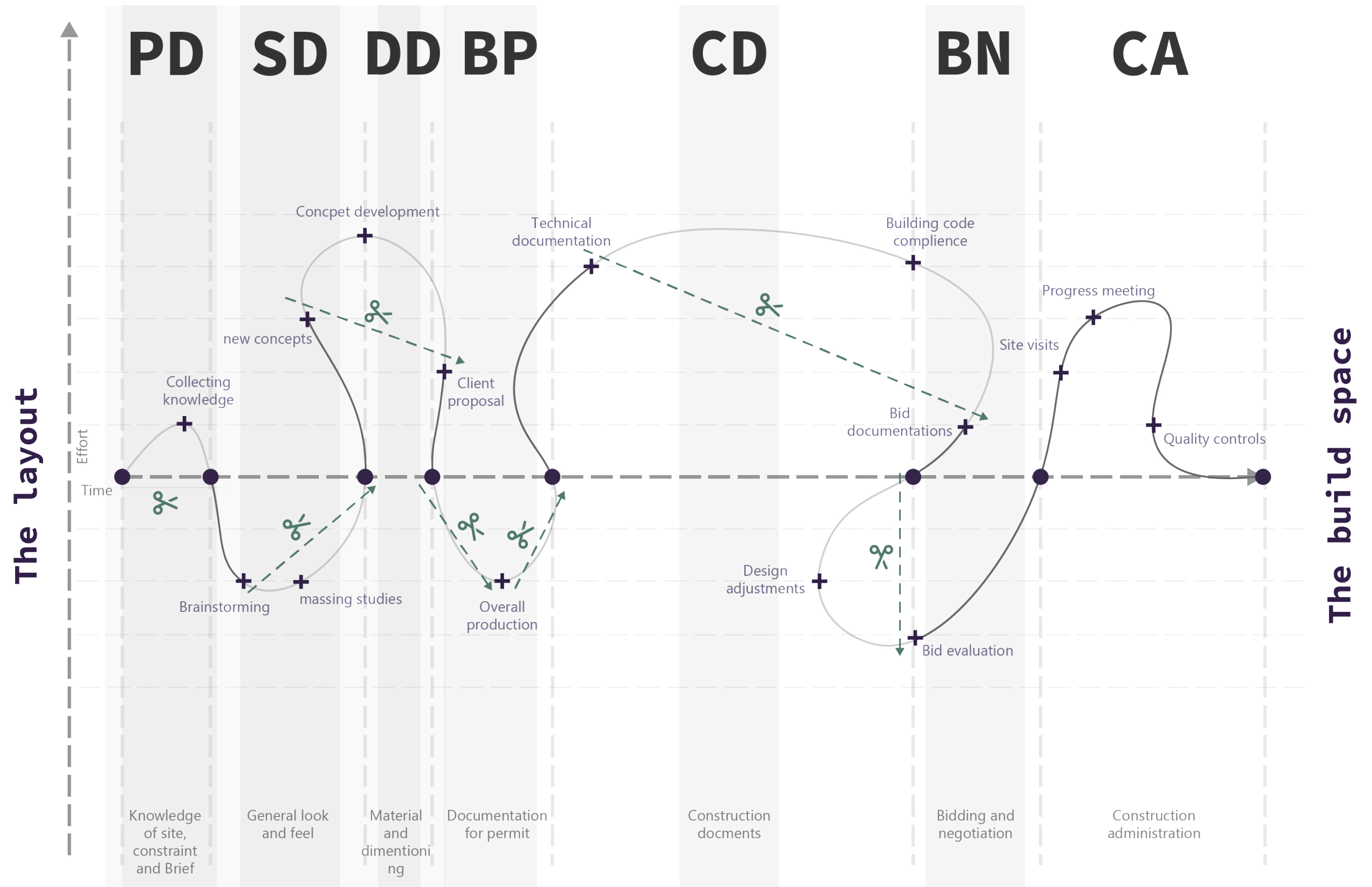
HELPER 02: The assistant

This process, on the other hand, aims to have continuous AI assistance at every step through various tools. These tools can be implemented within the existing software or as additional software, focusing solely on speeding up the process and making it more efficient.

This process therefore impacts the entire system, and this is done through the use of different agents. These are put into action during crucial stages of the process that are repetitive, tedious, and very rational. This increases the time spent in the most creative phase and thus increases the quality of the project itself.

So, unlike the previous one, this aims at including several small interventions by AI in the process, aimed at making it optimal and quicker, at the same time giving the full control of the project, in the design and creative phases, back to the architect.

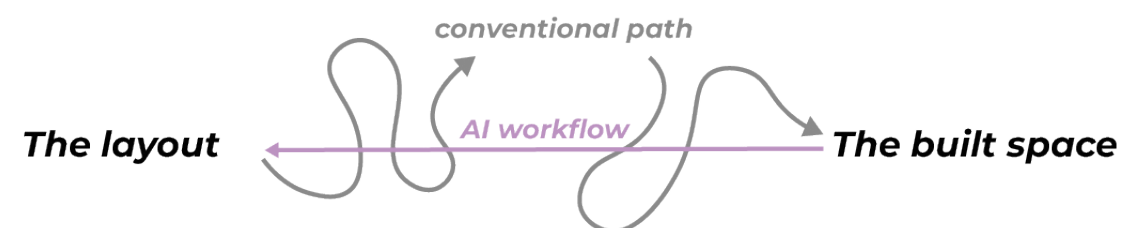


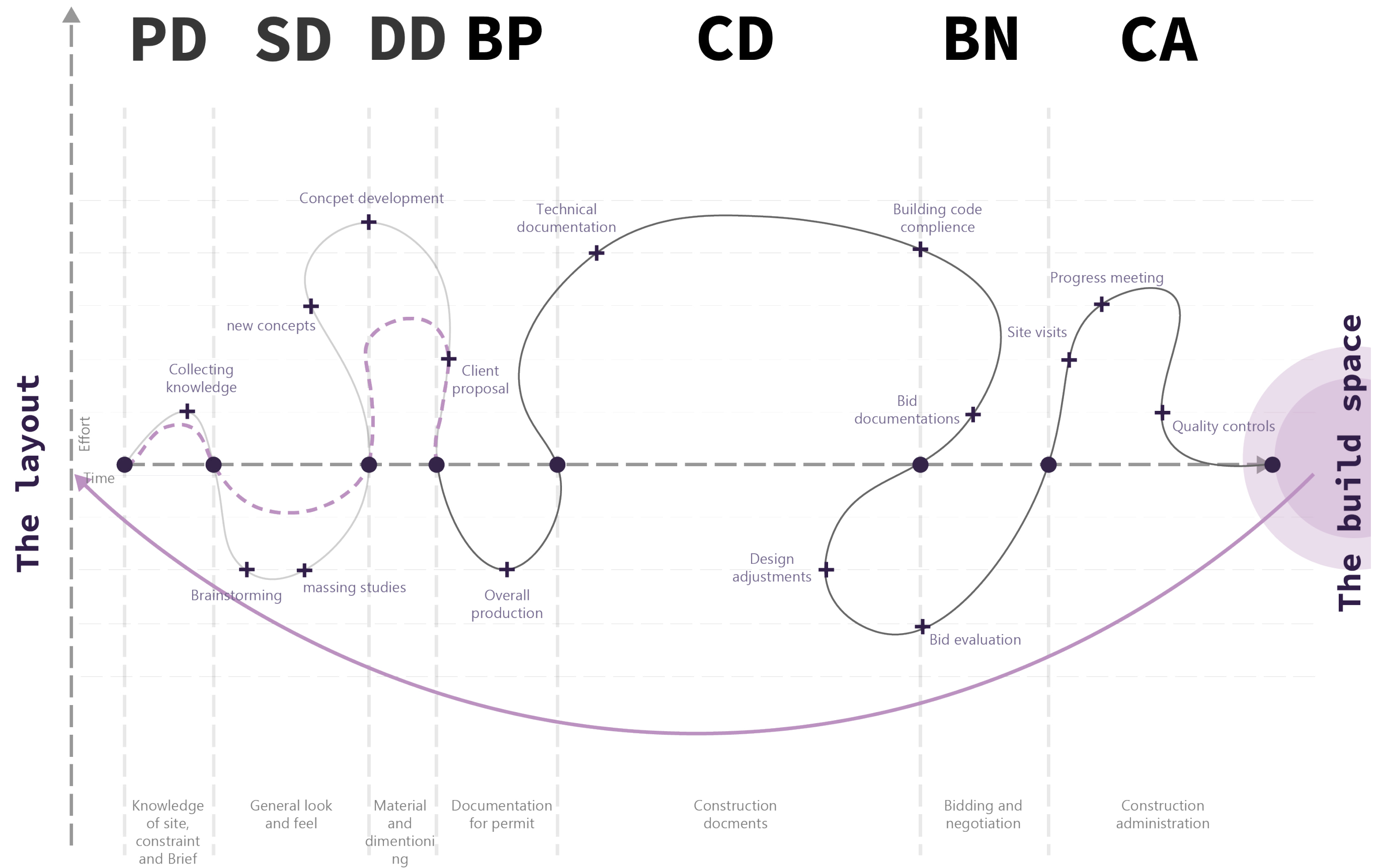


SHAPER 01: The inversion

This process aims to reverse the typical approach used in architectural offices. It starts with images depicting the internal and external spatiality of the building and the near-final result of the product, then works backwards to the composition of the 3D model and its derivatives: plans, sections, and elevations.

The developed process stems from intuition gained in the profession and experimentation of the tool within the profession, evolving after a particularly evocative period of brainstorming, sometimes wrapped in an idea that represents to the letter the evocations we have given to artificial intelligence. The working through Prompts and trying to bring the Diffuse Model to a very high degree of conditioning can result in a final image of high quality, either in design or representation. So, once the image has been developed with the spaces clearly defined, the actual project development is to be carried out by referring to that particular image or that set of generated images. This testifies to the freedom of expression that the use of such a powerful tool can impart when applied appropriately and with the right knowledge.



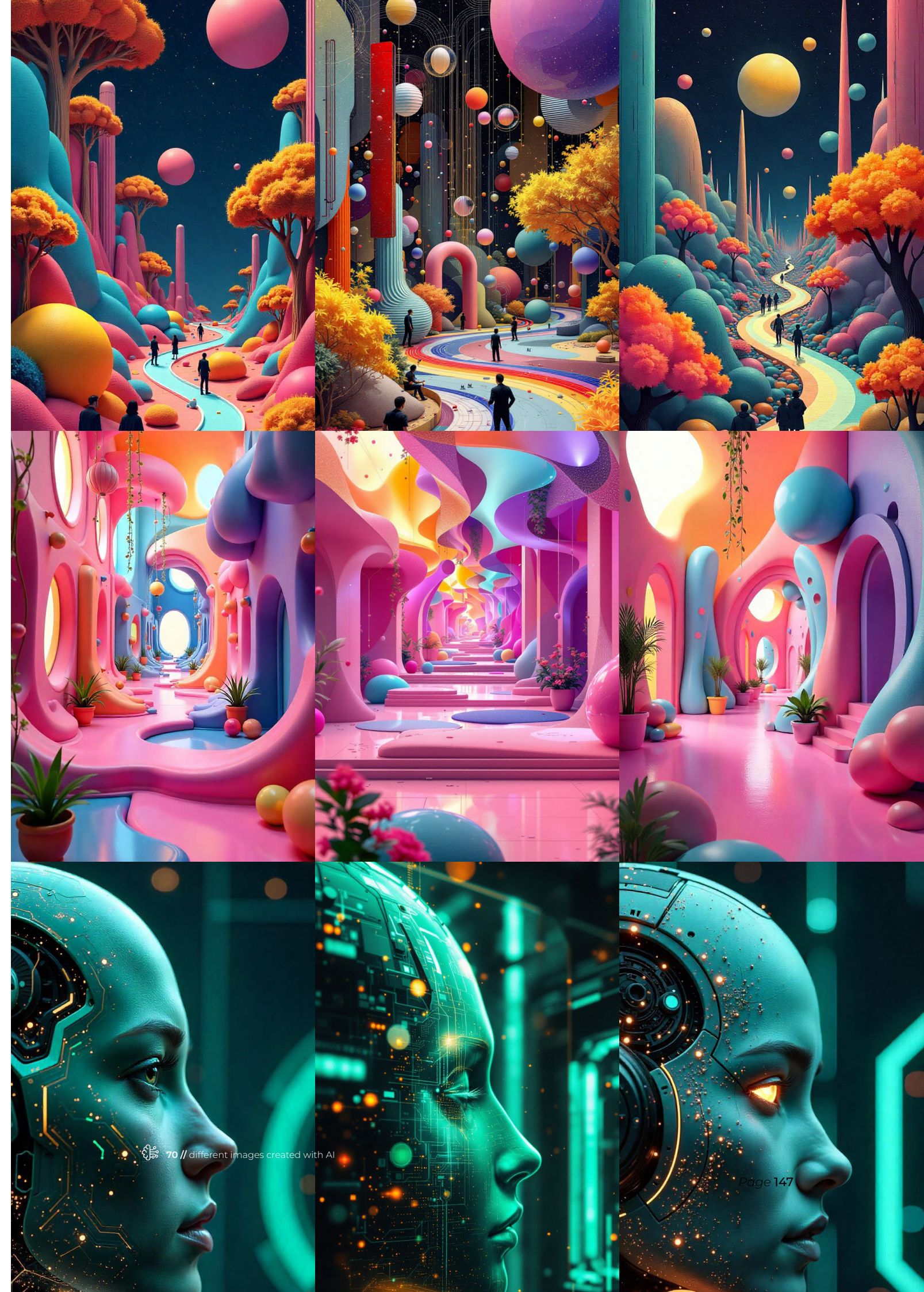
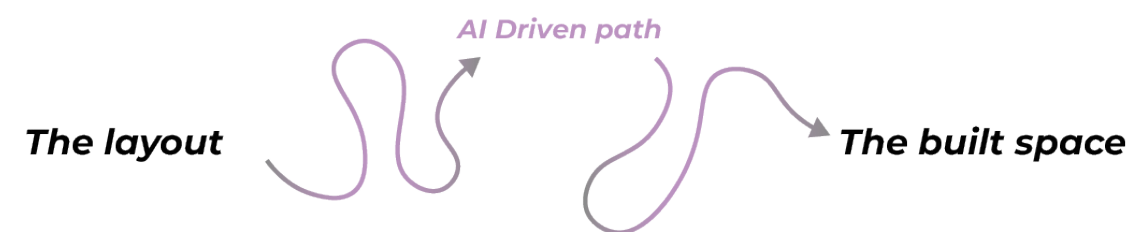


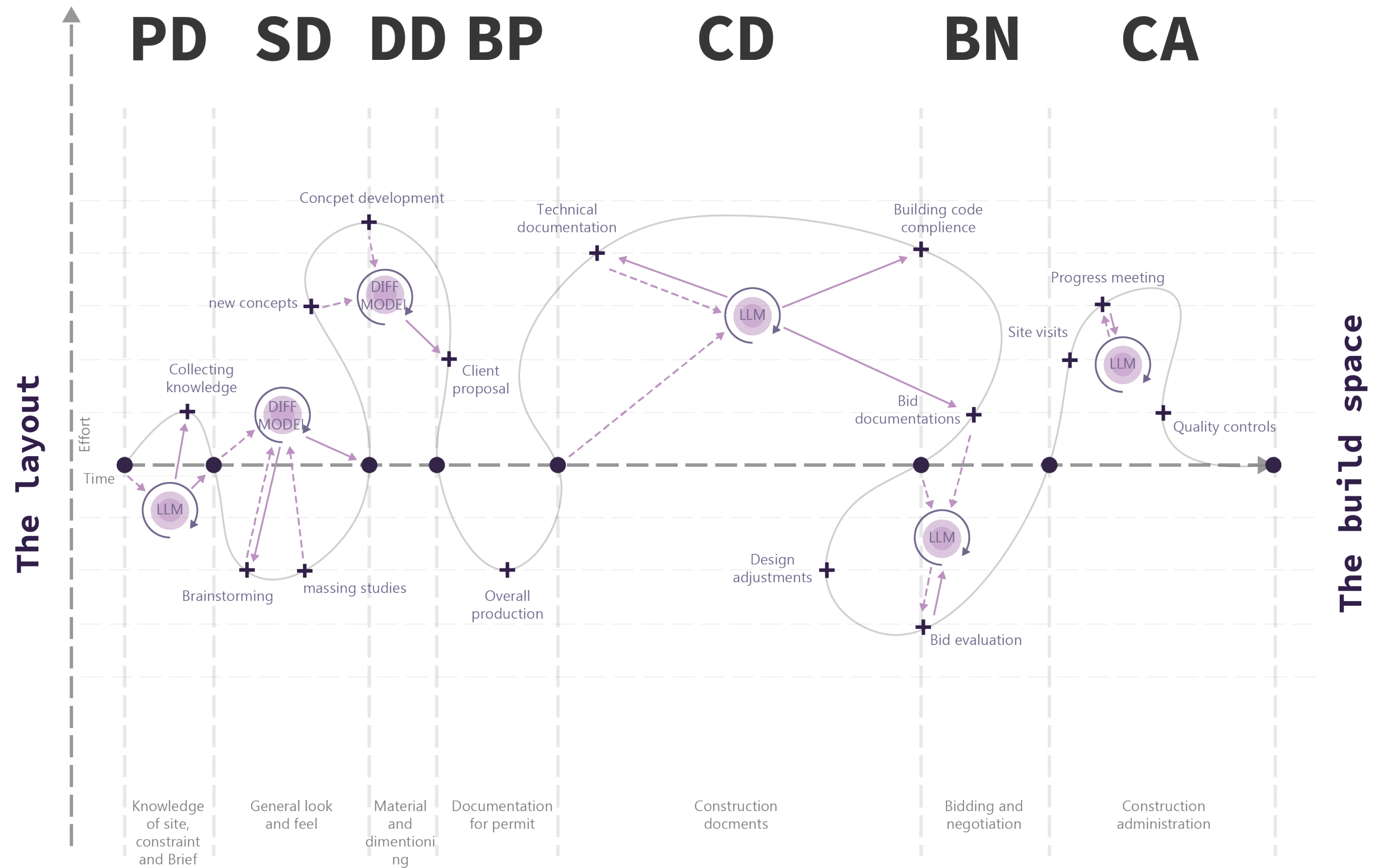
SHAPER 02: The Dialogue

This process involves research that empowers the process through a constant dialogue (ping pong) between the human intentions and the artificial creativity.

This alternative process deviates from assistant in dramatically change the canonical one, occurring within the continuous changes between the designer and the AI, in which both subjects are placed on the same level, both actors contributing to the developing the desing proposal. Experimentation and strong change are difficult to analyze within the overall process. Further we will demonstrate how this implementation can be archived.

If approached correctly, this customizes the process to conform with the design underway and would allow a high degree of originality and out-of-the-norm design. Instead, If not managed correctly it can lead to a veryt generic and superficial design.





How can we use artificial intelligence in architecture ?
What is ComfyUI and how does it work ?
How much influence can we exert on artificial intelligence ?

The tools: capabilities and limitations

Sitography

// <https://github.com/comfyanonymous/ComfyUI>
// <https://stable-diffusion-art.com/how-stable-diffusion-work/>
// <https://github.com/tencent-ailab/IP-Adapter>
// <https://www.labelerr.com/blog/unveiling-florence-2-microsofts-cutting-edge-vision-model-for-diverse-ai-applications>
// <https://www.tripo3d.ai/>
// <https://civitai.com/>
// <https://blenderneko.github.io/ComfyUI-docs/Core%20Nodes/Sampling/kSampler/>
// <https://comfyui-wiki.com/comfyui-nodes/sampling/k-sampler>

Bibliography

// u Osijeku, Sveučilište Josipa Jurja Strossmayera. "Text-to-image Stable Diffusion model."
// Ho, Jonathan, Ajay Jain, and Pieter Abbeel. "Denoising diffusion probabilistic models." Advances in neural information processing systems 33 (2020): 6840-6851

ComfyUI

The tool

Artificial intelligence allows designers and architects to create unique images that help them in their creative process. All this is supported by tools that enable image creation and manipulation. There are various good software tools on the market such as Stable Diffusion, DALL-E 2, Midjourney, and ComfyUI²⁵.

This thesis is mainly concerned with the use of ComfyUI as the main software.

ComfyUI is an open-source platform built on the control of artificial intelligence pipelines related to image creation. The interface, it is very similar to Grasshopper, as it is based on the Visual Scripting type of software.

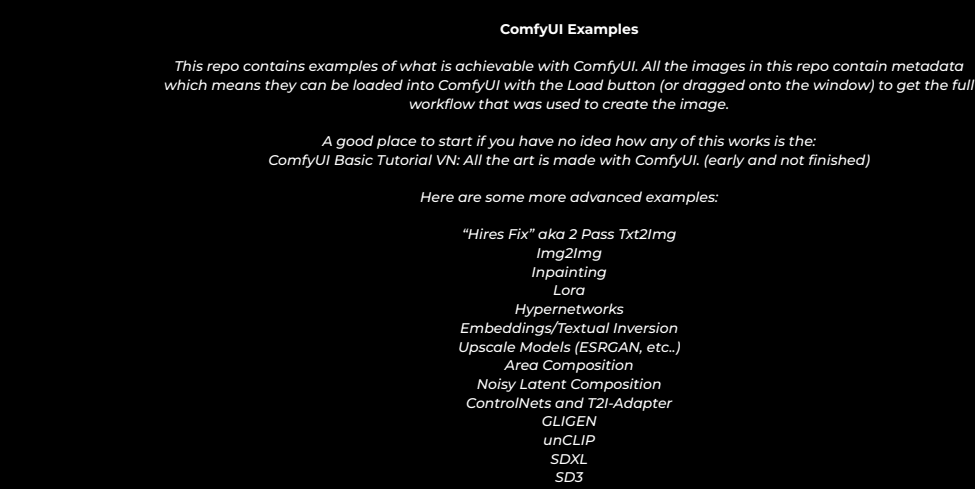
The most powerful and modular
diffusion
model GUI and
backend.

²⁵developed in 2023 by comfyanonymous to facilitate the modular use of AI models like Stable Diffusion
²⁶a platform that uses Git to help developers store, manage, and collaborate on source code. It offers tools for project management, issue tracking, and code reviews, and is used for both open-source and private projects
²⁷source : <https://civitai.com/>

The choice of ComfyUI was primarily driven by its accessibility, as it is completely free and available to everyone. The only limitation is related to the hardware power required.

Online, an enormous and readily accessible collection of resources allows you to craft fascinating visuals, exploring diverse image-creation methods and optimizing your production pipeline. GitHub's²⁶ ComfyUI section provides a thorough overview that details the software's many capabilities and several integrated workflows. Civitai²⁷ offers a thorough open-weight platform for exploring and using a huge array of AI models, providing readily available access to many pre-trained models. It is analytically important to note, however, that all of Civitai's source code is not publicly available for any modification or redistribution; therefore, the platform is fundamentally not open-source.

All this allows anyone to start experimenting in a relatively short time without the need for any specific skills or expertise, although the first steps may be more difficult than expected. Compared to other programmes, it is not as user-friendly and it may take some time to get to grips with its mechanisms, even if only in terms of the installation process, but for any problems you can find tutorials and guides online.



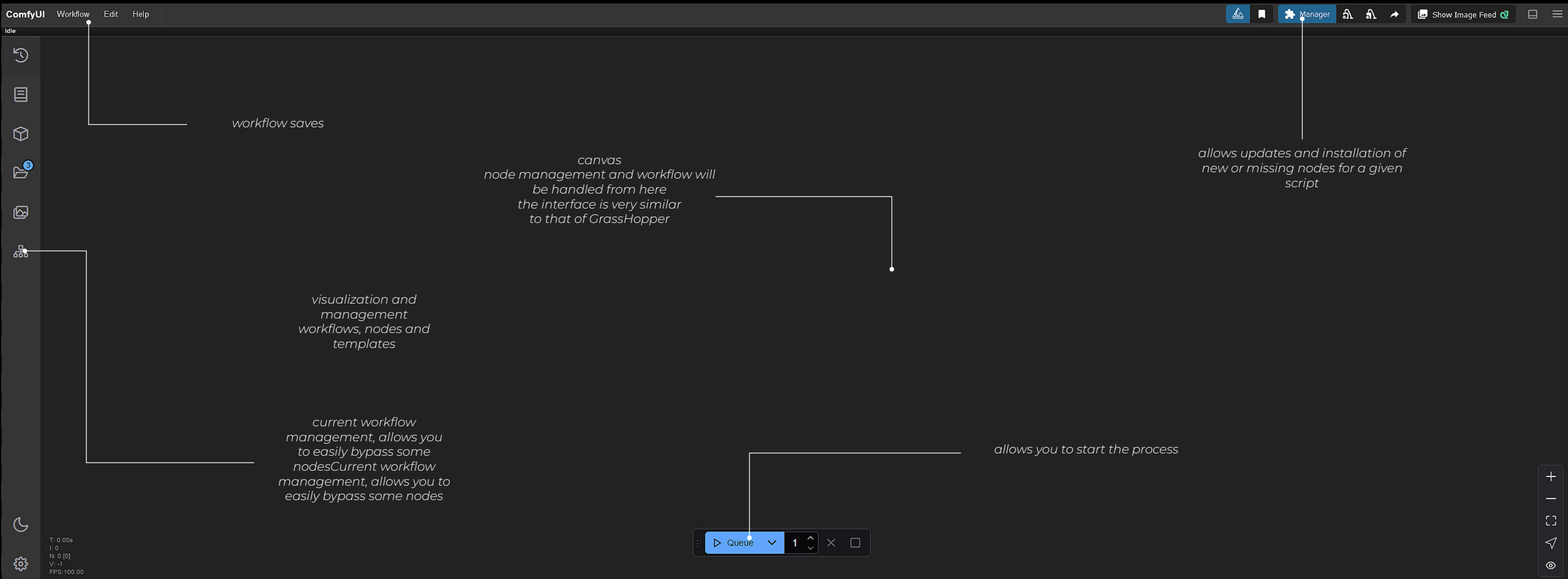
Interface

By mastering ComfyUI, users are able to have more control when creating images, its total freedom allows for an enormous amount of editable variables, and being an open-source programme, it is easy to find workflows and models generated by other users.

It has a similar interface to Grasshopper, but of course they have completely different functions (later we will see some ways to relate these programmes and thus create an interconnected process).

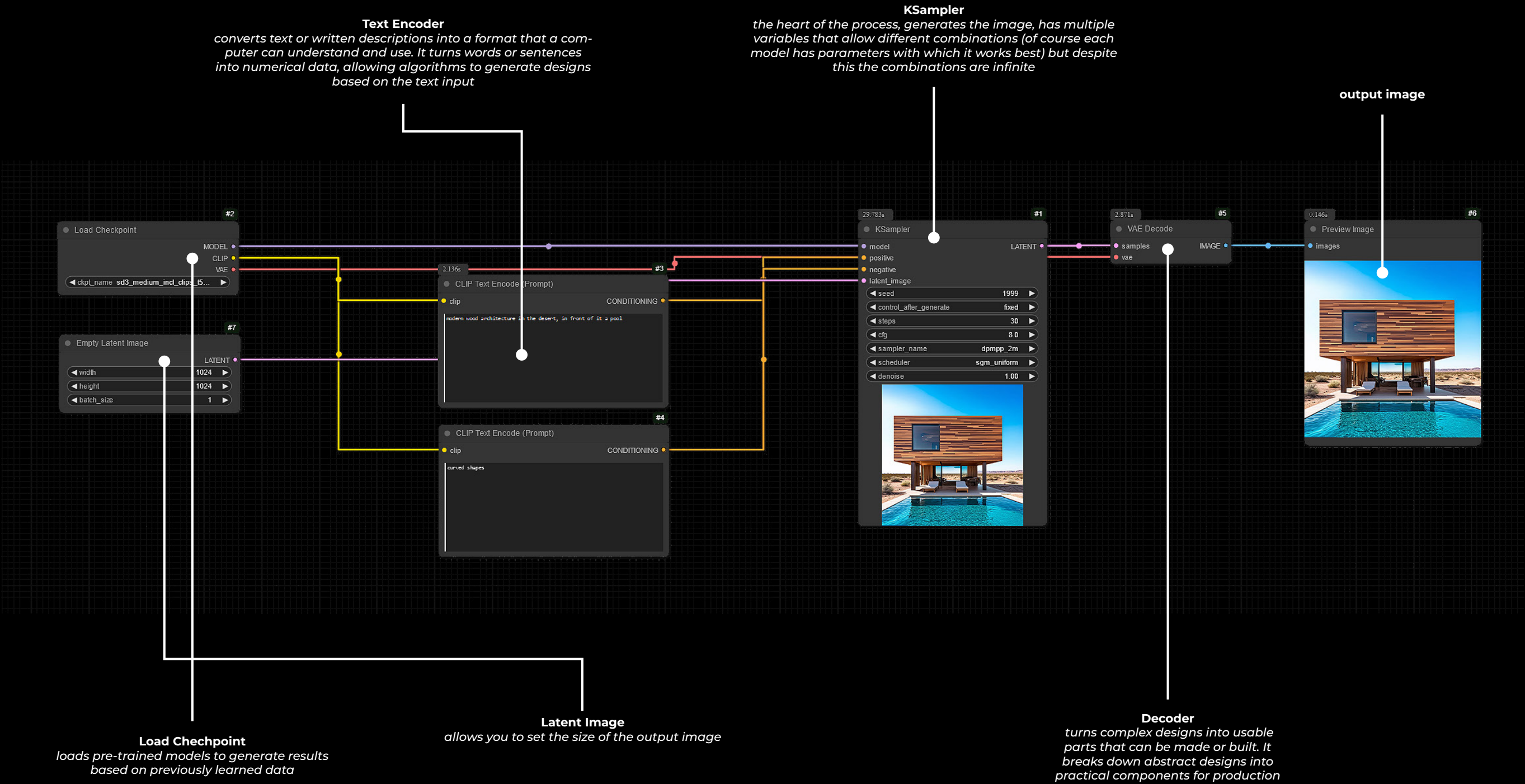
It has a modular structure that allows easy customisation by the user and the possibility of testing different processes without having to completely rewrite the code base.

It is based on the use of deep learning models for image generation; the user can choose to use previously trained machine learning models or to train his own model.



73 // interface of ComfyUI (2025)

Main Commands and Nodes



KSampler "The image generator"

The core of Image Generation

The KSampler uses the provided model and the conditionings to generate a new version of the given latent. First the latent is noised up according to the given seed and denoise strength, erasing some of the latent image, then this noise is removed using the positive and negative conditioning as guidance, “dreaming” up new details in places where the image was erased by noise.

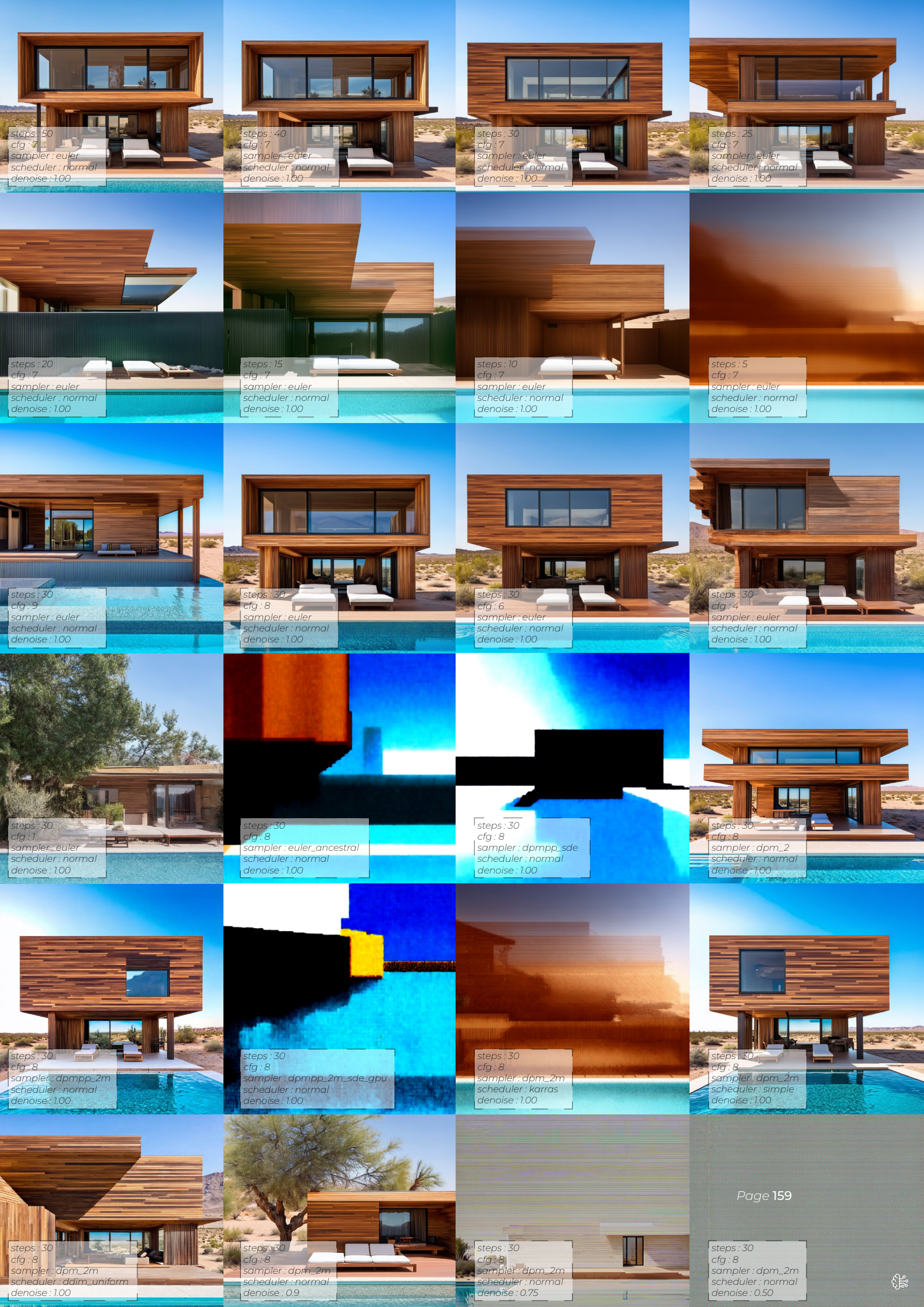
There are several variables that influence the generated image, before obtaining good results it is necessary to check these parameters and understand how the chosen model performs best, moreover model developers usually suggest optimal values.

model	specifies the generative model to be used for sampling, playing a crucial role in determining the characteristics of the generated samples
seed	controls the randomness of the sampling process, ensuring reproducibility of results when set to a specific value
steps	determines the number of steps to be taken in the sampling process, affecting the detail and quality of the generated samples
cfg	adjust the conditioning factor, influencing the direction and strenght of the conditioning applied during sampling
sampler_name	selects the specific sampling algorithm to be used, impacting the behavior and outcome of the sampling process
scheduler	chooses the scheduling algorithm for controlling the sampling process, affecting the progression and dynamics of sampling
positive	defines positive conditioning to guide the sampling towards desired attributes of features
negative	specifies negative conditioning to steer the sampling away from certain attributes or features
latent_image	provides a latent space representation to be used as a starting point or reference for the sampling process
denoise	controls the level of denoising applied to the samples, affecting the clarity and sharpness of the generated images

75 // main nodes and terms on ComfyUI

Next, a study is shown to understand how the various inputs influence the image; for convenience and consistency of the study carried out, the same seed has been kept for all generations, while the model used is SD3.

The prompt used is: modern wood architecture in the desert, in front of it a pool.



Control Net "The Guidelines"

Workflow Explanation

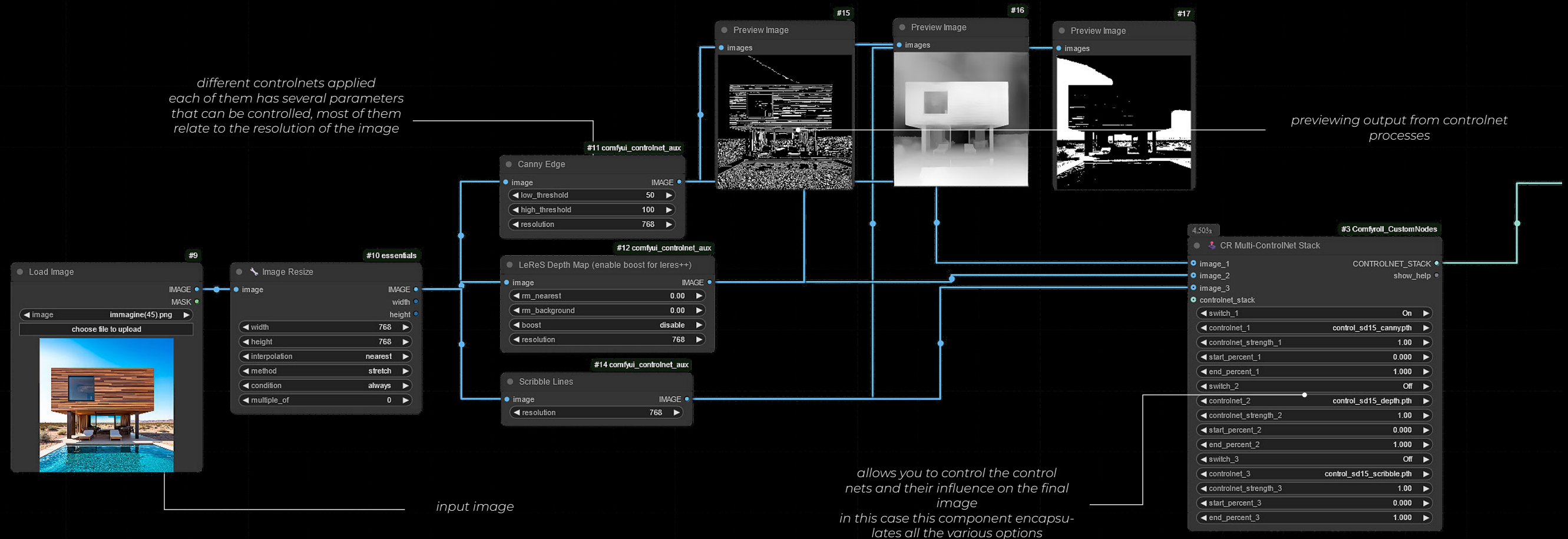
ControlNet is an extension within ComfyUI that assists in improving image generation. Specifically, it allows for more control over the image through specific inputs or conditions.

ControlNet allows to control the image through multiple ways and as far as the architectural field is concerned the most used are:

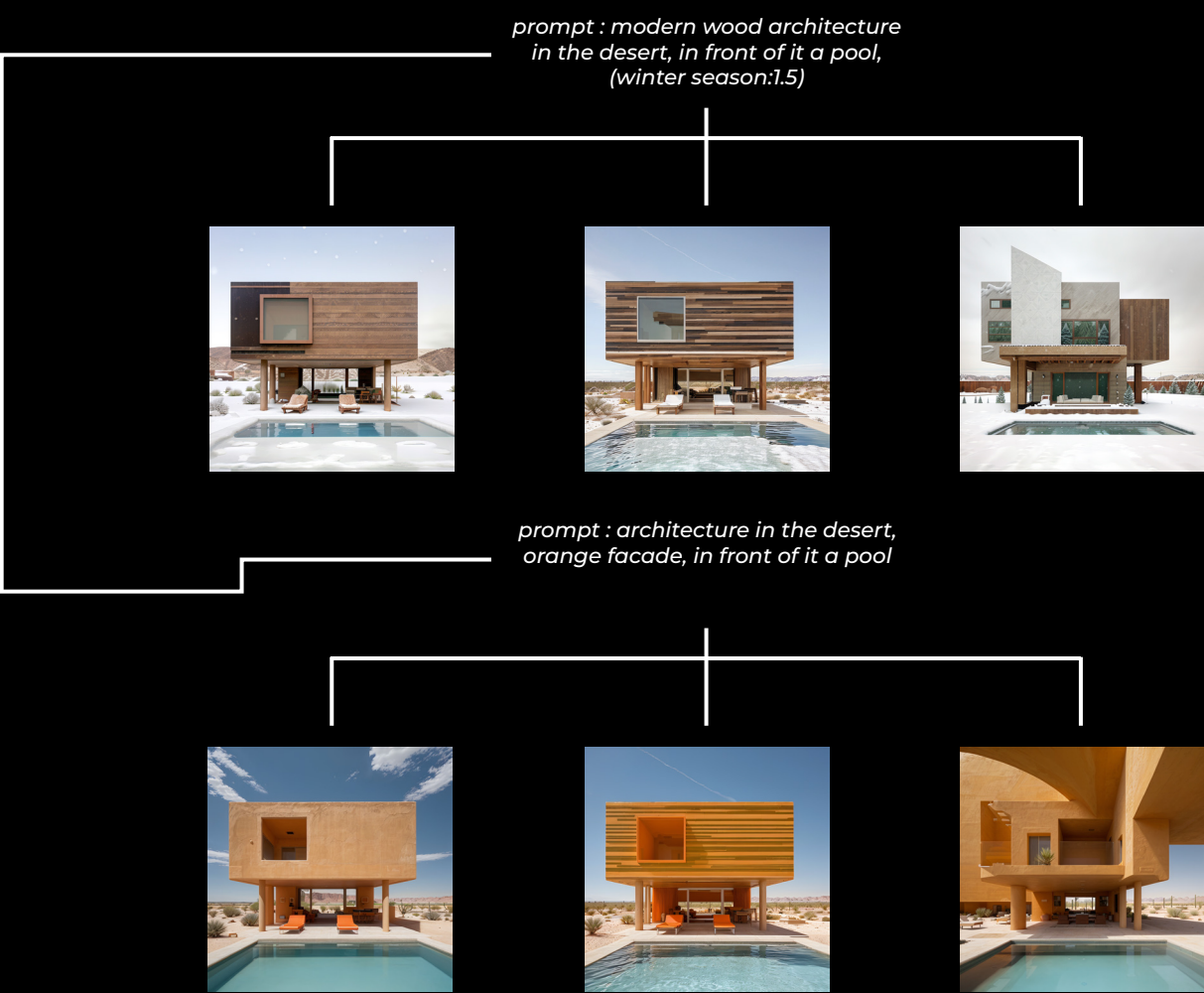
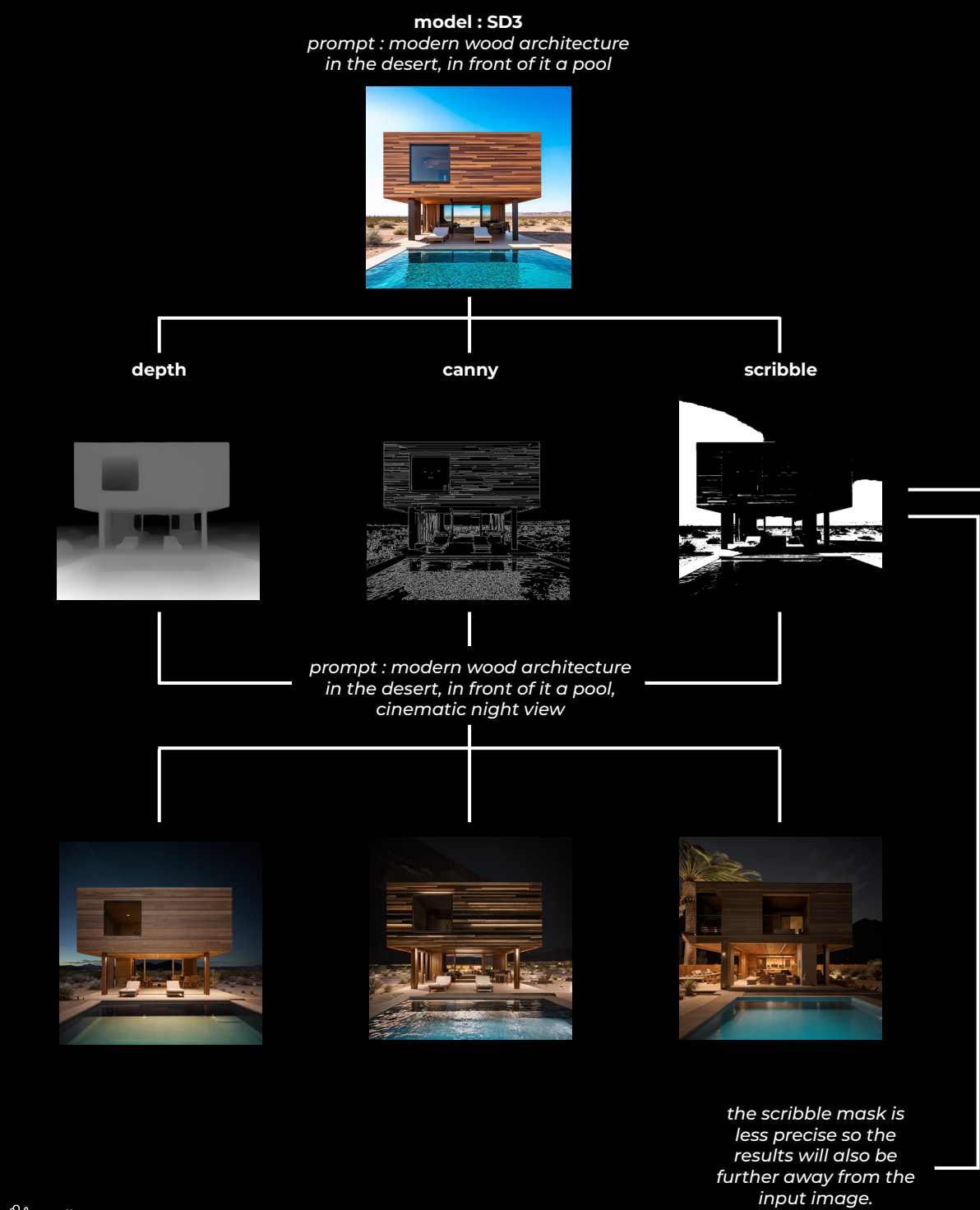
- scribble (image line detection)
- depth mask (depth of elements in the image)
- canny edge (more precise image line detection)

To understand how it works, a study was done with a starting image (for convenience the previously generated image was taken) that was controlled through ControlNet to get a different output. In this case it was decided to change the mood of the image by changing the time of day represented. The possible variations of course are many, from changing the materials represented to the lighting of the final image.

ControlNet is extremely powerful and is used in most workflows, and it is also possible to use more than one input at the same time to have more control over the image.



Iterations



In this specific study it was decided to modify the image by intervening on the time of day, season and materiality, the results obtained are acceptable in some cases and less so in others.

We see how with CannyEdge we can have an image almost identical to the original in shape and spatiality, while with the other components we lose more of the influence we have on the image, but this depends from image to image and it may be that with other inputs the other ControlNet components are more suitable.

canny edge



architecture in the desert, red facade, in front of it a pool



architecture in the desert, white facade



(black architecture:1.3) in the desert, in front of it a pool



concrete house, brutalism



modern house in the desert, (winter season:1.3), (snow:1.5)



architecture in the desert, sunset



architecture in the desert, grass in front



architecture in the desert, yellow facade, in front of it a pool



(green:1.3) architecture in the desert



(copper facade:1.3), modern architecture in the desert



architecture in the desert, cinematic, black and white



(red house:1.2) in the desert, night, cinematic, realistic



architecture in the desert, night, cinematic



(orange and black house:1.25) in the desert, pool



(blue house:1.2) in the desert, pool, light blue sky



(dark blue:1.3) architecture, (snow:1.3), winter season

depth anything



classic (brick facade:1.2), sunset, (cinematic:1.2)



glass and red concrete house



modern house (on a platform on the water:1.5)



white house, (le corbusier style:1.3)



modern house (in the forest:1.3)



blue house, (fog:1.3)



white (lamella:1.5), (forest:1.2), trees



yellow architecture in the night, cinematic view, (yellow:1.1) and



(orange house:1.3)



modern house (in a forest:1.2), (winter season:1.3), snow



blue house in the mountains, rgb(0,0,205)



(parametric facade:1.2), forest



modern house, (neon style:1.1)



minimal house, night view



steel house, (mountains in the background:1.3)



colorful house, (multicolor:1.2)

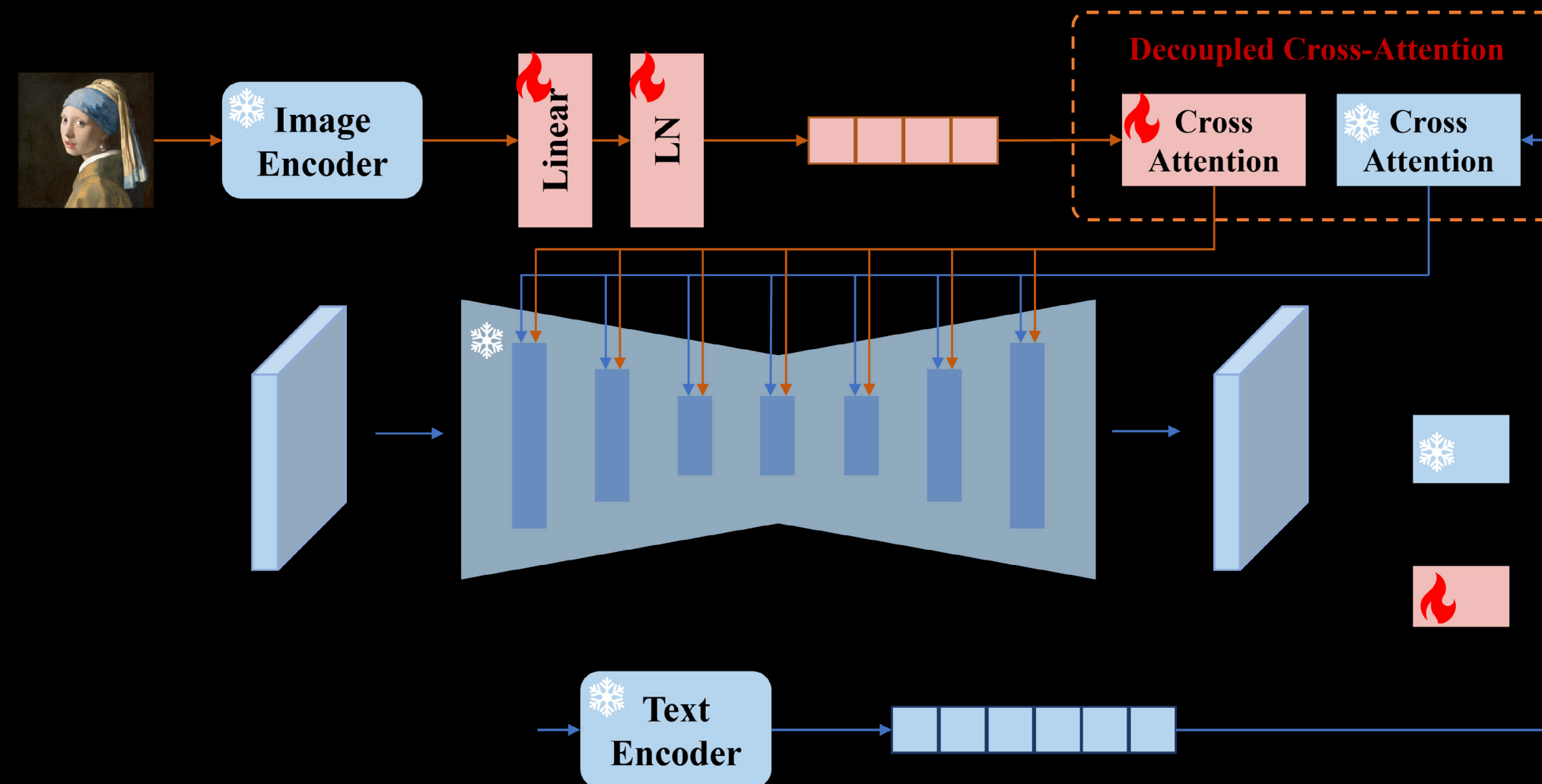
IPAdapter "The Style"

General understanding

The image prompt adapter is designed to enable a pretrained text-to-image diffusion model to generate images with image prompt. The proposed IP-Adapter consists of two parts: a image encoder to extract image features from image prompt, and adapted modules with decoupled cross-attention to embed image features into the pretrained text-to-image diffusion model.

The starting point is the image encoder which processes the initial image, in this case a portrait, and extracts its main features called 'image features' through a pre-trained and frozen architecture in order to obtain a stable result.

The text Econdor as already stated in the previous chapters deals with the analysis of a textual prompt thus obtaining 'text features' as output.



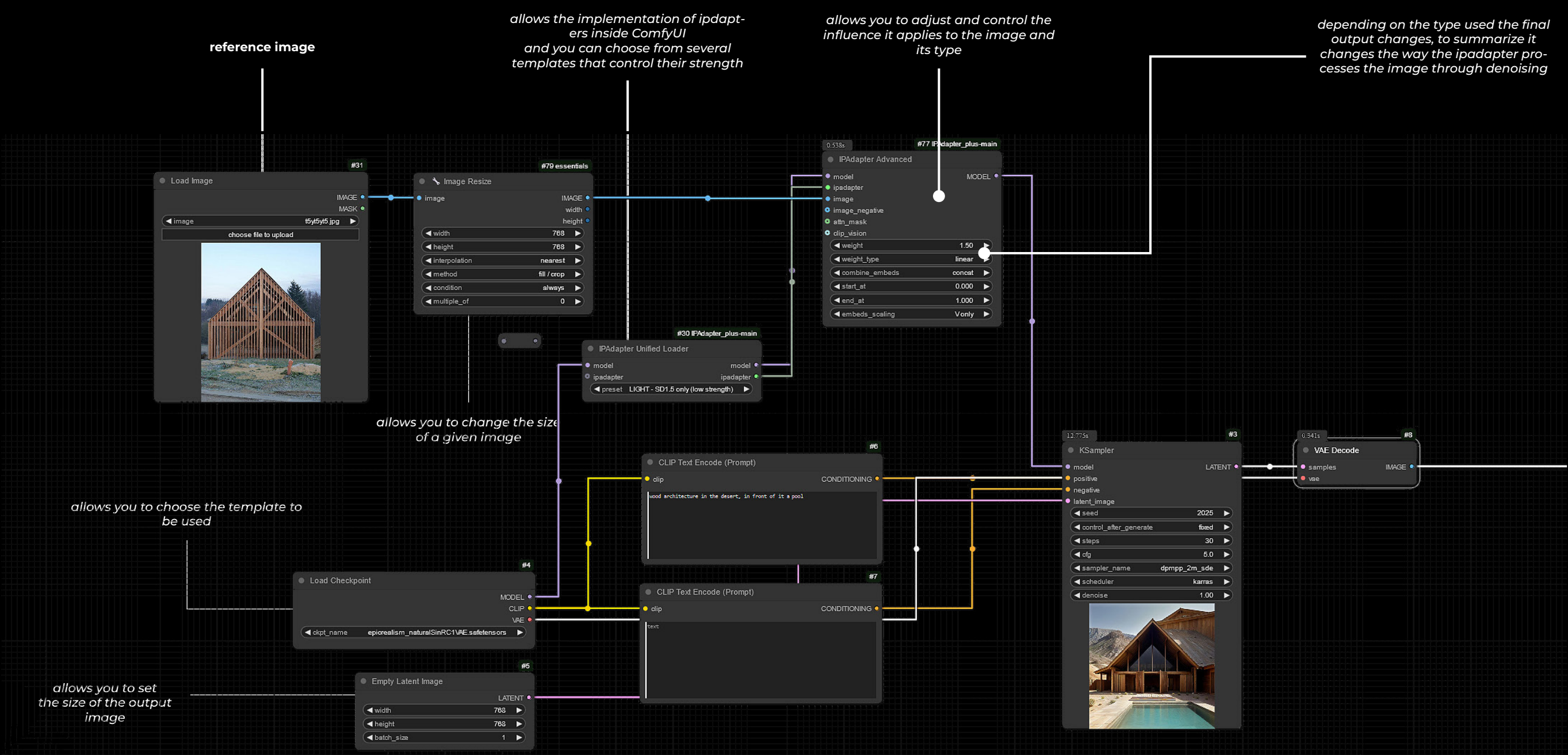
Through the subsequent mechanism of Cross Attention, the IP-Adapter manages to combine visual and textual features, and thanks to the 'decoupled' component it manages to keep these features divided so that the model preserves the particularities of both domains.

The core of the process, however, remains the U-Net Denoising, which generates the images.

The module obtains as its initial input a noisy image represented as X_t and proceeds with a series of transformations so as to reduce the noise, represented as X_{t-1} .

Modules that are tractable within the process allow the model to be adapted to specific requirements while those that are not provide the solid foundation for the process.

Workflow Explanation



Weight and Weight Type

The IPAdapter is a very complex tool with which you have the ability to transfer features of an input image to an output image. You can change various parameters that allow you to have different results.

In the next study, all process types of the IPAdapter were analyzed with two different weights : 1.00 and 1.50. The prompt used is identical to the previous one : modern wood architecture in the desert, in front of it a pool. model : epicrealism seed : 2025 steps : 30 cfg : 5 sampler : dpmpp_2m_sde scheduler : karras

w 1.00



reference image



linear



ease in



ease out



ease in-out



reverse in-out



weak input



eak output



weak middle



strong middle



style transfer



composition



strong style transfer

81 // example of IpAdapter output and different parameters with 1.00 as weight value

w 1.50



reference image



linear



ease in



ease out



ease in-out



reverse in-out



weak input



weak output



weak middle



strong middle



style transfer



composition



strong style transfer

82 // example of IpAdapter output and different parameters with 1.50 as weight value

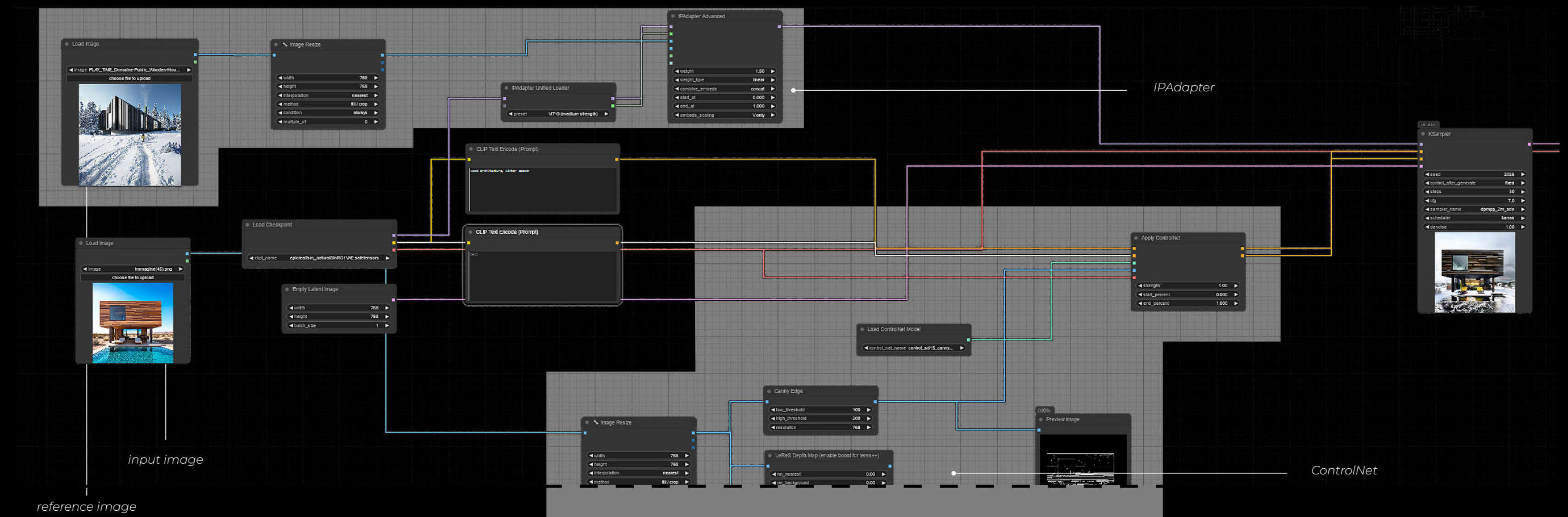
IPAdapter & ControlNet

Workflow Explanation

To further increase our influence and 'control' over the image it is possible to combine ControlNet and IPAdapter in the same process.

To get the best results we need to understand and calibrate the weight that each component exerts on the final output, it could become a long process and there is a risk of getting lost within the infinite results especially in the first experiments.

IPAdapter works with visual and textual input, it is important to adapt our prompt to the reference image and also important to choose our reference images well taking into account our source image that will be analyzed by ControlNet.




Iterations

The following study analyzes the same example used previously, the parameters remained unchanged except for the prompt which was modified slightly to fit the reference image : wood architecture, winter season.


The controlNet used in this case is CannyEdge.

w 1.00




reference image













<https://www.play-time.es/project/private-house-module/>



output without IPAdapter influence




84 // examples of the combined workflow output

 <p>linear</p>	 <p>ease in</p>	 <p>ease out</p>
 <p>ease in-out</p>	 <p>reverse in-out</p>	 <p>weak input</p>
 <p>weak output</p>	 <p>weak middle</p>	 <p>strong middle</p>
 <p>style transfer</p>	 <p>composition</p>	 <p>strong style transfer</p>


Page 174

w 1.50




reference image













<https://www.play-time.es/project/private-house-module/>



output without IPAdapter influence

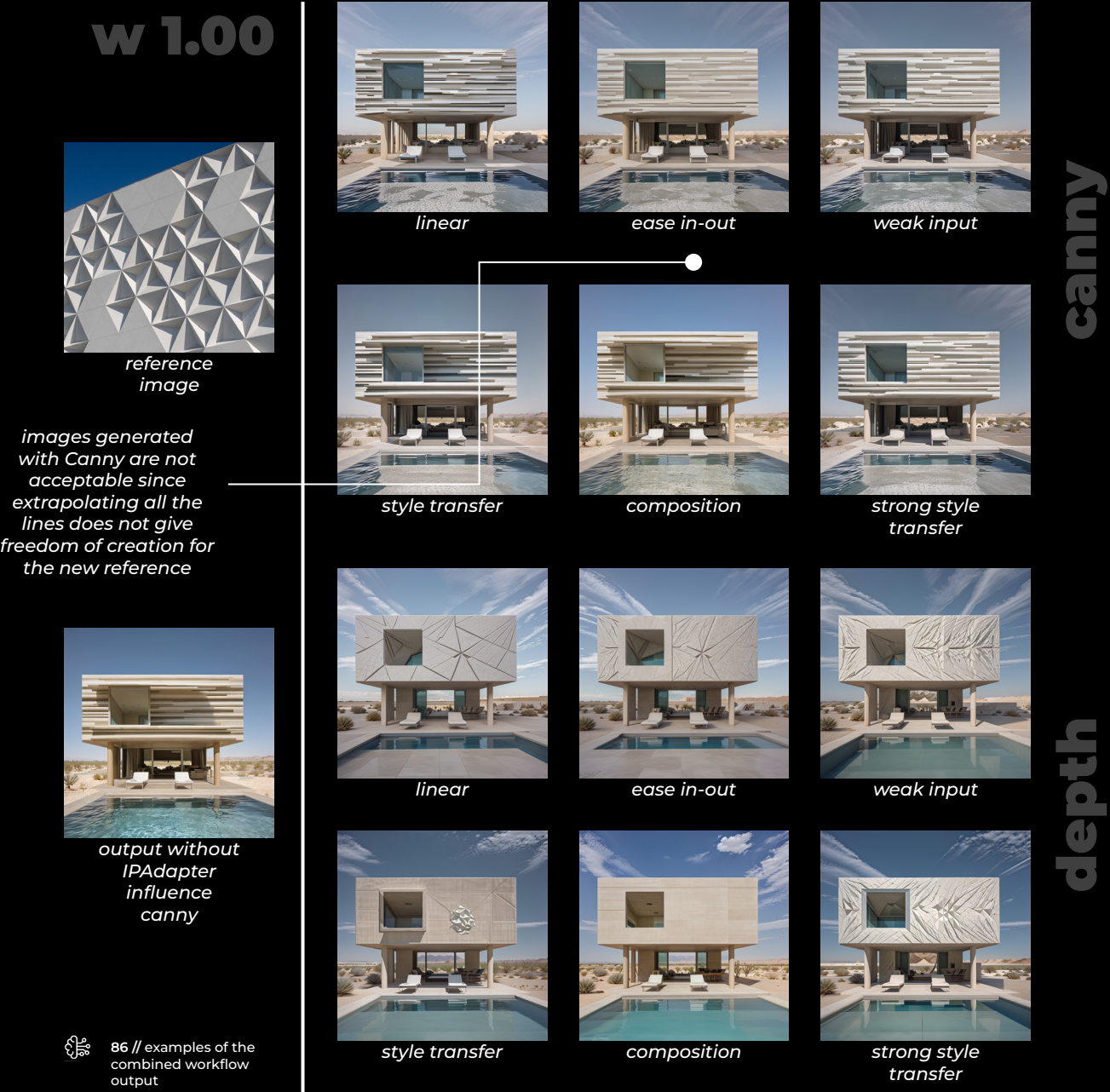


85 // examples of the combined workflow output

 <p>linear</p>	 <p>ease in</p>	 <p>ease out</p>
 <p>ease in-out</p>	 <p>reverse in-out</p>	 <p>weak input</p>
 <p>weak output</p>	 <p>weak middle</p>	 <p>strong middle</p>
 <p>style transfer</p>	 <p>composition</p>	 <p>strong style transfer</p>

Page 175

To demonstrate the effectiveness of this process, we continued the study by choosing as a reference an image that is completely detached from the input image architecture.

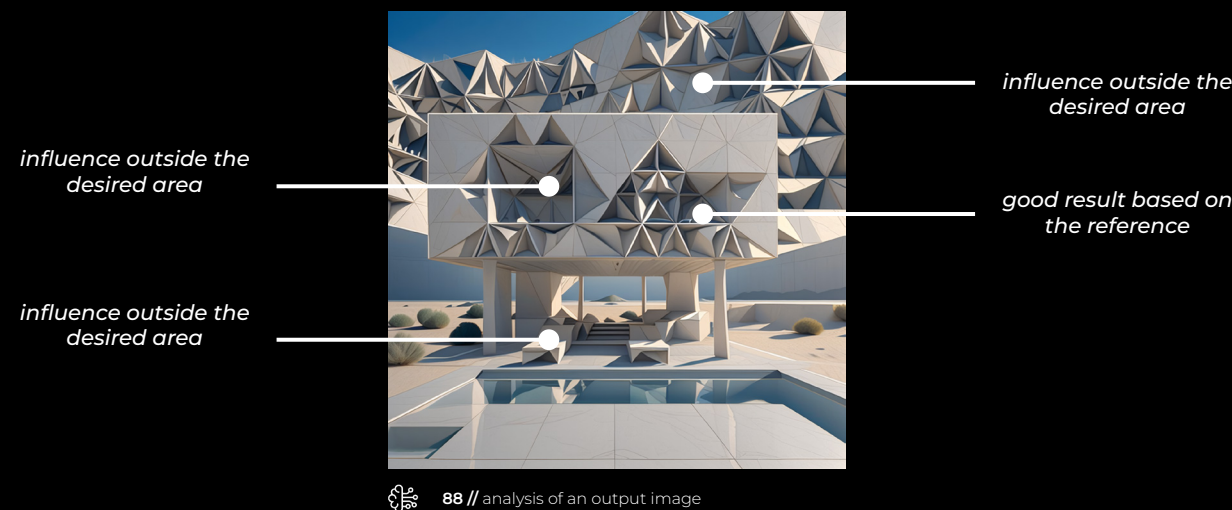


We chose some of the typologies to analyze and had them processed with ControlNet Canny and Depth.
prompt : parametric modern architecture in the desert, in front of it a pool.



Segmentation "The Recognizer"

General understanding

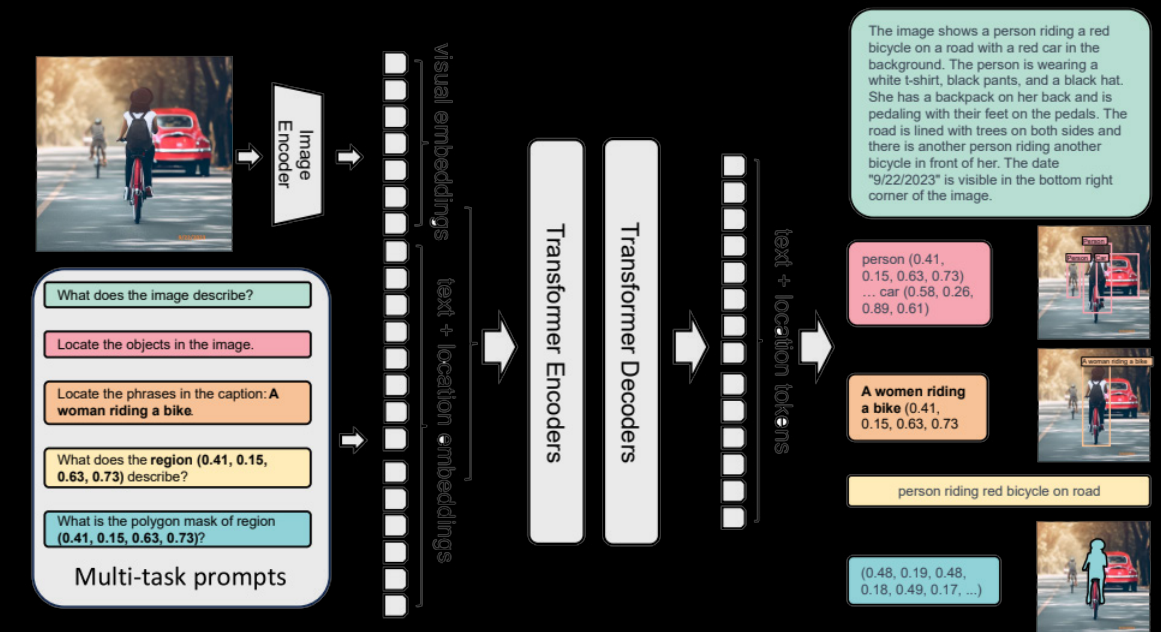


The latest results look promising but not yet acceptable since the influence of the IPAdapter has expanded everywhere in the image.

To have a realistic image we should expect to see the style transfer process only on the facade.
To be able to have this control through artificial intelligence we can rely on a process of Segmentation.
This process allows us to identify and isolate parts of the image and then create masks of them, which can then be processed using previously trained segmentation models such as SAM (Segment Anything Model).

We will make use of the new component developed by Microsoft : Florence 2.

'Florence-2 is a lightweight vision-language foundation model developed by Microsoft Azure AI and open-sourced under the MIT license. It aims to achieve a unified, prompt-based representation for diverse vision and vision-language tasks, including captioning, object detection, grounding, and segmentation. Its strength lies in the extensive FLD-5B dataset, which consists of 126 million images and 5.4 billion annotations, enabling robust zero-shot and fine-tuning capabilities.' (Ashtekar, 2024)²⁸.



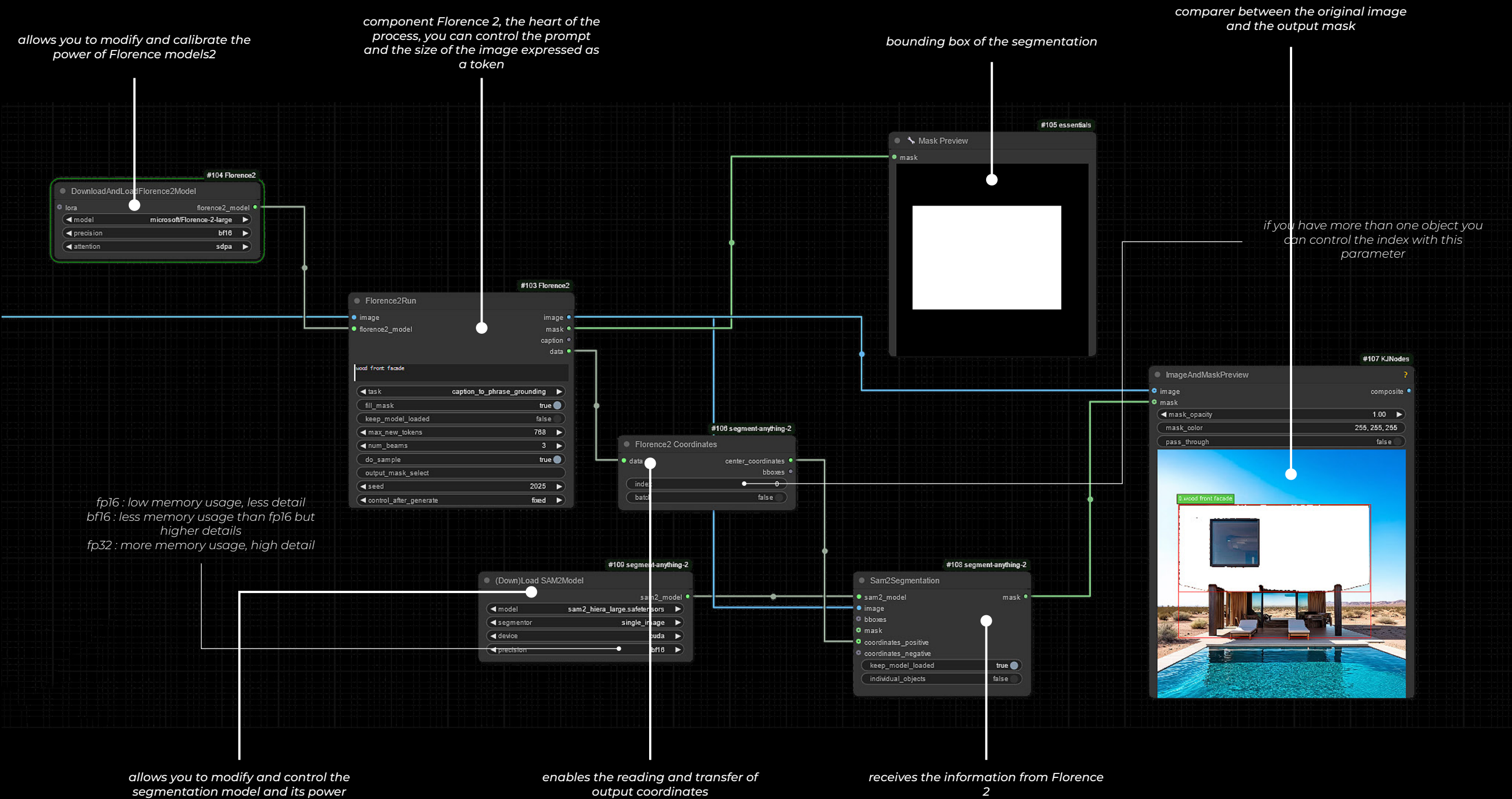
89 // segmentation process explanation (source : <https://www.labellerr.com/blog/unveiling-florence-2-microsofts-cutting-edge-vision-model-for-diverse-ai-applications>)

Florence-2 is designed in a simple way - to take in textual prompts (in addition to the image being processed), and generate textual results. Unifying the way in which diverse types of information - masked contours, locations, etc. - are input to the model permits (i) a unified training procedure and (ii) easy extension to other tasks without the need for architectural modifications, which are the two hallmarks of a foundational model.

In particular, Florence-2 adopts a classic seq2seq transformer architecture into which both visual, textual, and location embeddings are fed. The input image and prompt are mapped into embeddings which are then simply concatenated and passed into the standard Transformer encoder-decoder.

²⁸source : <https://medium.com/%40sunidhi.ashtekar/all-you-need-to-know-about-florence-2-348e23f75768>

Workflow Explanation



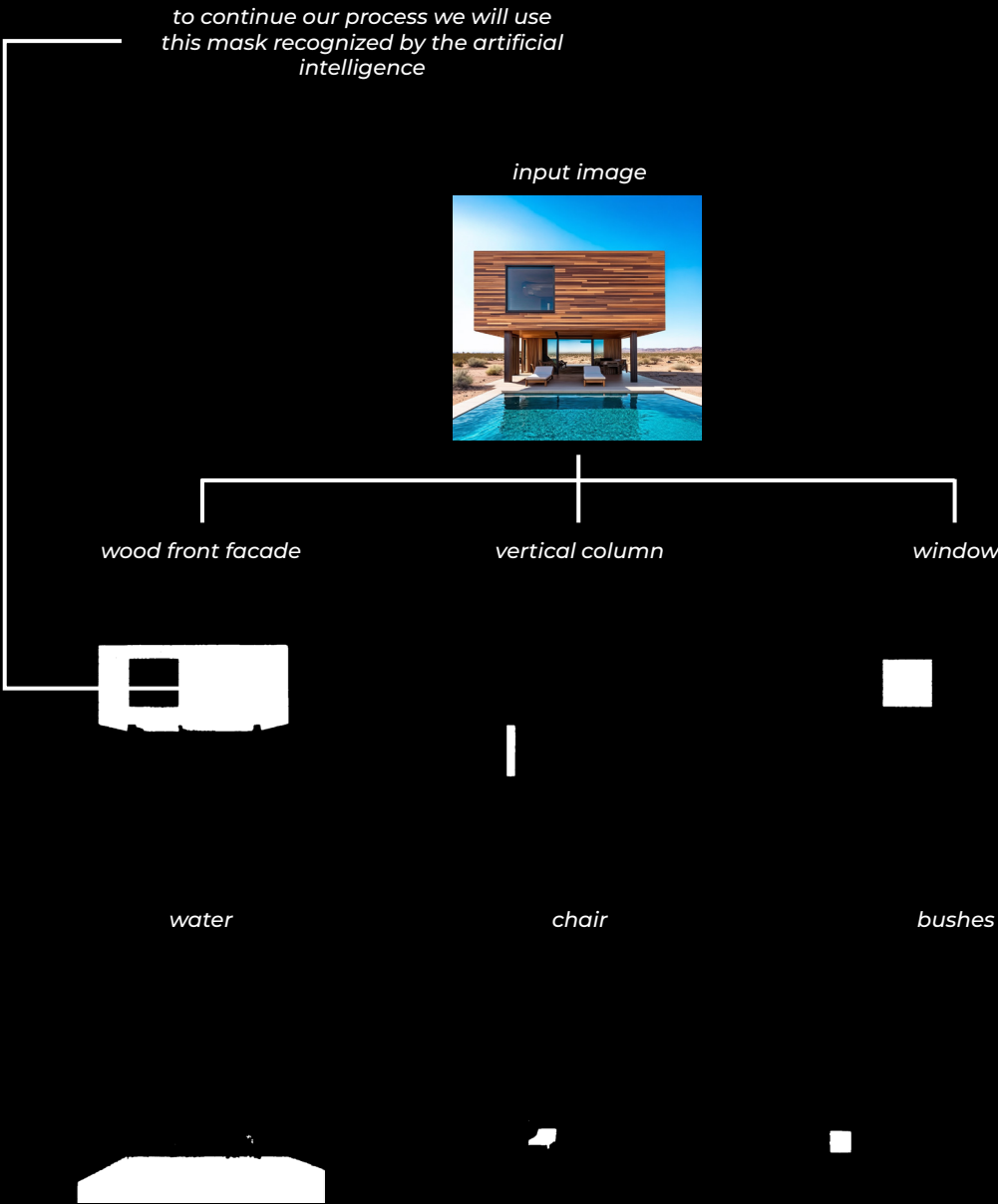
Iterations

To understand how Florence 2 works, we have inserted inside the workflow the image created earlier with the goal of recognizing and isolating the facade.

Of course it can recognize most objects that are present in the input, it all depends on the prompt we want to use.



What we are interested in are the various masks that we can extrapolate from the segmentation and later use as input in the same process we used previously.



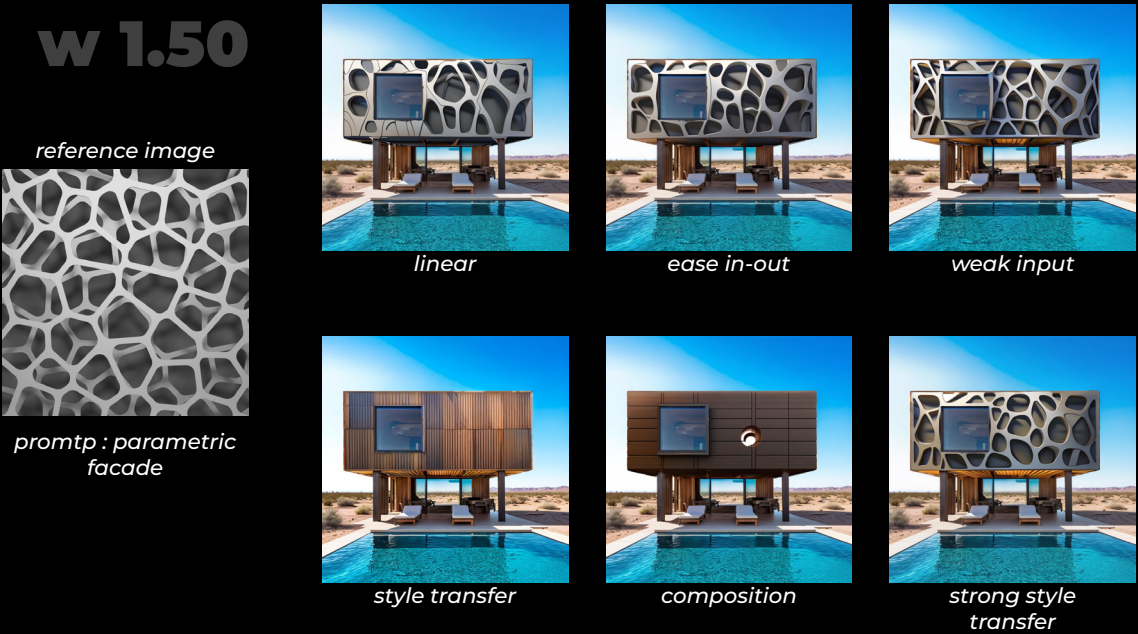
IPAdapter, ControlNet & Segmentation

Iterations

The extrapolated mask will be integrated inside the empty latent image so as to give the ksampler references on where to act.



As found previously for this specific case it is better to use the ControlNet Depth since it gives us more freedom of movement in the area of influence.



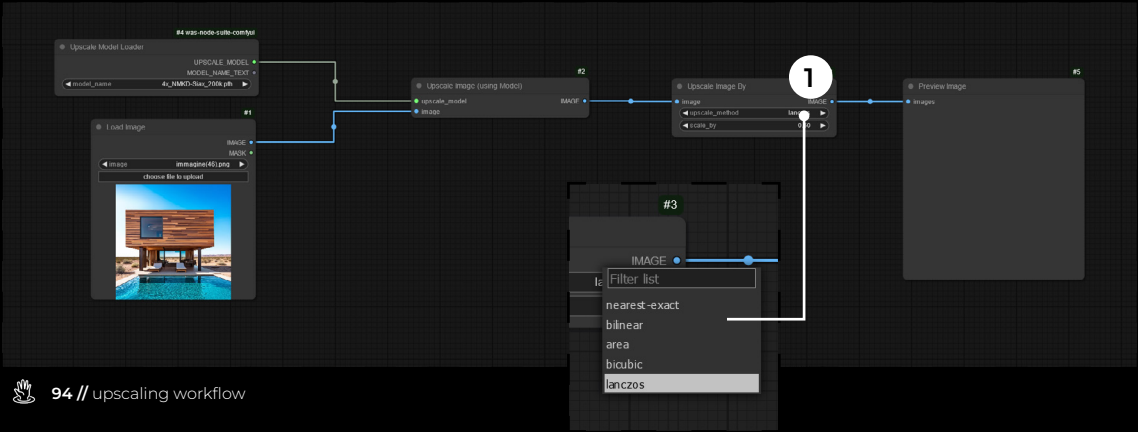
The experimentation can continue indefinitely, and the combinations between the parameters of the various nodes, the models used, and the possible inclusion of Lora to add references is infinite.

This lengthy study aims to explore the main conditioning methods that can be used and implemented in the architectural process.

As can be seen not all outputs can be considered acceptable and it depends heavily on the parameters we set during the process.

Upscaler "The enhacer"

Workflow and Iterations



The upscaling process makes it possible to improve the quality of the image, increasing its detail and adjusting parts of the image that are unclear. It is a very simple process that is usually linked in workflows to the final output.

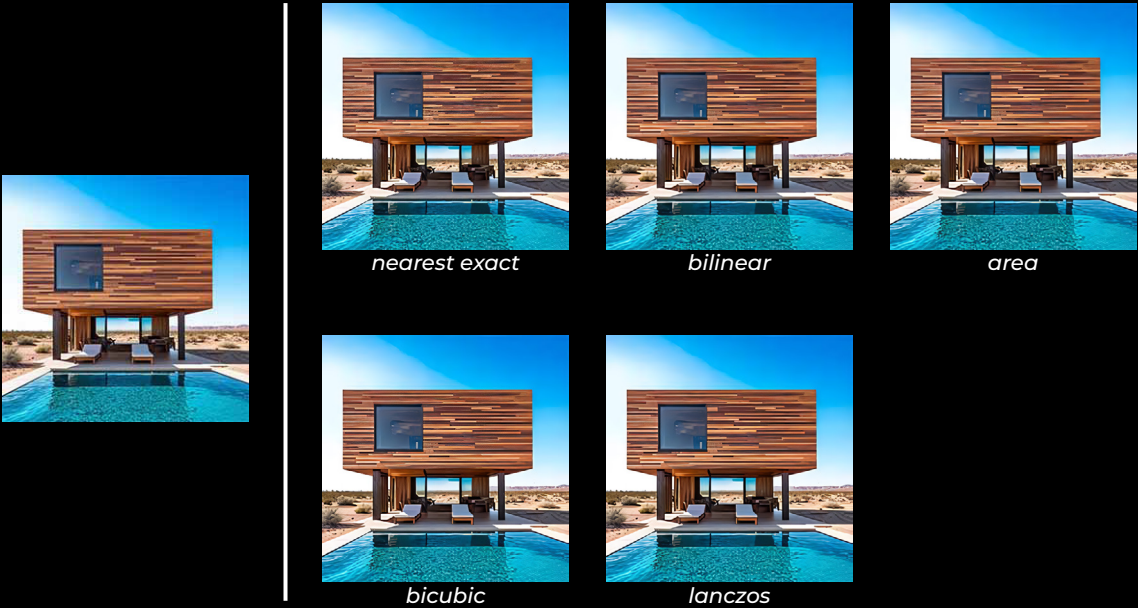
The main node that allows the control, besides the upscaling model used, is the `upscale image by (1)` that allows to resize the image according to some specific parameters, the `scale by` allows to control this value, in this case it is 0.50 which means that the image will be resized by 50 %, we have to be aware of the dimensions of our images and understand well how to manage them during the various workflows.

Different types of upscaling are available:

- nearest exact creates a copy of the pixels in the following ones creating a 'pixelated' effect
- bilinear uses an average of neighbouring pixels
- area calculates the area of pixels
- bicubic is more advanced and calculates pixels with cubic interpolation (more detailed results)
- lanczos uses a weighted average of neighbouring pixels to calculate (out of all, it is the one that produces the best results)

Since the starting image already has a very high quality, we use the same image but modified to a much lower quality, otherwise the results would not be comparable here, the sharpness would be further increased.

In this case, a scale factor of 0.50 was used.



95 // examples of different upscaling outputs

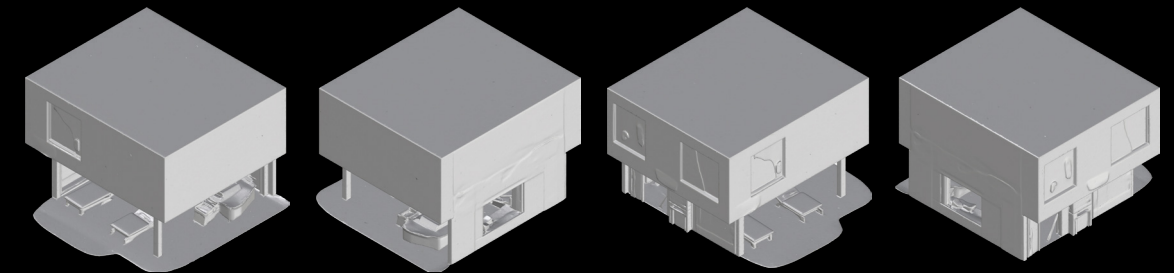
2D to 3D: Transforming Images into Objects

Tripo AI

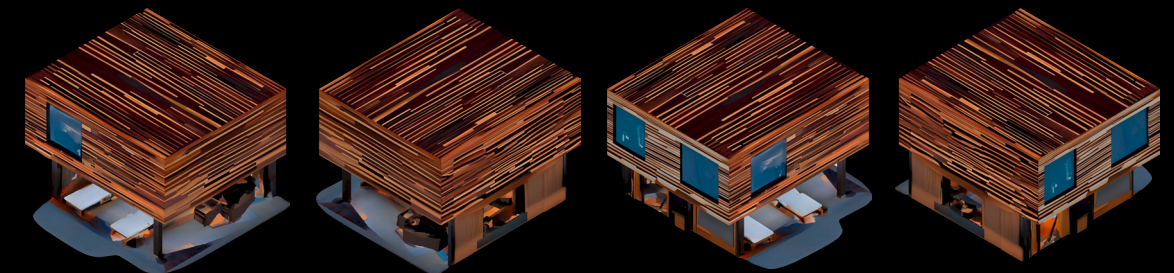
The next big step would be the generation of 3D models from simple 2D images, at the time of writing this thesis research is advancing to obtain similar results (January 2025) and some online platforms such as Krea.ai, Tripo.ai or CGDream allow this.

For the test we did we kept the same starting image, a clear image with simple shapes, and put it on Tripo.ai to convert it into 3D, on the right side we see the results which can be described as very good at first glance.

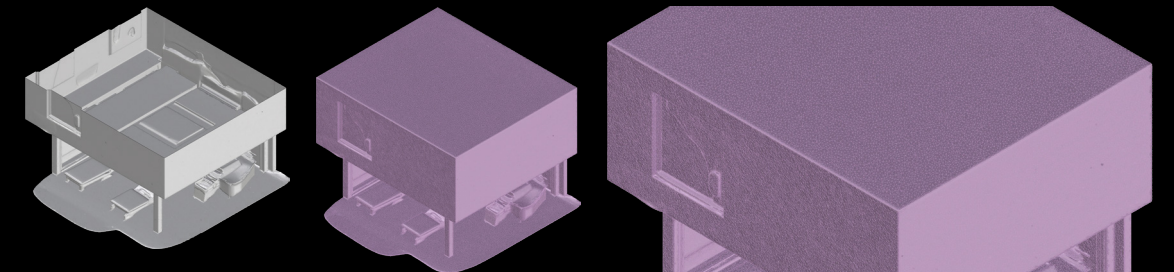
The main problem lies in the impossibility of modifying the geometry, which consists of hundreds of tiny meshes that are practically impossible to control. Nevertheless, the result is promising, although the image is clear and simple in shape.



96 // 3d views of the generated model



97 // 3d textured views of the generated model



98 // sectioned 3d model

99 // 3d view of the model and zoom to see the density of the mesh structure

scan for
3d video
interaction



How can artificial intelligence help us in generating ideas ?
AI in generative design, is it just a tool or not ?
How much influence can we exert on artificial intelligence ?
Can we use AI to generate and enhance our ideas ?

01 - The Last Nuclear Bomb Memorial

Experimentation for Efficiency

Sitography

<https://architecturecompetitions.com/?open-competitions>

Page 190

Page 191

The Competition

Key dates and overview

The Last Nuclear Bomb Memorial / Edition #5 is the fifth annual competition in which participants are tasked with creating a memorial located within a decommissioned nuclear weapon testing site.

In response to the global silence surrounding the issue of nuclear weapons, participants in “The Last Nuclear Bomb Memorial / Edition #5” competition must submit their designs with NO DESCRIPTION TEXT.

The architecture ideas must be communicated strictly with visuals. We are asking participants to design a building or structure in which the architecture would do all the talking.

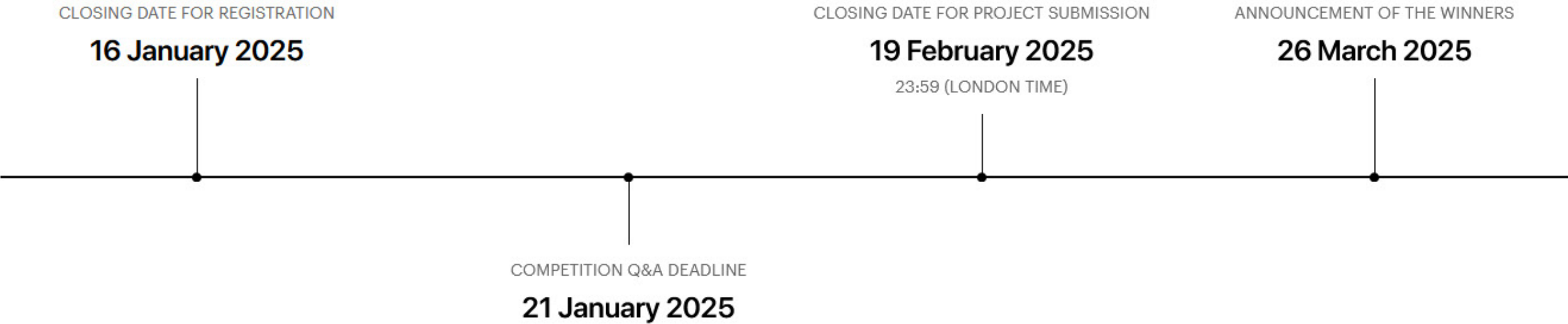
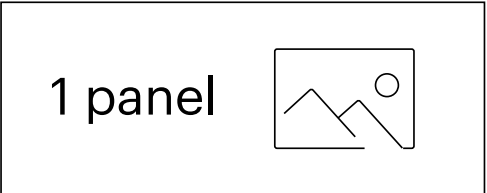
- Participants are required to upload **one (1) A2 landscape-orientated presentation board** (must not exceed 10MB per jpg) with sketches, renderings, plans, sections, elevations, diagrams, and/or other presentation tools to explain their proposal.



NO DESCRIPTION TEXT!

- No video files are accepted.
- All submissions must be uploaded via the architecturecompetitions.com upload panel. Access information and instructions on how to upload the presentation boards will be issued to participants via email immediately after successful registration.
- Presentation boards must not indicate any information related to an individual's/team's identity.

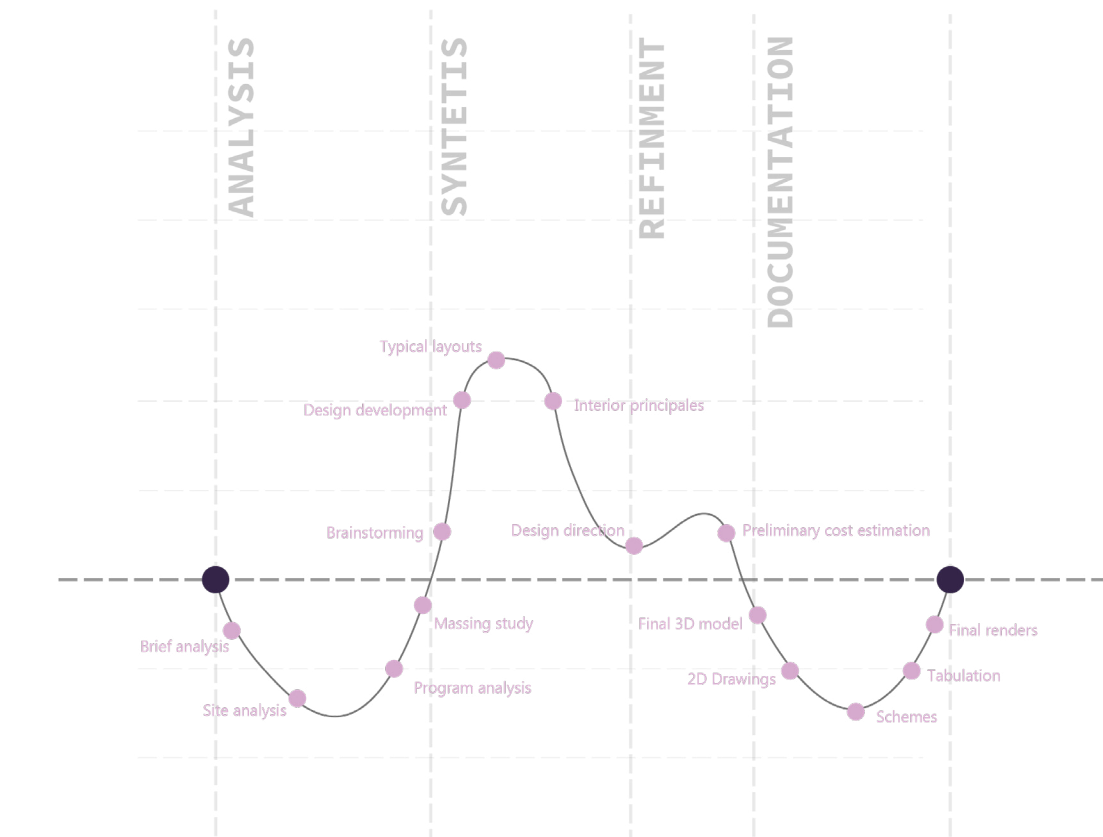
Participants who do not comply with the requirements will be disqualified without refund.



Lowering of the Curve

Before starting the process, we zoomed in on the phase that is mostly taken into consideration in this thesis and how artificial intelligence can be integrated within it, i.e. the Schematic Design phase, which has been divided into four categories: analysis, synthesis, refinement and documentation.

Each project is unique, but we have focused on the standard operations that an architectural process entails during this phase, following the guidelines drawn up by the NCARB²⁹ and AIA³⁰.



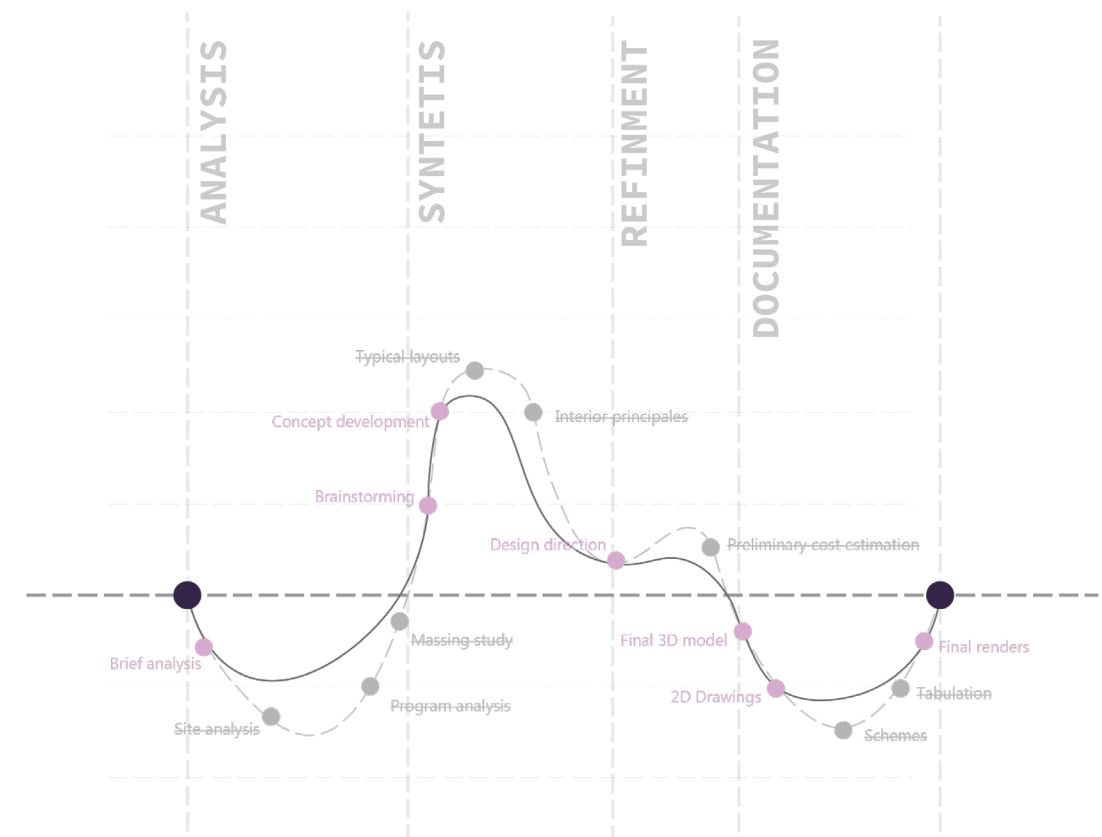
101 // traditional architectural process

²⁹NCARB stands for the National Council of Architectural Registration Boards, which oversees the licensing and credentialing of architects in the U.S.

³⁰AIA stands for the American Institute of Architects, a professional organization that supports architects through advocacy, education, and resources

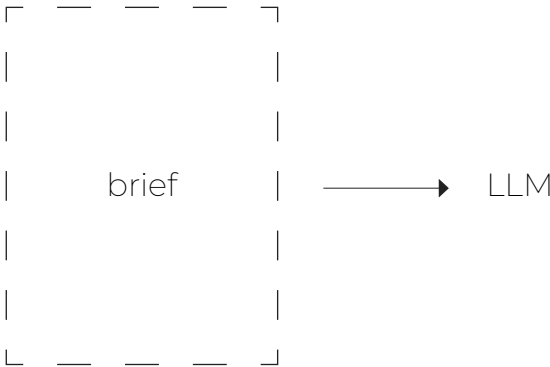
As the project we are going to tackle now is slightly different from a traditional project because it is a concept of ideas and the deliverables required are different, we have modified the scheme by adapting it to the requirements of this process.

In the final part we will see how this process has been influenced by artificial intelligence and how it has changed.



102 // variation of the traditional architectural process

Brief Analysis

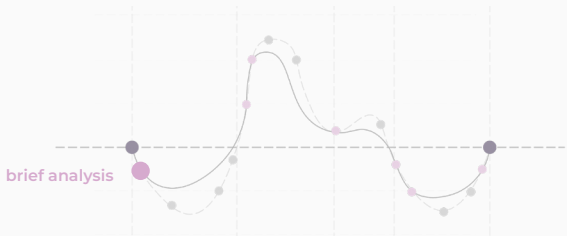


The first step in our process was to have the project brief analysed by an LLM, in this case ChatGPT.

This is an extremely quick process to get the key points of the brief.

In addition to the formal aspects, we also asked for initial suggestions regarding the project so that they would perfectly meet the competition’s requirements.

Today, this type of artificial intelligence integration in the architectural world is widely used, from the preparation of an email to the analysis of a document, thanks to AI everything is much faster.



Summary :

The competition, “The Last Nuclear Bomb Memorial,” seeks innovative architectural proposals for a memorial located on a former nuclear weapons test site. The design must convey a powerful narrative about the devastating impact of nuclear weapons and their lasting consequences, while inspiring reflection and promoting global peace.

Key aspects :

The memorial should communicate its message solely through visuals without accompanying descriptive text.
Participants are free to determine the scale and function of the memorial.
The site is a disused nuclear test area, tying the proposal to the history of nuclear testing.
The submission must align with the theme of “The Last Nuclear Bomb”, symbolising an end to nuclear conflict.

Suggestion for approach :

Embrace Symbolism: Use architectural elements that evoke the duality of destruction and hope (e.g., forms resembling mushroom clouds transformed into symbols of renewal).
Experiential Design: Incorporate interactive or sensory elements (e.g., sound, light, or shadows) to engage visitors emotionally and physically.
Abstract Representation: Avoid literal depictions; instead, focus on abstract forms that allow for individual interpretation.
Universal Language: Use geometry, light, and materiality to craft a design that resonates across cultures without the need for words.
Ethical Message: Highlight sustainability and the preservation of the environment as part of the design’s narrative.

The Design Concept

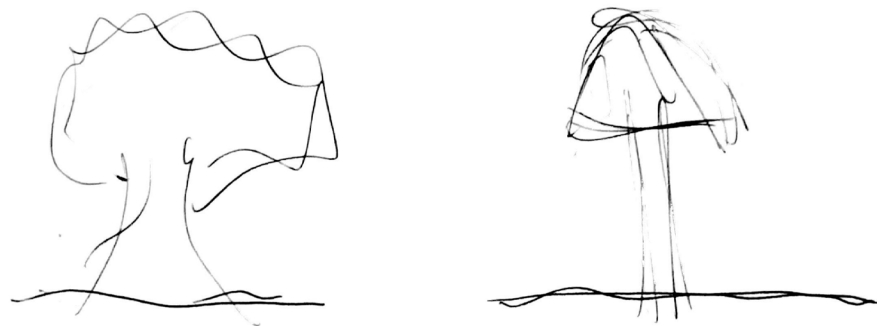
The prompt

The approach to the project was structured in a very definite way, first of all we searched for an idea, in this case the idea came from us and therefore from human creativity, but an idea could be generated directly with the help of artificial intelligence.

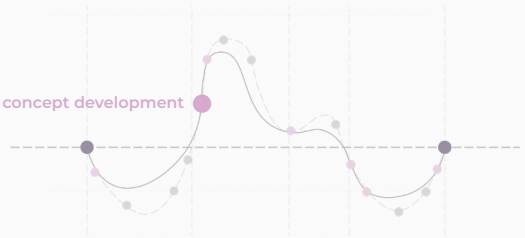
We started with the decomposition of the nuclear bomb, which led us to want to define a shape similar to that of the explosion, so a kind of atomic mushroom.

We then fed this idea to the AI so that it would generate a valid prompt for image generation, you can ask the AI to modify or change the prompt if it seems unsuitable.

In software such as ComfyUI it is possible to calibrate the power of individual words or phrases so that that part of the prompt is more prevalent, this can be done by using CTRL + UP/DOWN on the parts of the prompt where you want to intervene.



 103 // hand-drawn sketch



Prompt_1

'A dramatic, cinematic black-and-white shot with high contrast, capturing an award-winning aerial view of an abandoned atomic bomb test site. In the center of the crater, a surreal architecture installation creates a thick fog. Barely visible through the mist, a massive metal frame in the shape of a mushroom rises, its underside glowing with dazzling lights. Small, scattered figures on the ground are bathed in these lights, which pulse rhythmically, as if following the heartbeat of the earth.'

Prompt_2

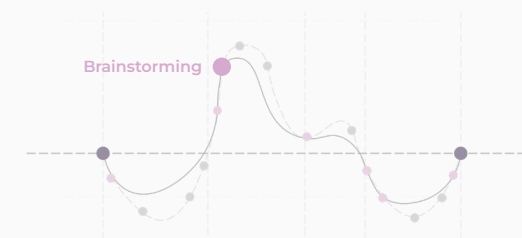
A high-contrast cinematic photograph, capturing a forgotten atomic test site. In the core of the barren crater, a mysterious architectural structure with the shape of a nuclear bomb explosion. Shadows of tiny figures below are illuminated by the rhythmic flashes, creating an otherworldly dance in the mist.

Prompt_3

A striking, black-and-white cinematic image with sharp contrasts, showcasing an award-winning aerial perspective of a deserted atomic test site. At the heart of the crater, an enigmatic installation shrouded in dense fog reveals a towering structure shaped like a mushroom, its underbelly aglow with vivid, rhythmic lights. Distant, scattered silhouettes of people on the ground reflect the surreal, pulsing illumination.



01 - The Last Nuclear Bomb Memorial

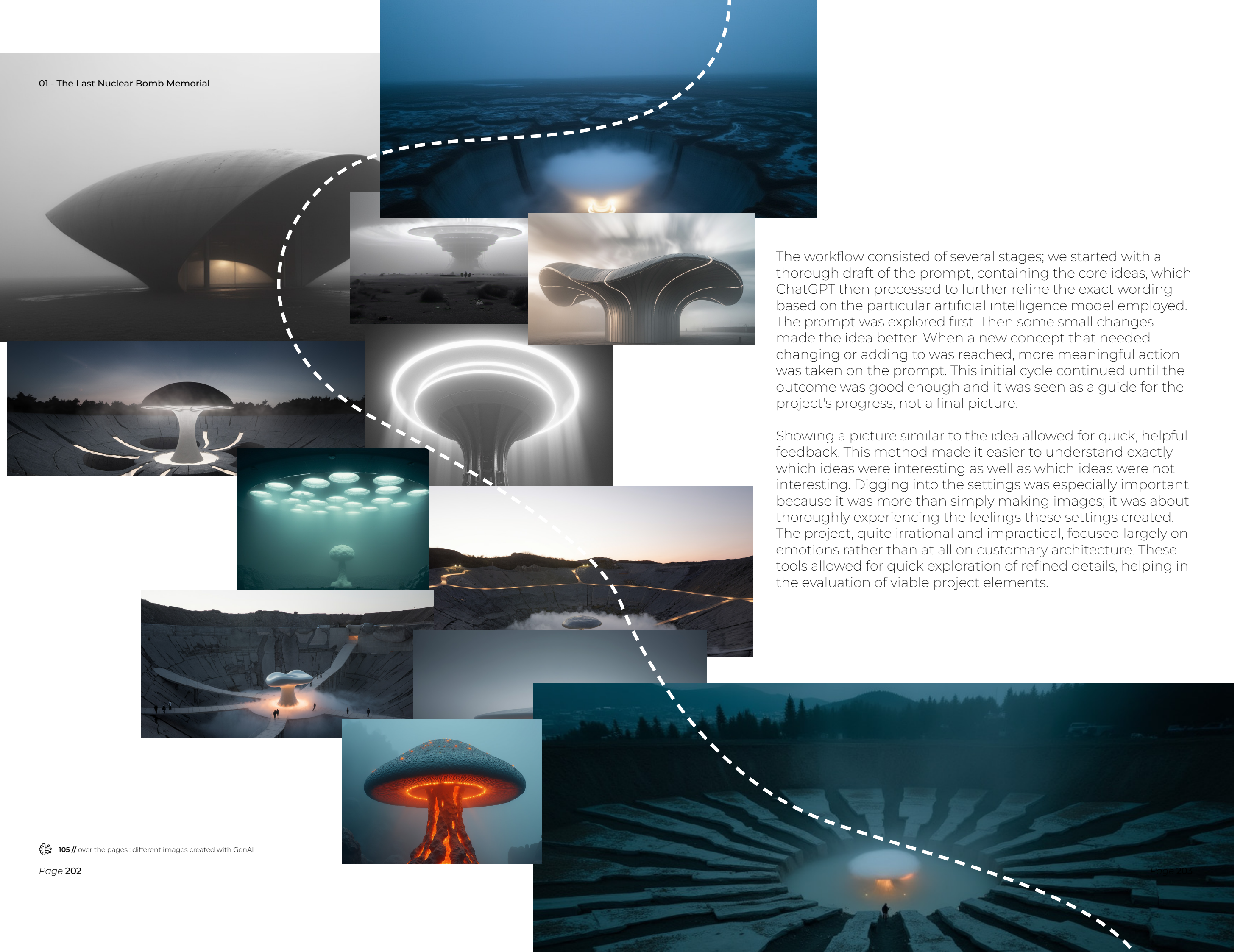


Following the flow, an uncontrolled brainstorm

The conceptualisation process began with the use of a simple and straightforward tool, such as a classical large language model (LLM), specifically ChatGPT. This tool made it possible to explore and articulate an initially simple basic idea that had been proposed. The main idea was to recall the image of the atomic mushroom generated by the atomic explosion, but with a radically different meaning: instead of representing death and destruction, it was to be transformed into a symbol of life.

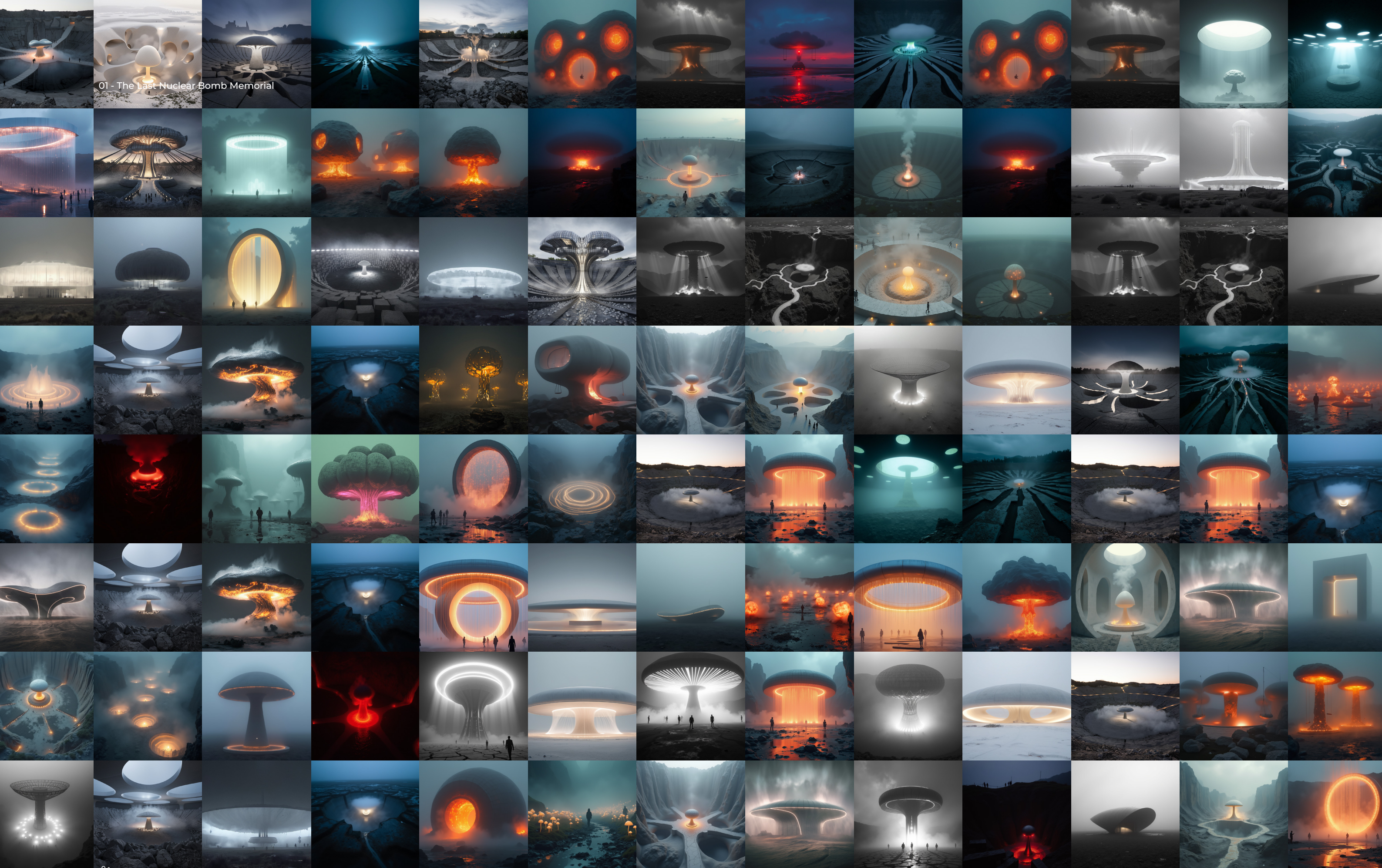
Out of this intuition came a creative journey that explored how the shape of the mushroom, instead of being associated with devastation, could instead become a key element in creating unique experiences within the project site. During this exploration phase, tools such as Flux and other Stable Diffusion models were instrumental in developing a coherent and continuous flow of ideas. This process led to the generation of images that, through the use of prompts, allowed the initial ideas to be visualised and refined.





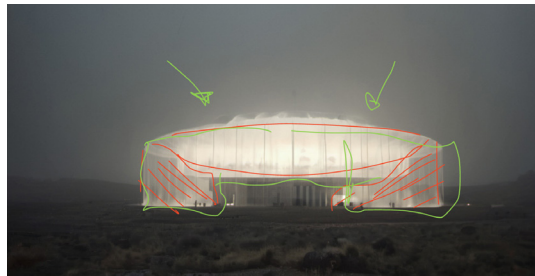
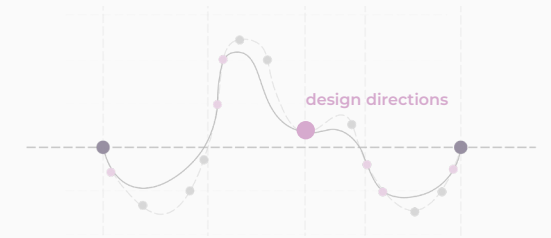
The workflow consisted of several stages; we started with a thorough draft of the prompt, containing the core ideas, which ChatGPT then processed to further refine the exact wording based on the particular artificial intelligence model employed. The prompt was explored first. Then some small changes made the idea better. When a new concept that needed changing or adding to was reached, more meaningful action was taken on the prompt. This initial cycle continued until the outcome was good enough and it was seen as a guide for the project's progress, not a final picture.

Showing a picture similar to the idea allowed for quick, helpful feedback. This method made it easier to understand exactly which ideas were interesting as well as which ideas were not interesting. Digging into the settings was especially important because it was more than simply making images; it was about thoroughly experiencing the feelings these settings created. The project, quite irrational and impractical, focused largely on emotions rather than at all on customary architecture. These tools allowed for quick exploration of refined details, helping in the evaluation of viable project elements.

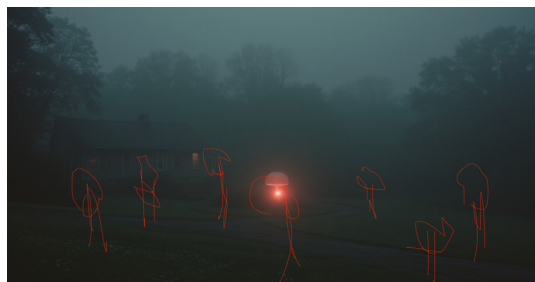


01 - The Last Nuclear Bomb Memorial

Analyzing the output



The image has this semi-transparent element on the façade, hence the idea to create concentric veils that recall the shock waves of the nuclear explosion, and this element can also be a symbol of lightness and therefore of the fragility that is present after an explosion.



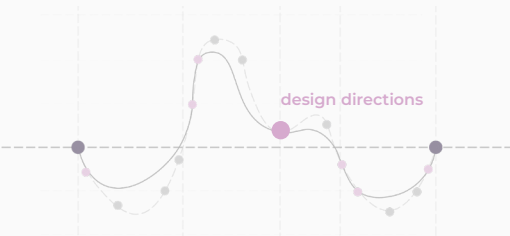
This output generated this very dark image with a small, illuminated, central element, and was a cue to come up with an idea where the crater can be filled with many illuminated, solitary elements.



In the picture you can see the marks on the floor of the crater, and it was a good idea to come up with an idea where the crater has one main element and work on what is around it at floor level by creating paths or circles, so that they would recall the rift or shock waves.



The sharp shape of the element in the image gave rise to another idea in which there could be a circular, massive element embedded in the crater wall, this element acting more like a pavilion could be accessible to have a panoramic view of the crater.



The segmentation, a controlled brainstom

During the general brainstorming and the more in depth analysis of the interested images, we proceeded to deepen the study of some of the ideas that come in during the process. This allowed us to futher study some aspect of different ideas. This was very interesting specially because of the topic and the aim of building very specific enviroinment.



'A dramatic, cinematic black-and-white shot with high contrast, capturing an aerial view of an abandoned atomic bomb test site. In the center of the crater, a surreal architecture installation creates a thick fog. Barely visible through the mist, a massive metal frame in the shape of a mushroom rises, its underside glowing with dazzling lights. Small, scattered figures on the ground are bathed in these lights, which pulse rhythmically, as if following the heartbeat of the earth.'



'A dramatic, cinematic black-and-white shot with high contrast, capturing an aerial view of an abandoned atomic bomb test site. In the center of the crater, a surreal architecture installation creates a thick fog. Barely visible through the mist, a massive metal frame in the shape of a mushroom rises, its underside glowing with dazzling lights. Small, scattered figures on the ground are bathed in these lights, which pulse rhythmically, as if following the heartbeat of the earth.'



'Envision a minimal project that embodies destruction and rebirth. A crater is filled with solid concrete, representing immobilized devastation, it has voids that form interconnected pathways with leading to the center. At the heart stands a symbolic structure shaped like a mushroom, a reminder of atomic destruction. Soft lights and artificial mist emanate from the base, shrouding the center in an atmosphere of mystery and awe. Visitors traverse the pathways to reach this illuminated core, experiencing a space that symbolizes both catastrophe and the potential for renewal.'



'A dramatic, cinematic black-and-white shot with high contrast, capturing an aerial view of an abandoned atomic bomb test site. In the center of the crater, a surreal architecture installation creates a thick fog. Barely visible through the mist, a massive metal frame in the shape of a mushroom rises, its underside glowing with dazzling lights. Small, scattered figures on the ground are bathed in these lights, which pulse rhythmically, as if following the heartbeat of the earth.'

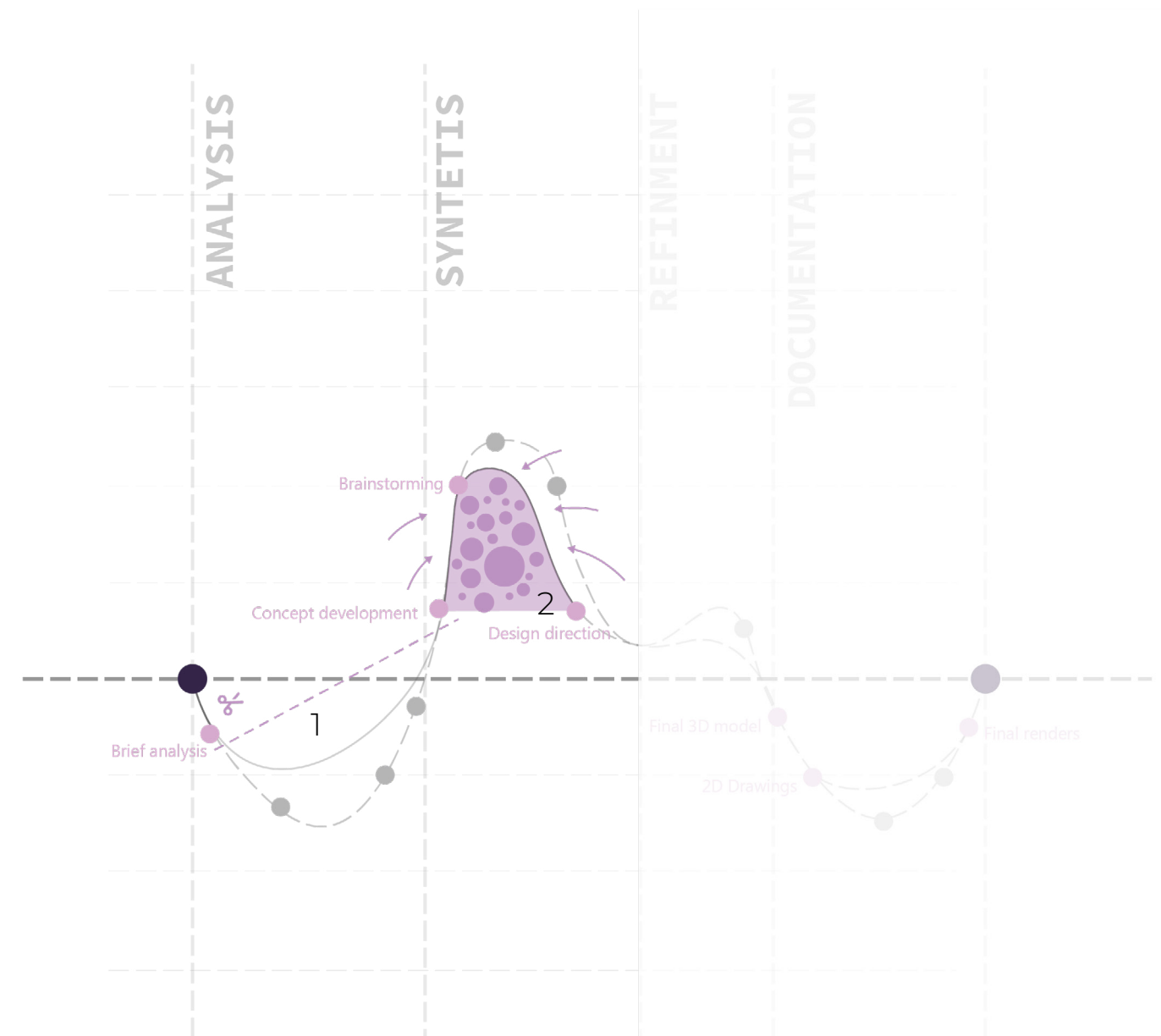


The concept conclusions

In the first steps, which mainly involve the analysis of the project in all its forms and the first brainstorming phase in which there is the first exploration of possible concepts and design of the project, we used artificial intelligence through single actions that we previously defined as helpers.

In this precise case, as mentioned above, we used an LLM (1) to analyse the project brief and transform this information into possible project concepts, this allowed us to speed up this process of project awareness and extrapolate the key parts of the requirements, of course this can also be supplemented by personal creativity and knowledge.

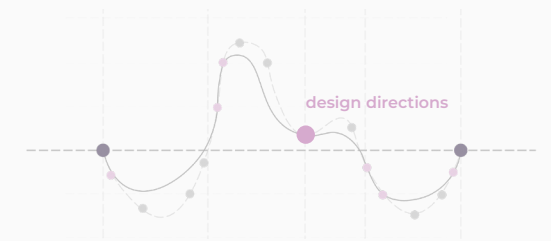
As far as brainstorming is concerned, the Conceptualizer process (2) has been used, which allows, through a large amount of image generation, to explore multiple options of concept and design of the project; this is a phase that, unlike the traditional phase, can lead to innovative ideas and to the instantaneous visualisation of the processed concepts, on the one hand it can shorten the time and increase creativity, but on the other hand you can get lost in the image generation as you can continue endlessly searching for new proposals, but of course in the first case if used with awareness it can lead to greater creative expression and shorter time.



AI Process - The Dialogue

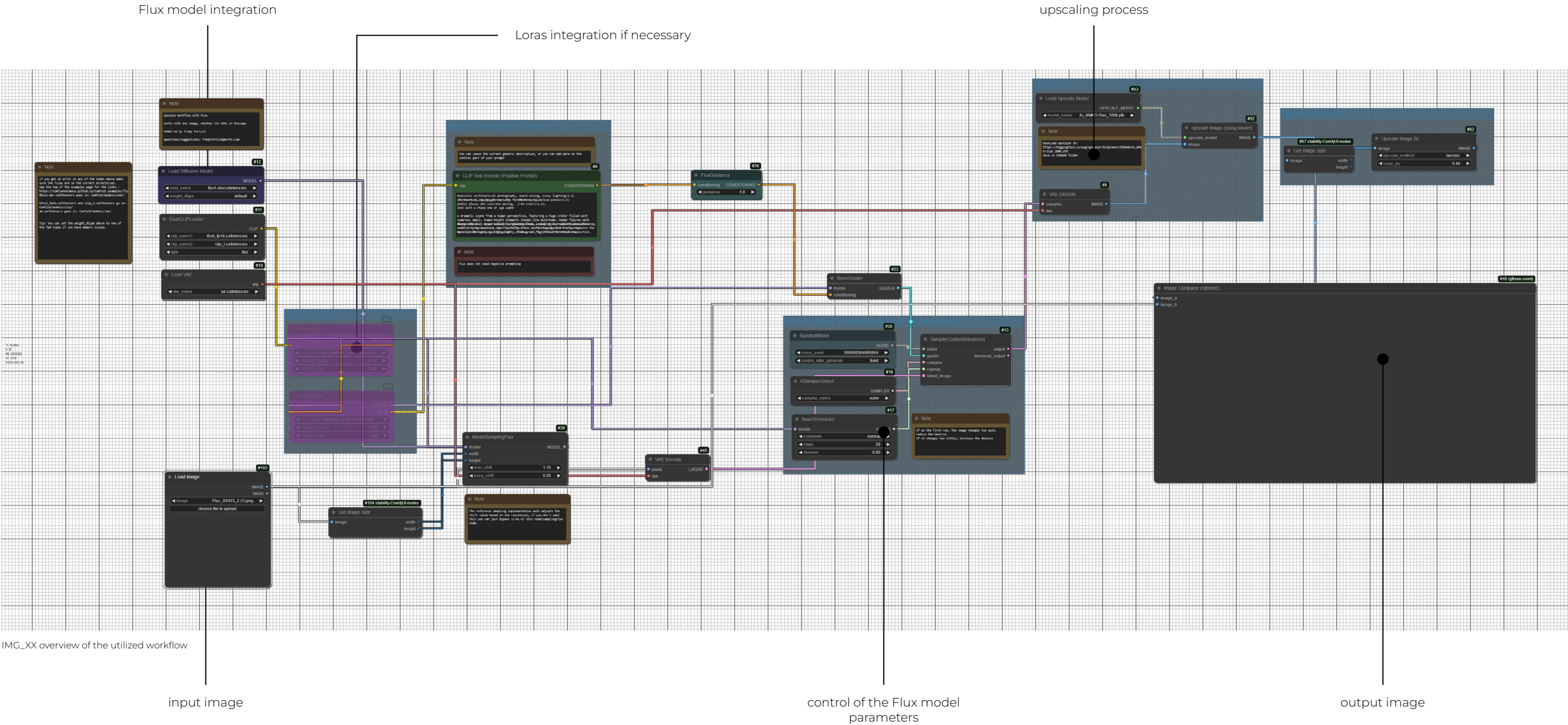
The approach

After the first phases we entered the design definition phase, in which case we kept the whole process focused on AI generation. We focused mainly on prompting and upscaling the created images in order to obtain better output. The starting image is unchanged and the design concept is the same. From the starting image some errors due to AI are visible, in this case our mushroom structures are wavy in some parts and this is not the final result we want to achieve. Through a process of image to image and upscaling using the Flux model, which currently requires a lot of computational power, we processed the image to obtain a new output. As for the other components in the previous chapters, we have explained the basic node and how it works, but many times to obtain better results it is necessary to make further additions and adjustments.



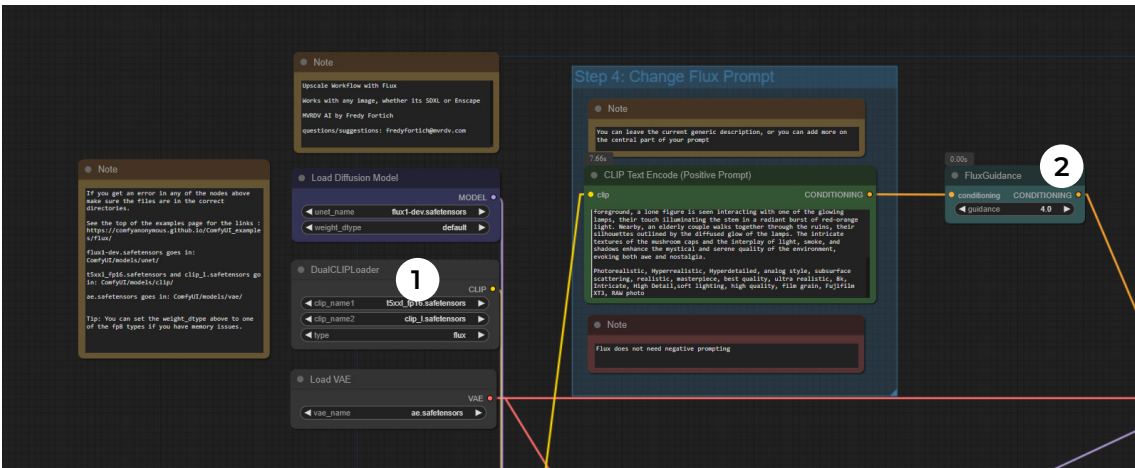
110 // starting image

Workflow explanation



IMG_XX overview of the utilized workflow

On the previous page, the general workflow was introduced in order to get an overview of the links and the various nodes used, now let's go into more detail about the main parameters manipulated and see how they affect the final result, as for the rest of the thesis, the same seed was used here to try to reduce the random aspect of generation.



112 // main Flux nodes

First of all, it is important to make an introduction to Flux, a model not used until now.

Flux is a model that allows the generation of images by diffusion, like the other models, the difference is that Flux requires a high computational power, which allows the creation of much more detailed and precise images.

One of the main features of Flux is the DualClipLoader (1) optimised for this model, it allows the reading and understanding of the prompt thanks to more than one CLIP model, moreover Flux works very well with very descriptive prompts unlike other models such as SD1.5 or SDXL.

The FluxGuidance node (2) allows you to calibrate how much the prompt affects the final output, higher values intensify the conditioning.

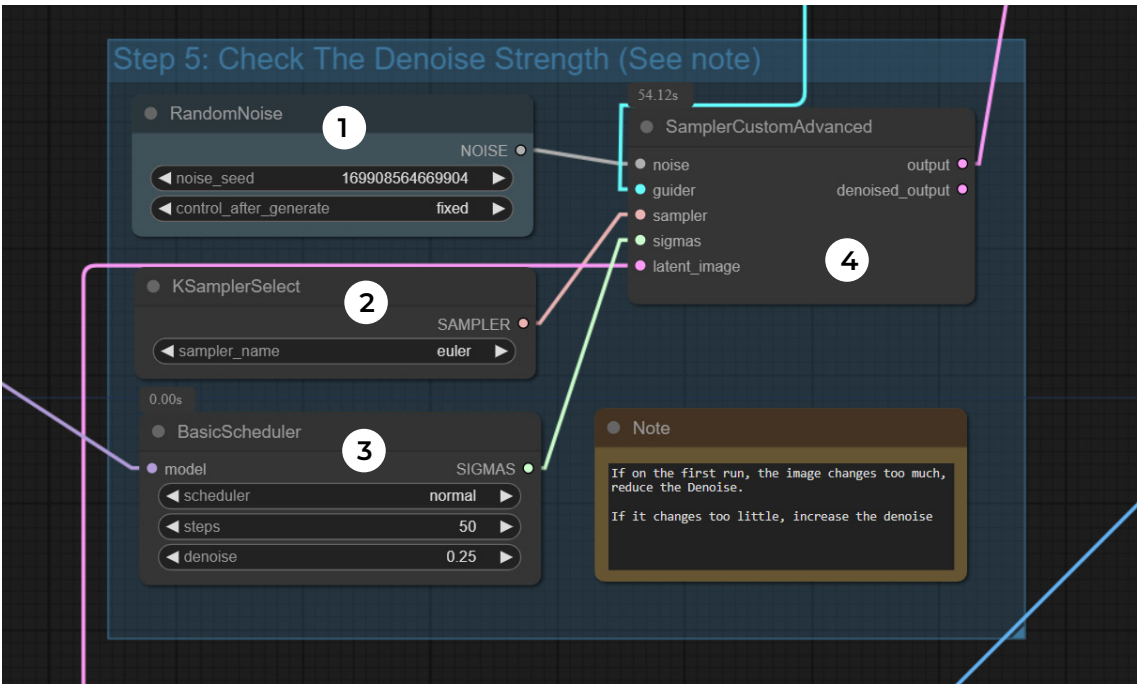


113 // different images created with Flux model and guidance test

Above we see how the FluxGuidance affects the final output, obviously this is greatly influenced by the denoiser which we have kept at a low average value of 0.25 to see the differences.

We can see how a high Guidance value fits the prompt, in the image with guidance 1 on the left there is a column that slowly becomes a person as the guidance increases because we have specified 'human figures' in the prompt.

prompt : a dramatic scene from a human perspective, featuring a huge crater filled with numerous small, human-height elements shaped like mushrooms. Human figures walk among. The mist wraps around the glowing forms, creating a dreamlike atmosphere under a dark, overcast sky. The reflections on the damp ground further enhance the mystical ambiance, blending light, shadow, and fog into an ethereal composition.



114 // Flux denoising nodes

We can say that the nodes that are manipulated the most are those relating to the denoiser. First of all, we have the RandomNoise (1), which generates the initial noise that will be processed later. The KSamplerSelect (2) is simply an extension of the sampler node seen above which allows us to change the sampler used.

The BasicScheduler (3) allows you to control the number of steps and the strength of the denoise applied during the generation process; this is the node where the parameters have been modified and tested the most.

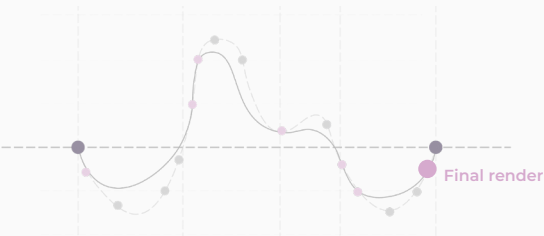
While the SamplerCustomAdvanced (4) is the box that contains the previously explained nodes and enables the denoising process.



115 // different images created with Flux model and denoise test

For the images generated above, an average FluxGuidance of 4.0 was maintained, while the prompt remained unchanged from the previous one.

We can say that as the denoise increases, the appearance of the image changes completely and no longer follows that of the input. To remain fairly consistent with the starting image, it is important to maintain a denoise value between 0.25 and 0.50.



116 // input image



117 // img2img with Flux model, guidance : 4.0, denoise : 0.5

The process of image to image and upscaling using flux gave good results and was mainly used to maintain consistency between the images, in practice the artificial intelligence in this case was used as if it were a boosted Photoshop, and this is because we needed a high degree of conditioning on the input image, which we can hardly achieve with a simple text to image

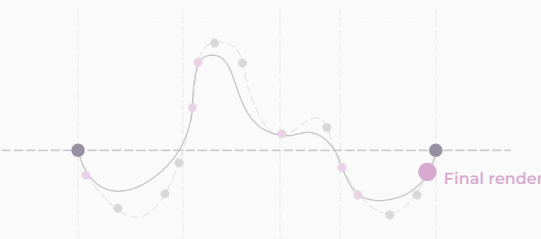
On the left we see the substantial difference between the two images, what we were looking for was a consistency in the shapes and the removal of anything that might look wrong or unrealistic.

In addition to correcting the shapes of our structures, it added a lot of detail to the image, especially in the foreground, making it much more realistic, even the people have been corrected and are less deformed than before.

The average time it takes to generate one of these corrected images is about 3 minutes with an RTX 4090 GPU, of course at the beginning there can be various tests that increase the time it takes, but once the correct workflow is set it's a very fast process.



118 // comparison of the initial image and the final image

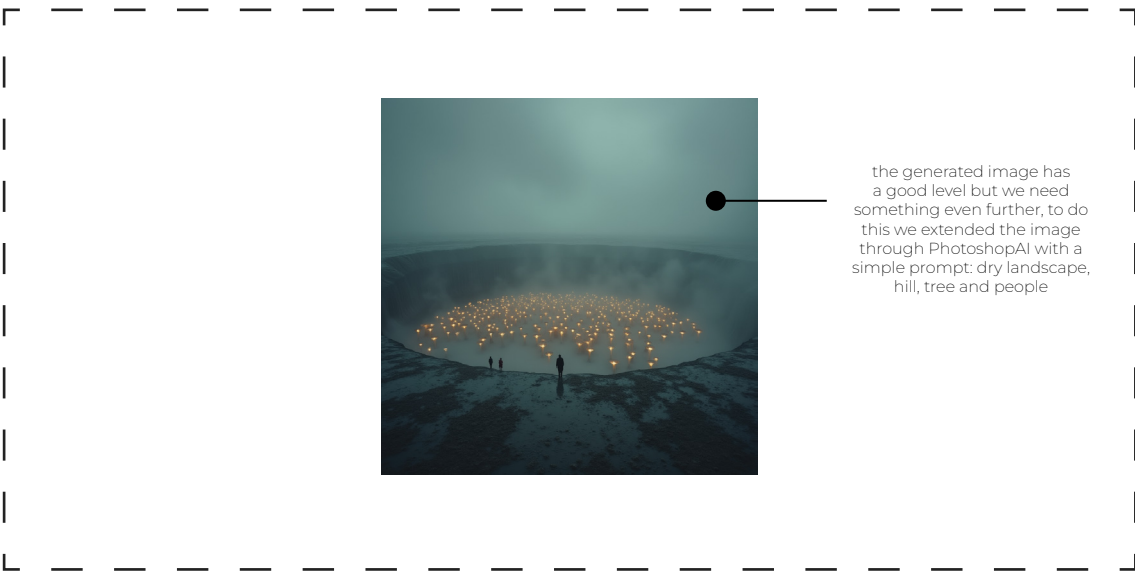


For the continuation of our project we need different views of the same project, namely we need a far view and a near view.

As far as the distant view is concerned, it was simple because many details are not visible, the starting image was created as done previously by a Text to Image using the same prompt but with slight modifications to arrive at the desired output.

prompt img01 : a dramatic scene from a human perspective, featuring a huge crater filled with numerous small, human-height elements shaped like mushrooms. Human figures walk among. The mist wraps around the glowing forms, creating a dreamlike atmosphere under a dark, overcast sky. The reflections on the damp ground further enhance the mystical ambiance, blending light, shadow, and fog into an ethereal composition.

prompt img02 : **an aerial perspective** showcasing a desolate **flat landscape** and a huge **round** crater filled with countless elements shaped like mushrooms with different heights. The scene is shrouded in mist, with glowing forms illuminating the fog, creating a dreamlike atmosphere. Human figures are visible, walking among the mushroom-like structures, appearing small and insignificant from above. The dark, overcast sky looms over the landscape, while reflections on the damp ground blend seamlessly with light, shadow, and fog, creating an ethereal, mystical composition.



119 // birdseye image of the project input

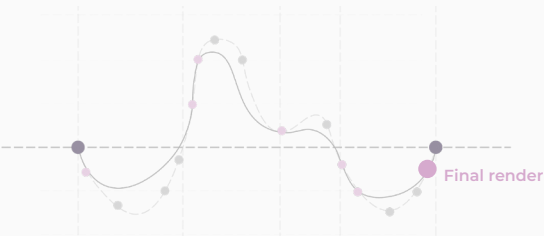
PhotoshopAI works practically like an InPainting, i.e. a Text to Image that is applied only to the indicated area and takes into account what is around it trying to match the rest of the image.



120 // input image after photoshopAI implementation



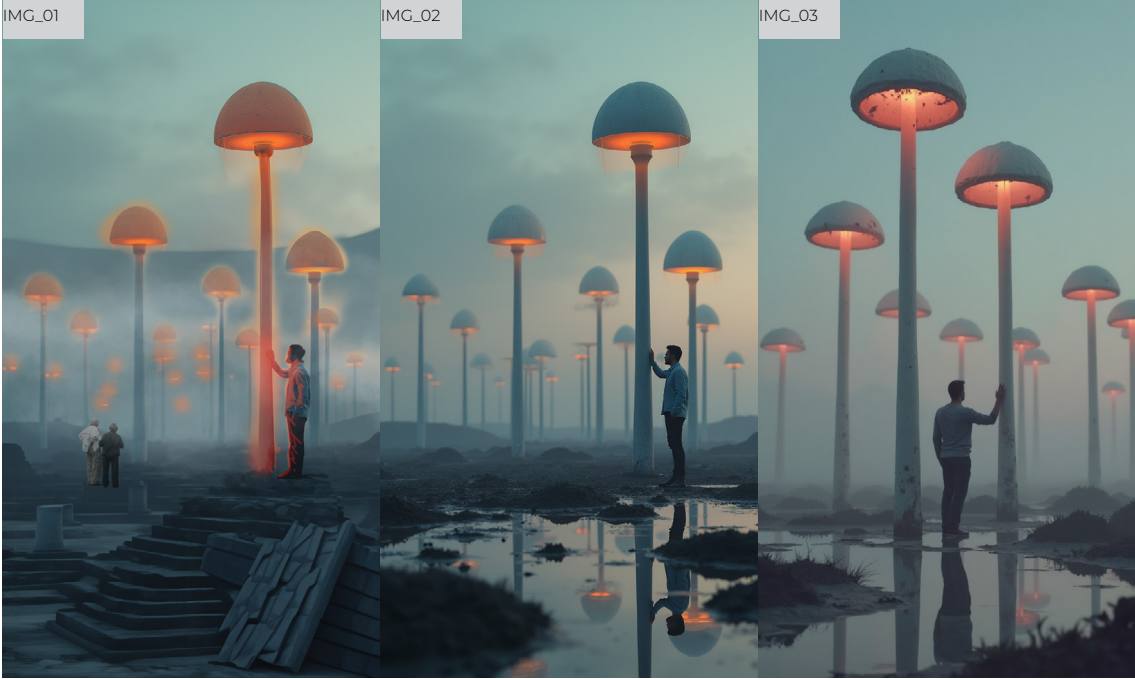
121 // image after Flux Img2Img and upscaling process



For the distance view it was very easy, but when we got to the close up view we started to encounter the first difficulties due to the limits of AI today, it is difficult to represent the same elements with a different view or angle, and due to this we had to do a bit more work behind the scenes, a constant ping pong between Photoshop and ComfyUI.

As operated before, we generated new images by readjusting the prompt with some keywords to represent the new scene.

prompt img03 : create a surreal, futuristic landscape at dusk featuring a vast field of slim towering with a **mushroom head as top part**, with random heights with soft orange lights emitting from their rounded caps and penetrating also the column. The ground is reflective, resembling shallow water, with scattered dark soil mounds and subtle mist in the background. **A man stands next to one of the structures, touching it.** The atmosphere is tranquil and otherworldly, with a muted, soft gradient sky transitioning from blue to warm hues



122 // different images generated with Flux with the new prompt

They were then processed with the same workflow as before, then an image to image and an upscaling with Flux to try to improve the output and make it consistent with the other project images. To achieve the final result, parts of different images were merged in photoshop, and a method of blending the different images was used.



The final objective was to create a close-up image in which it is clearly visible that the mushroom-shaped structures can light up when touched by humans, through quick manual work we combined several images in order to get the effect and vibe of the image just right, in the images above we have highlighted what was taken from each image.

123 // blending different AI images



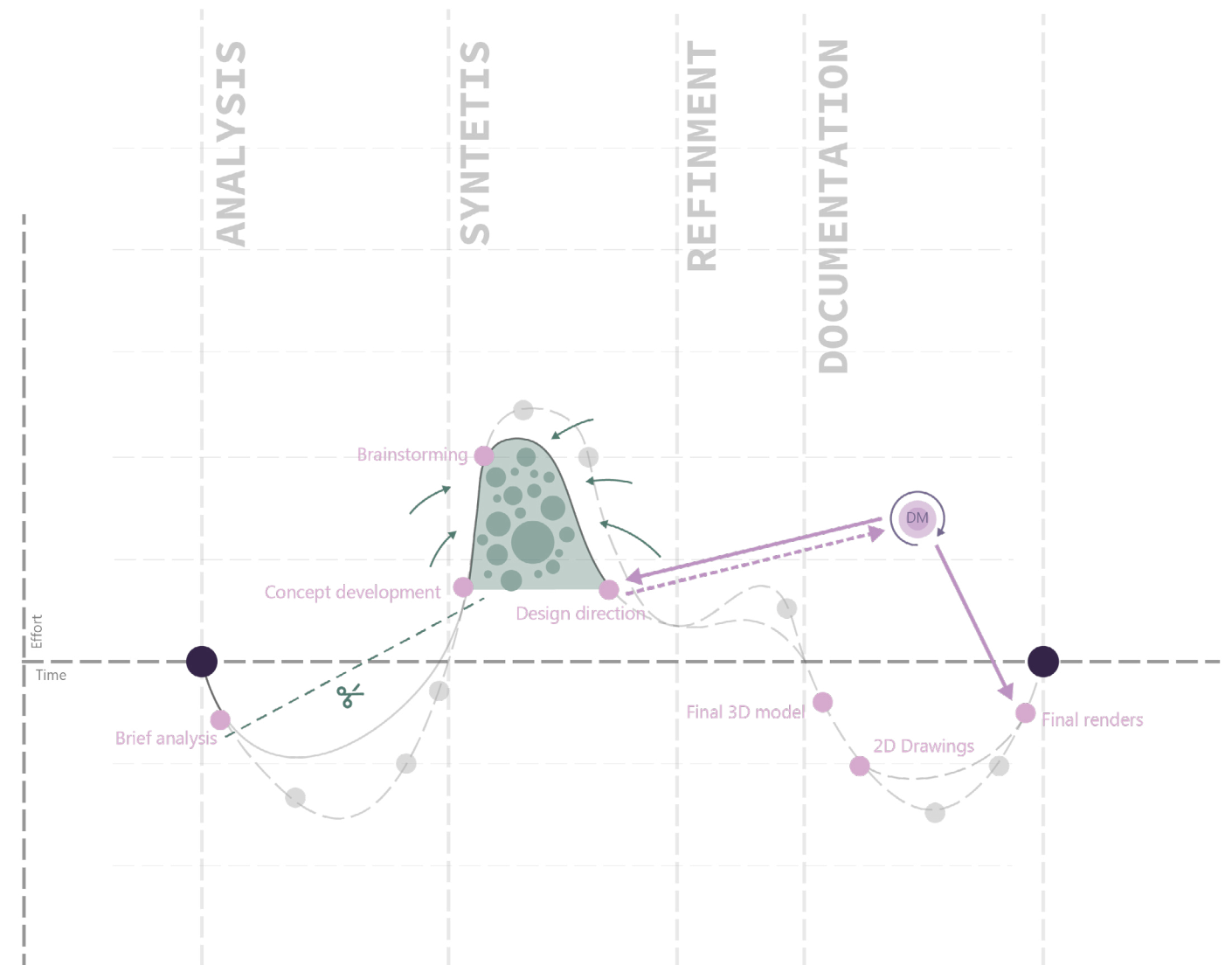
The Dialogue Conclusions

In the diagram on the right, we see how the dialogue process has changed the traditional process.

As we have already seen, the first part of the analysis and concept could lead to a creative increase in the proposed solutions and a shortening of timeframes.

In this case for this specific project, we do not have the requirement for the production of floor plans, elevations and sections for example, as well as various diagrams, which is why we can consider the dialogue process as a process that can be considered valid, but we must point out that this competition does not require the production part, except for the final renders.

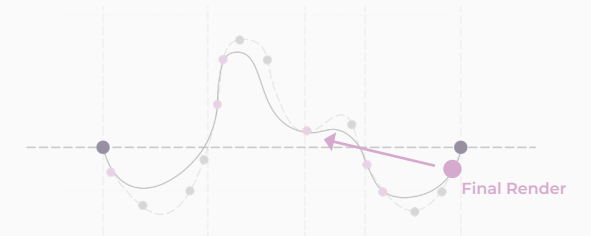
In the following examples, we will further see how the architectural creative process can change with the integration of artificial intelligence.



125 // result of the process with AI integration project 1 process 1

AI Process - The Inversion

The Approach



126 // starting image created with artificial intelligence

In order to meet the requirements in the production phase, we must create a model that allows us the subsequent technical production.

We saw how the first process failed, so we modified the process by integrating a traditional manual modelling part.

The image created with Flux was the starting point for this process, and from it we took the inspiration to create all the necessary deliverables. In the first avenue explored on how to integrate AI into the architectural process, we decided to follow the process of inversion in which we create images using artificial intelligence and then create our 3D model.

We opted for the idea of having numerous lighting elements that resembled mushrooms of different heights, we then created a small GrassHopper script that would allow us to quickly model and create these elements, from here we set up and exported completely white views on Enscape that would act as an input image for the subsequent manipulation and generation by artificial intelligence.

With this process it is easier to set up new views in this way, having total control over our views, it is difficult and laborious to find new views that talk about the same project through artificial intelligence.

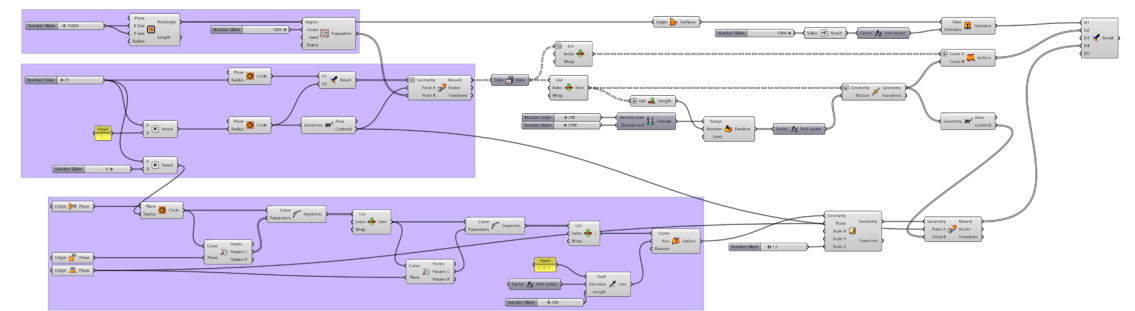
The 3D Model

Applying this process from the final design image, we create our 3D model that enables production.

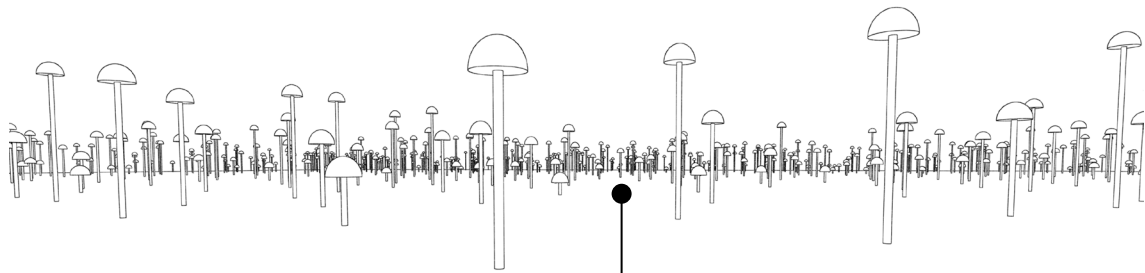
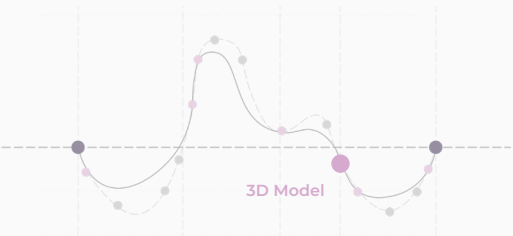
In this case we created our model using a Grasshopper script and then imported it into Rhino and generated the input images using the Enscape rendering engine.

As we see in the following images we have at first faithfully reproduced the 'atomic mushrooms' present in the reference image, but having complete control we can also modify their structure and shape and proceed with the one we are most interested in.

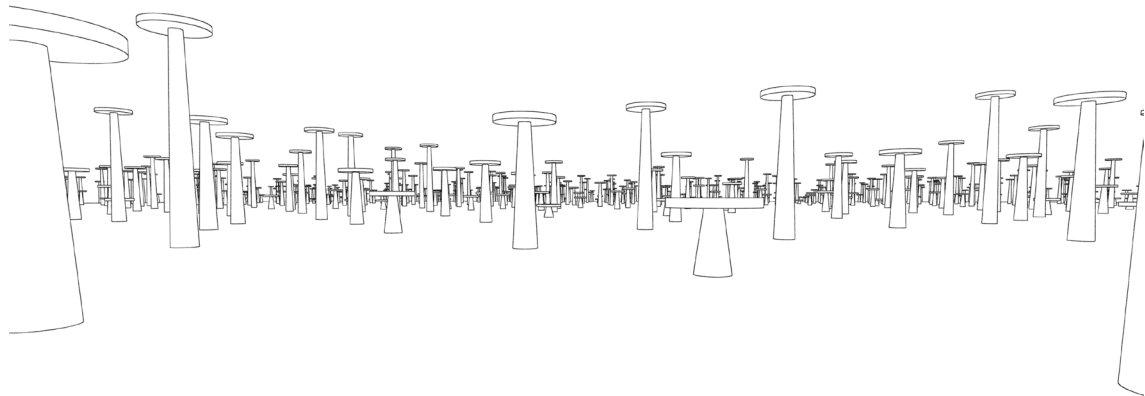
In this case to explore different possible actions with the artificial intelligence we have recreated the final renders again, but if the level of the final image is adequate there is no need to go into this process but to use the model only to produce the required technical drawings.



127 // GrassHopper script for the 3d modeling manipulation

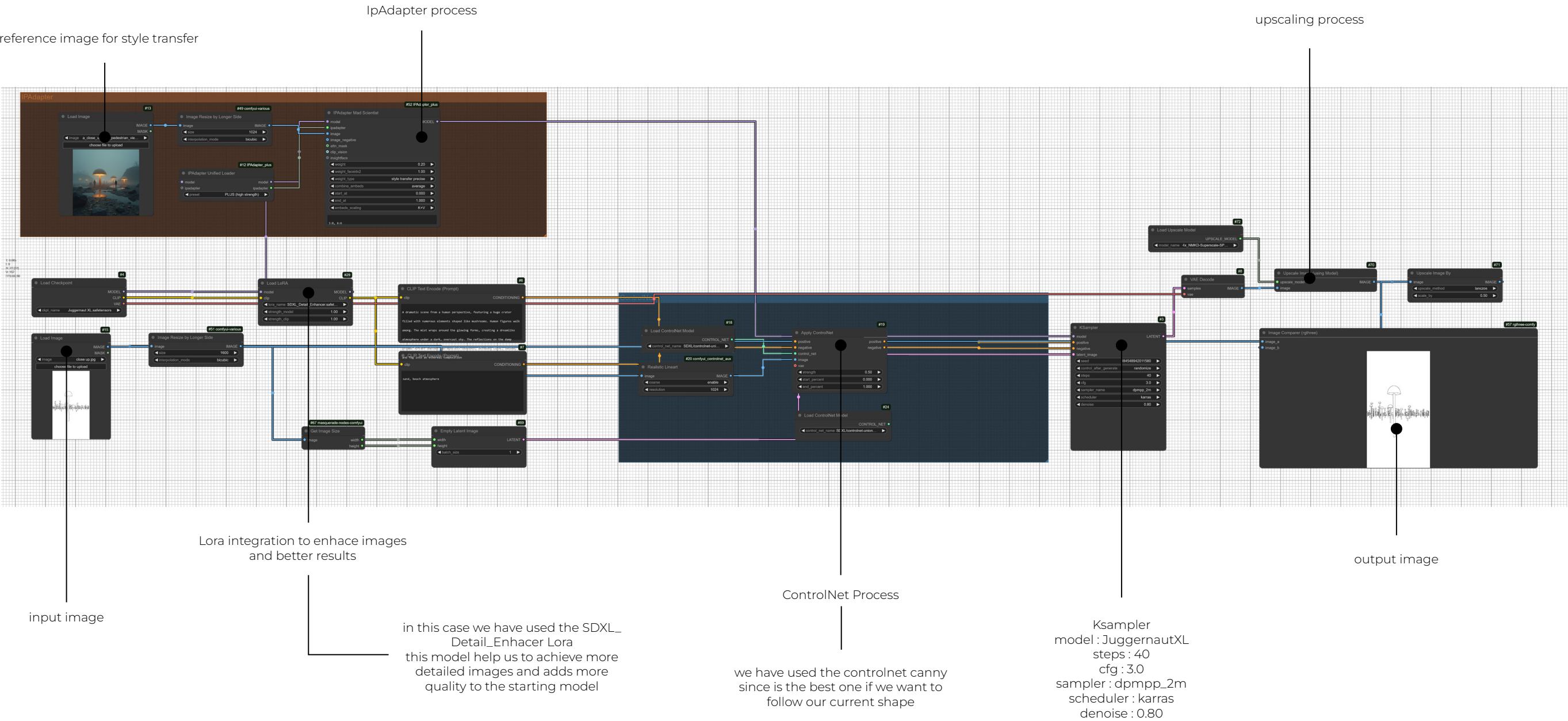
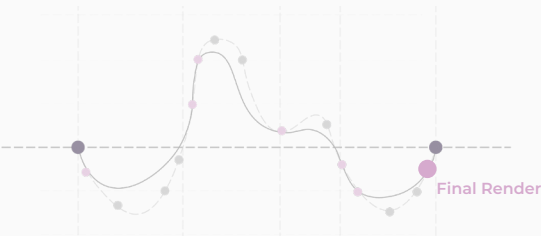


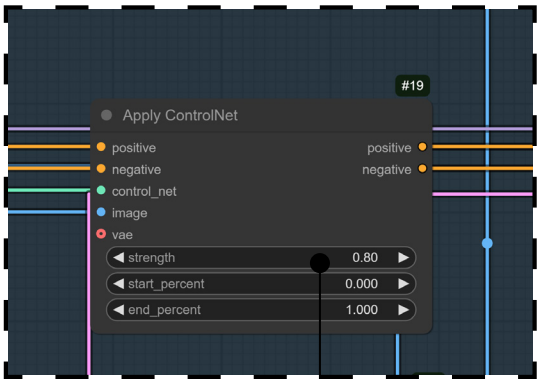
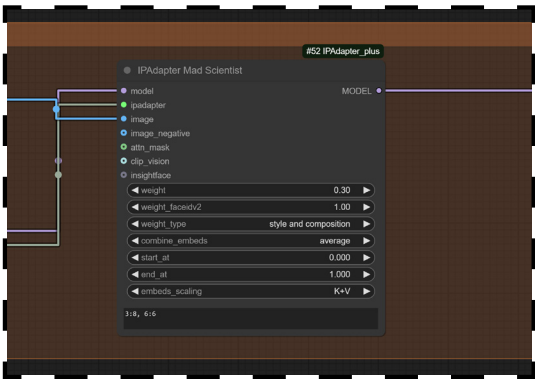
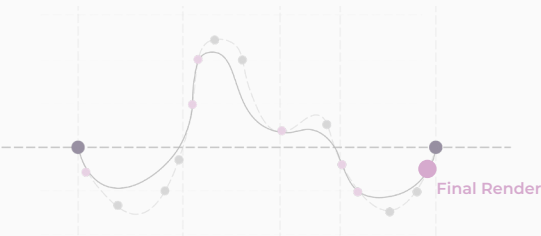
it is very important that the image has well-defined lines and shapes, while the style and mood of the image does not matter, we opted for a basic white image for its speed of creation



129 // input image 02

Workflow Explanation





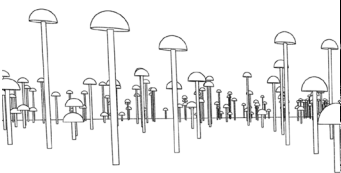
main parameter controlled in the controlnet process, the strength influences how much the parameter affects the final image, a high value will respect the visual constraints imposed by the controlnet

strength < 0.50 = low influence
strength 0.50 - 1.00 = medium/high influence
strength > 1.00 = high influence

reference image



3



prompt : glowing elements shaped like mushrooms, dramatic atmosphere, 4k, UHD, crater landscape



ipadapter model : PLUS
weight ipadapter : 0.2
weight type ipadapter : strong style transfer
controlnet strenght : 0.80
denoise : 1.00



ipadapter model : PLUS
weight ipadapter : 0.2
weight type ipadapter : strong style transfer
controlnet strenght : 0.80
denoise : 1.00



ipadapter model : PLUS
weight ipadapter : 0.2
weight type ipadapter : strong style transfer
controlnet strenght : 0.80
denoise : 1.00



ipadapter model : PLUS
weight ipadapter : 0.2
weight type ipadapter : strong style transfer
controlnet strenght : 0.80
denoise : 1.00

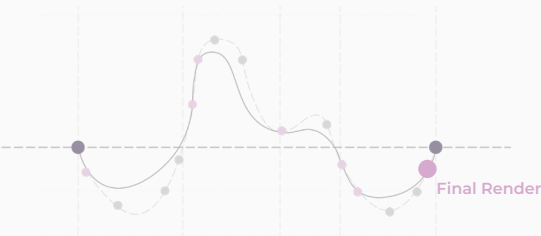


ipadapter model : PLUS
weight ipadapter : 0.2
weight type ipadapter : strong style transfer
controlnet strenght : 0.10
denoise : 1.00

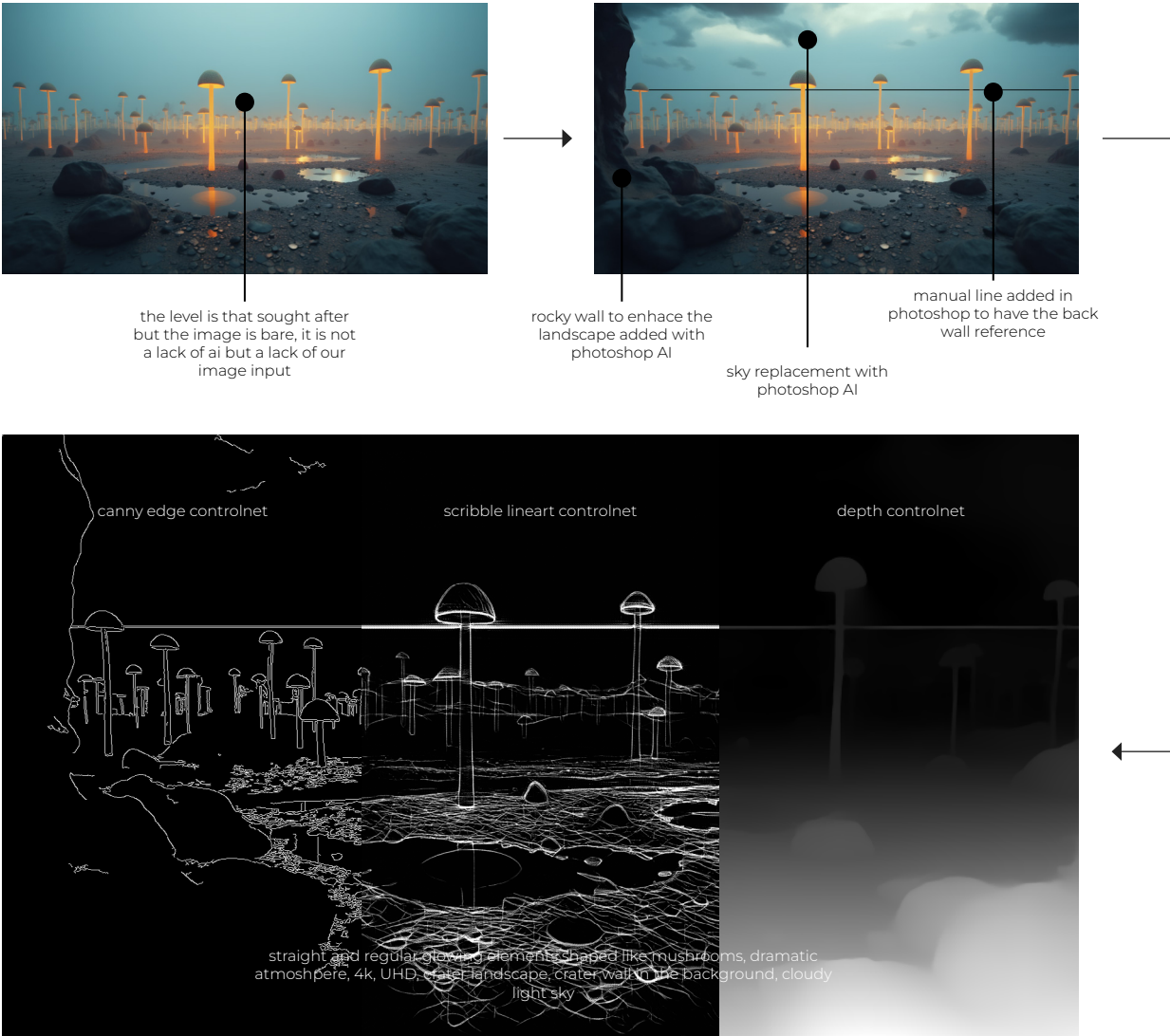
minore forza del controlnet
significa minore coerenza con l'immagine input



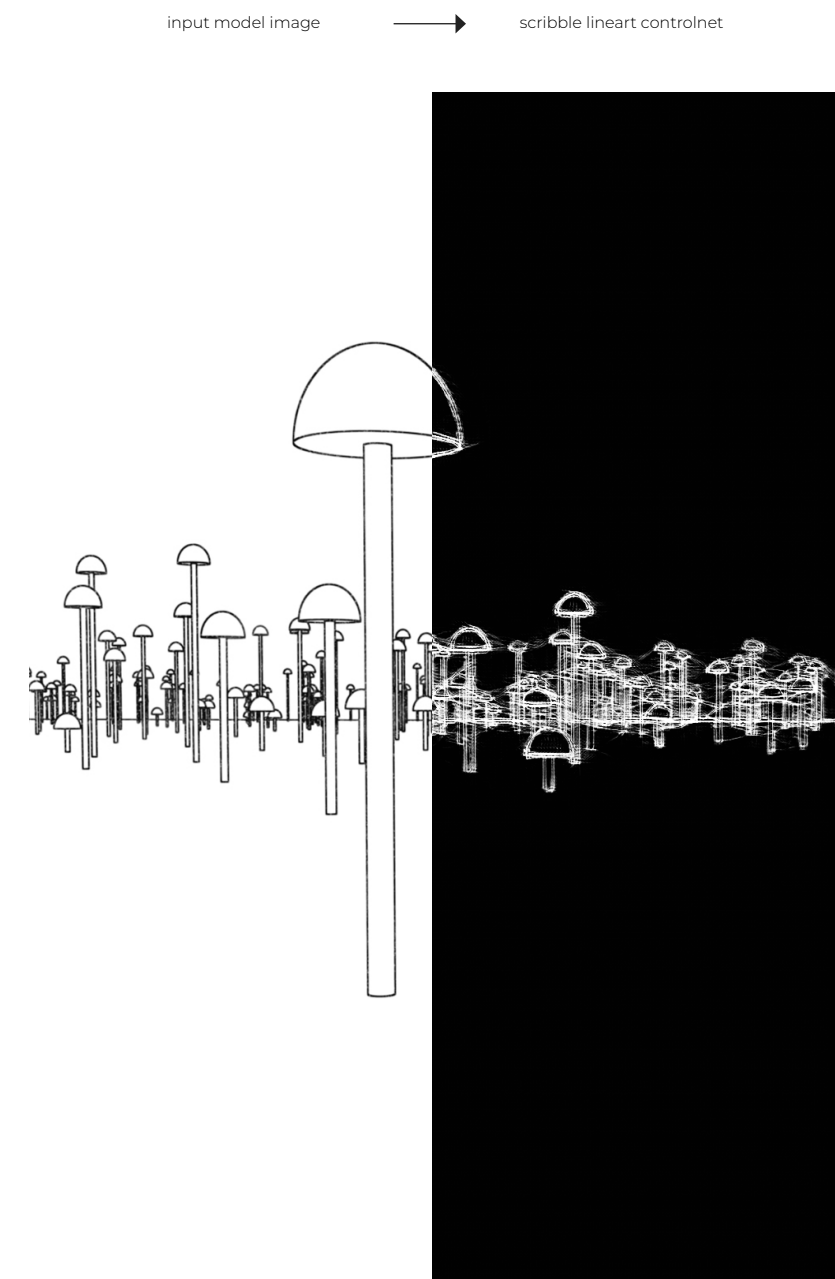
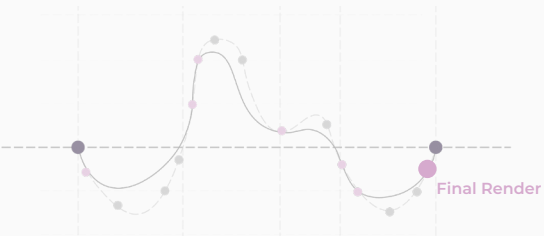
ipadapter model : PLUS
weight ipadapter : 0.2
weight type ipadapter : strong style transfer
controlnet strenght : 0.80
denoise : 1.00



In the model we have produced the level of detail is practically minimal, but thanks to AI we are able to increase this, however it may not be enough, in our case details are missing that would commonly be present in such a project, we quickly created guides in photoshop and re-processed everything by the AI with the exact same process, changing the basic prompt slightly as we are trying to add elements.

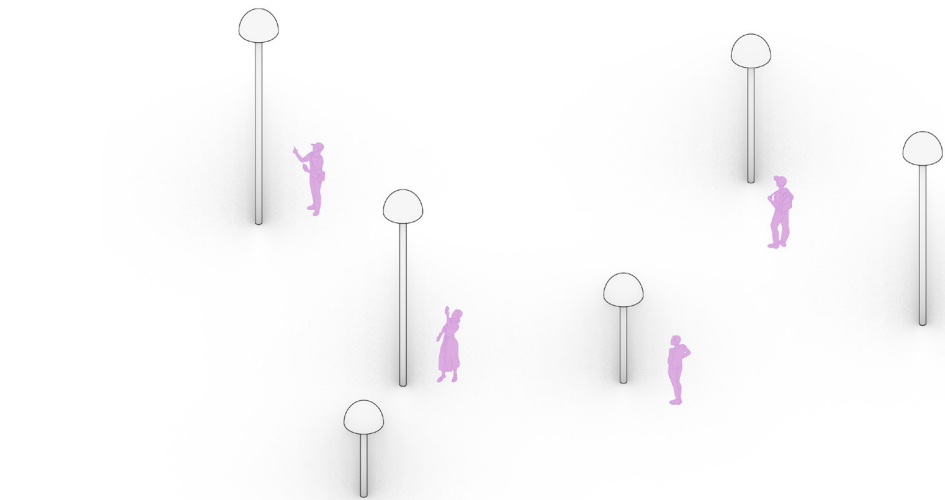


Our goal is not to demonstrate the perfect project or concept but to understand how AI can be integrated into the architectural process today. As far as this image is concerned, we consider ourselves quite satisfied with the results achieved. Only 5 years ago a process of this kind was unthinkable in the world of architecture, obviously there are small defects and details that are not purely correct but we must remember that at the moment we are using models and nodes that have not been developed specifically for architecture, furthermore the manual work to fix details or parts of the image can be easily integrated and does part of the process.



The Production

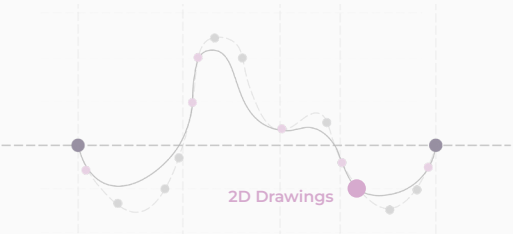
Thanks to modelling, we can proceed with the production of diagrams, floor plans and various sections, which in our case we have also produced in order to go into detail about further processes possible with artificial intelligence as seen on the previous pages.



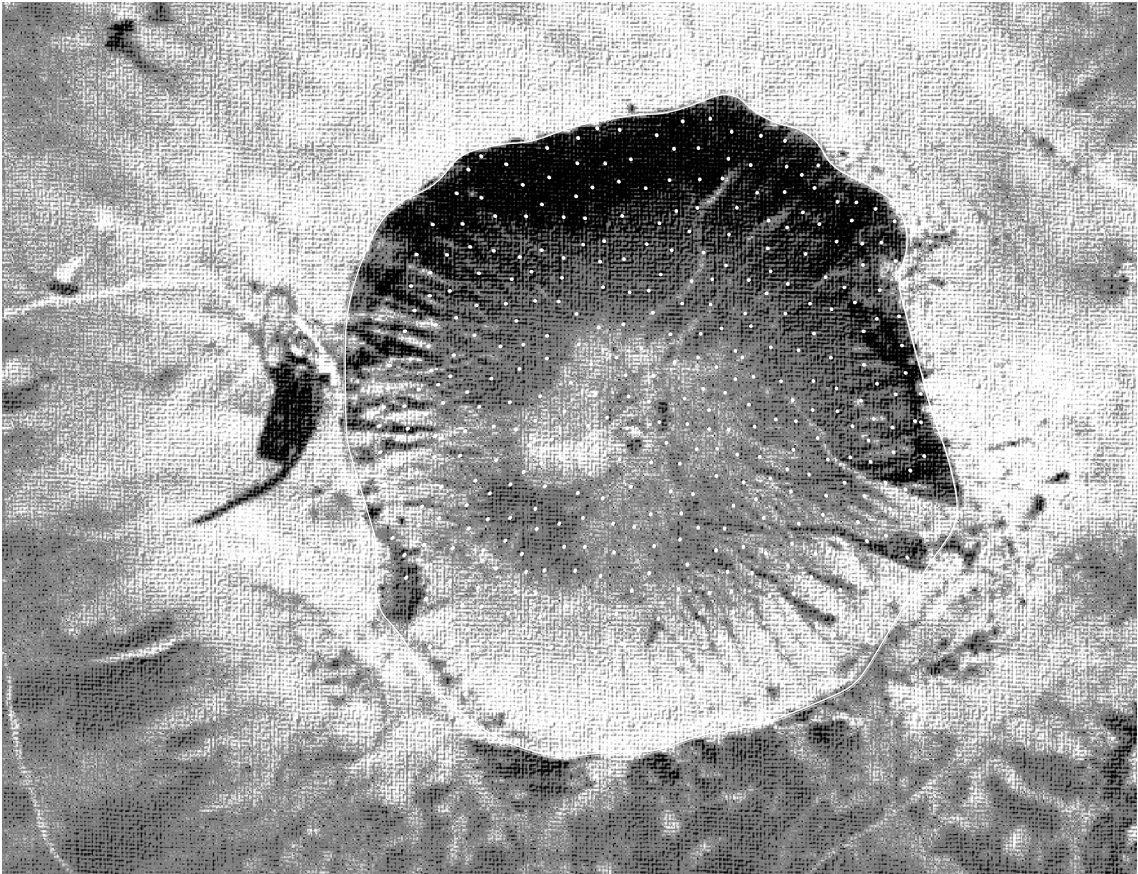
134 // axonometric scheme



135 // section



2D Drawings



136 // plan



The Inversion Conclusions

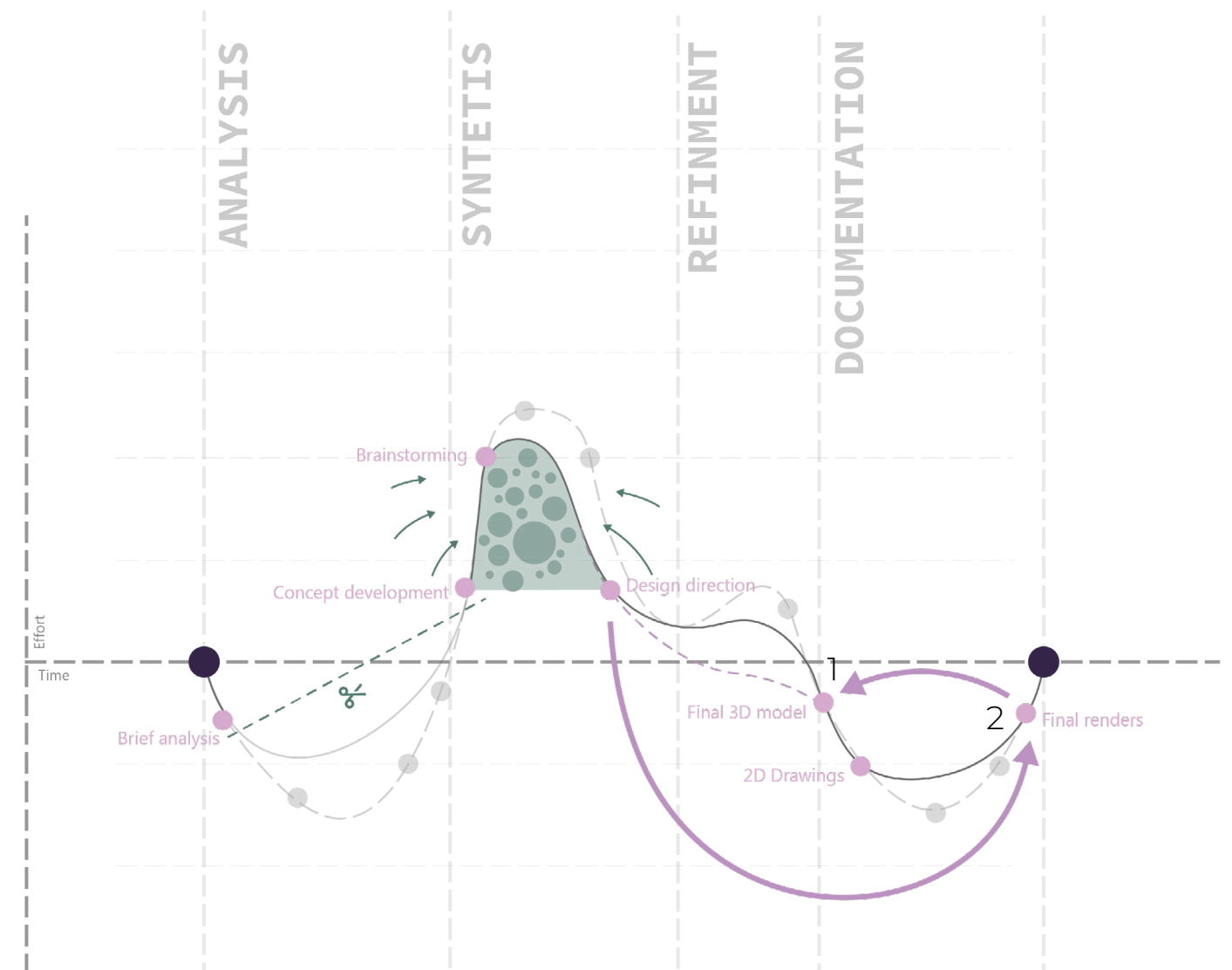
The sidebar shows how the traditional process is integrated with The Inverted process that includes the integration of artificial intelligence.

This is the process that has been successful because, as mentioned above, we have integrated parts of the traditional process to meet all requirements.

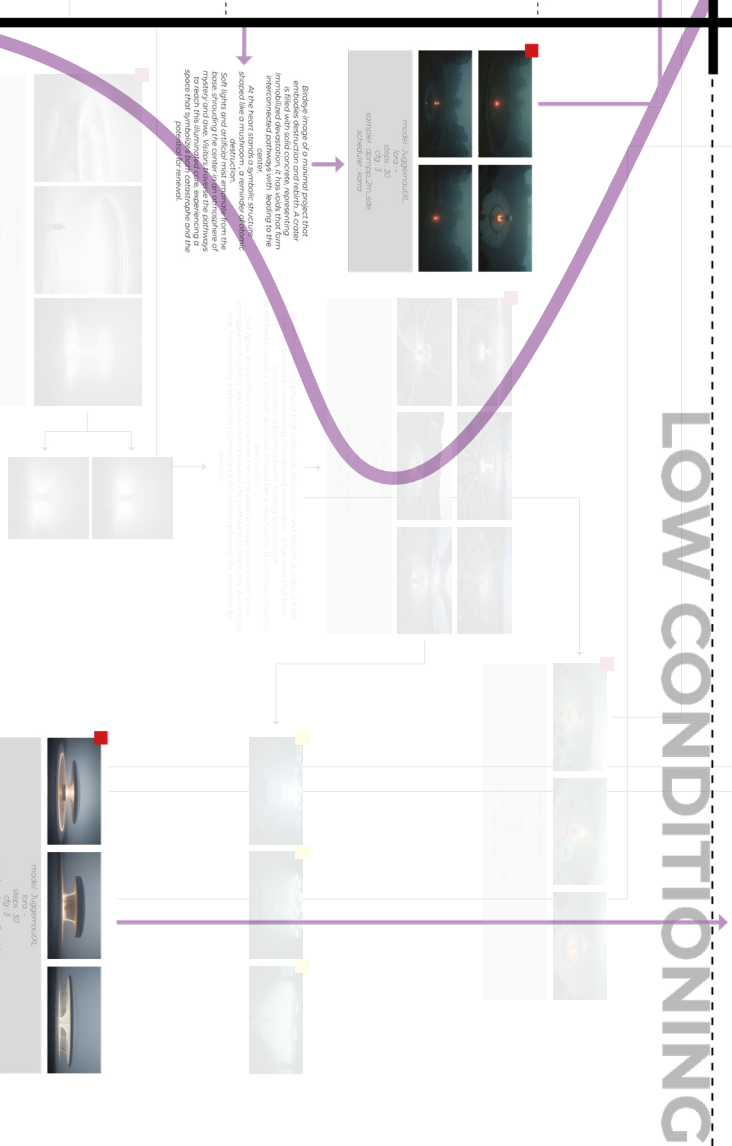
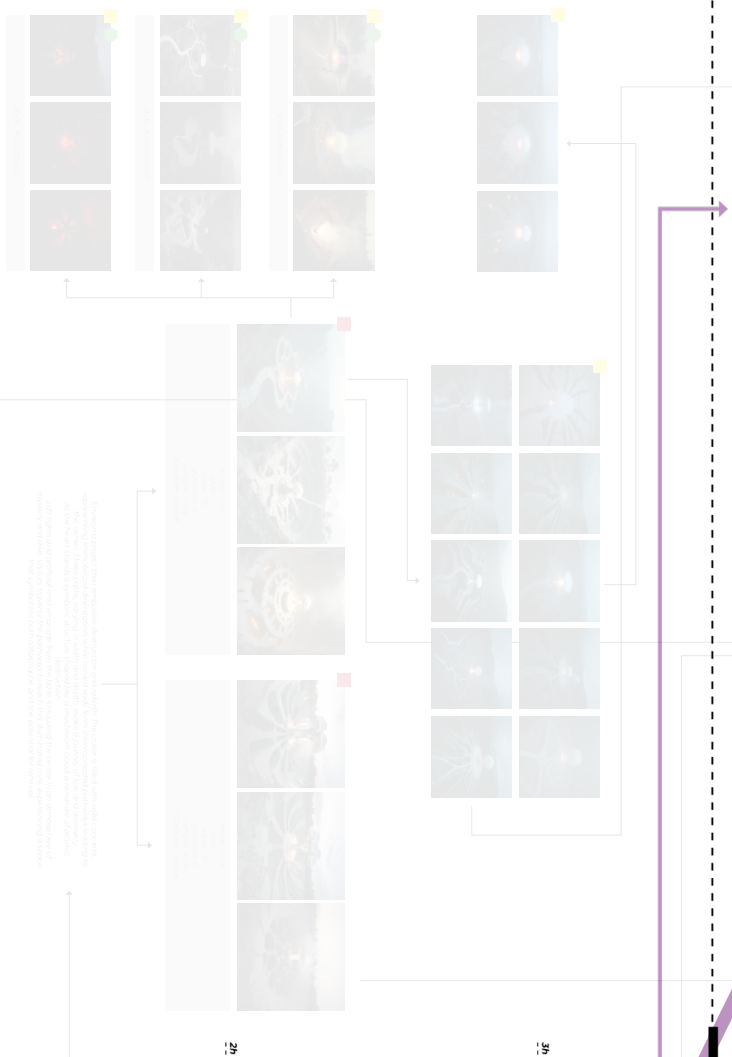
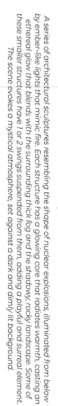
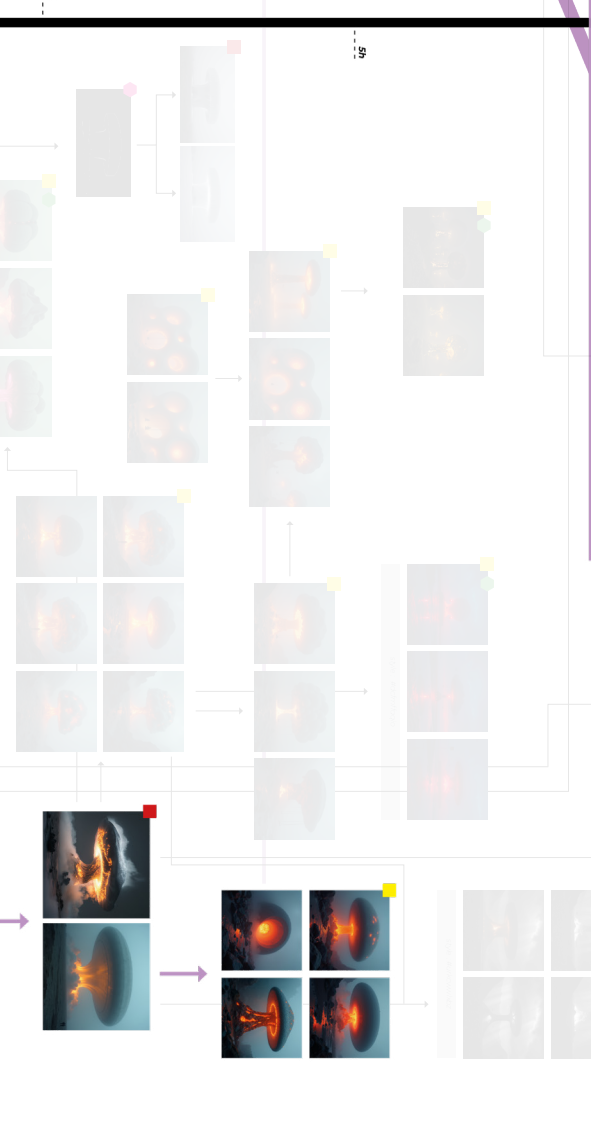
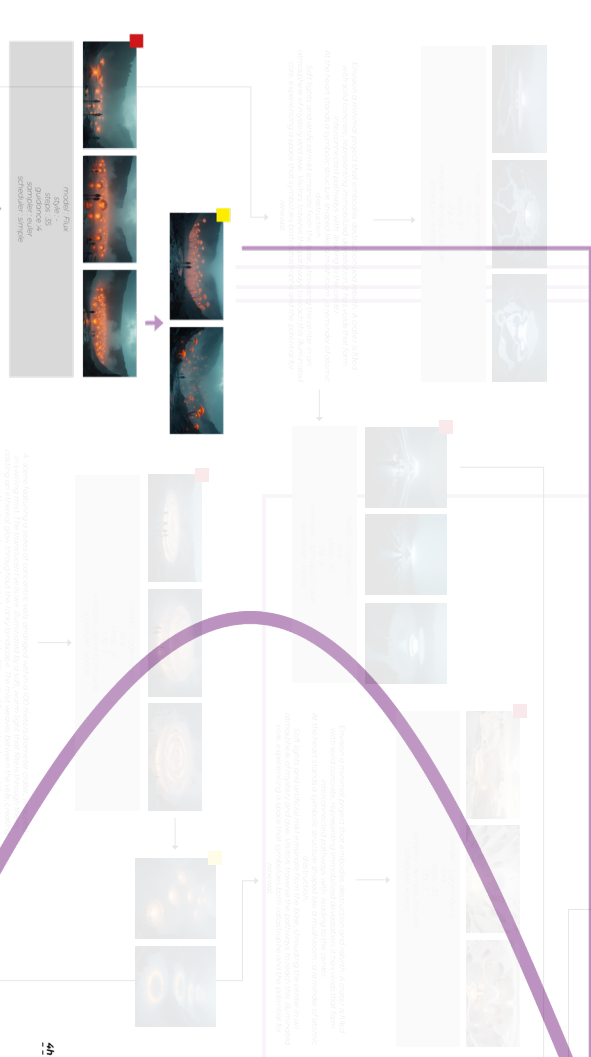
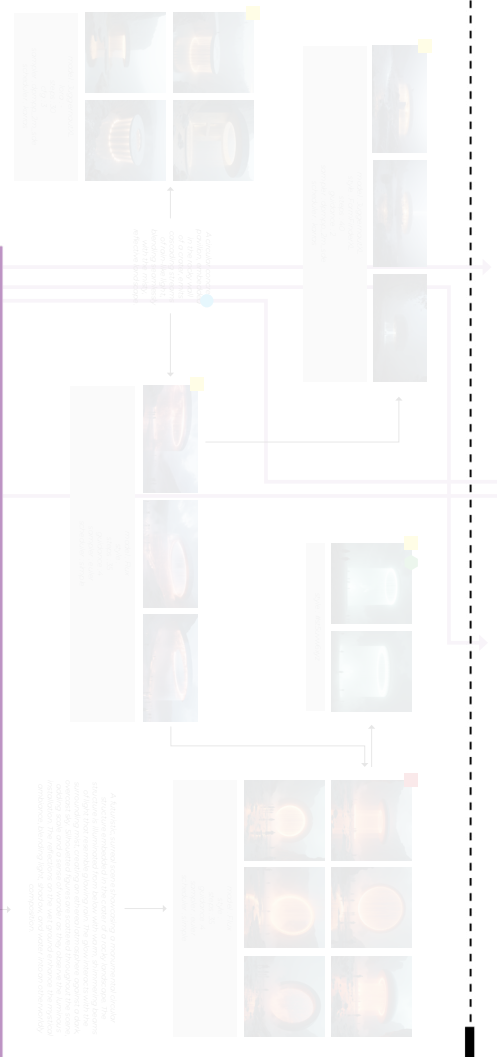
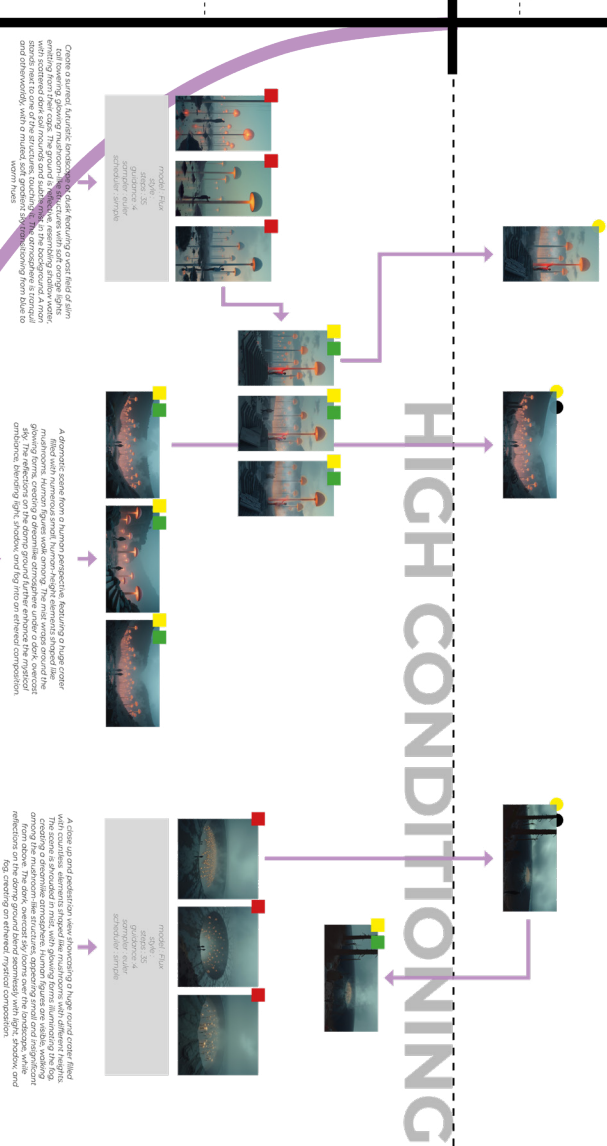
To the question we asked ourselves at the beginning whether artificial intelligence can be integrated into the architectural process, we say yes, but with limitations, at certain times and points we are not yet able to use it, architectural processes require high degrees of precision that we are not yet able to achieve.

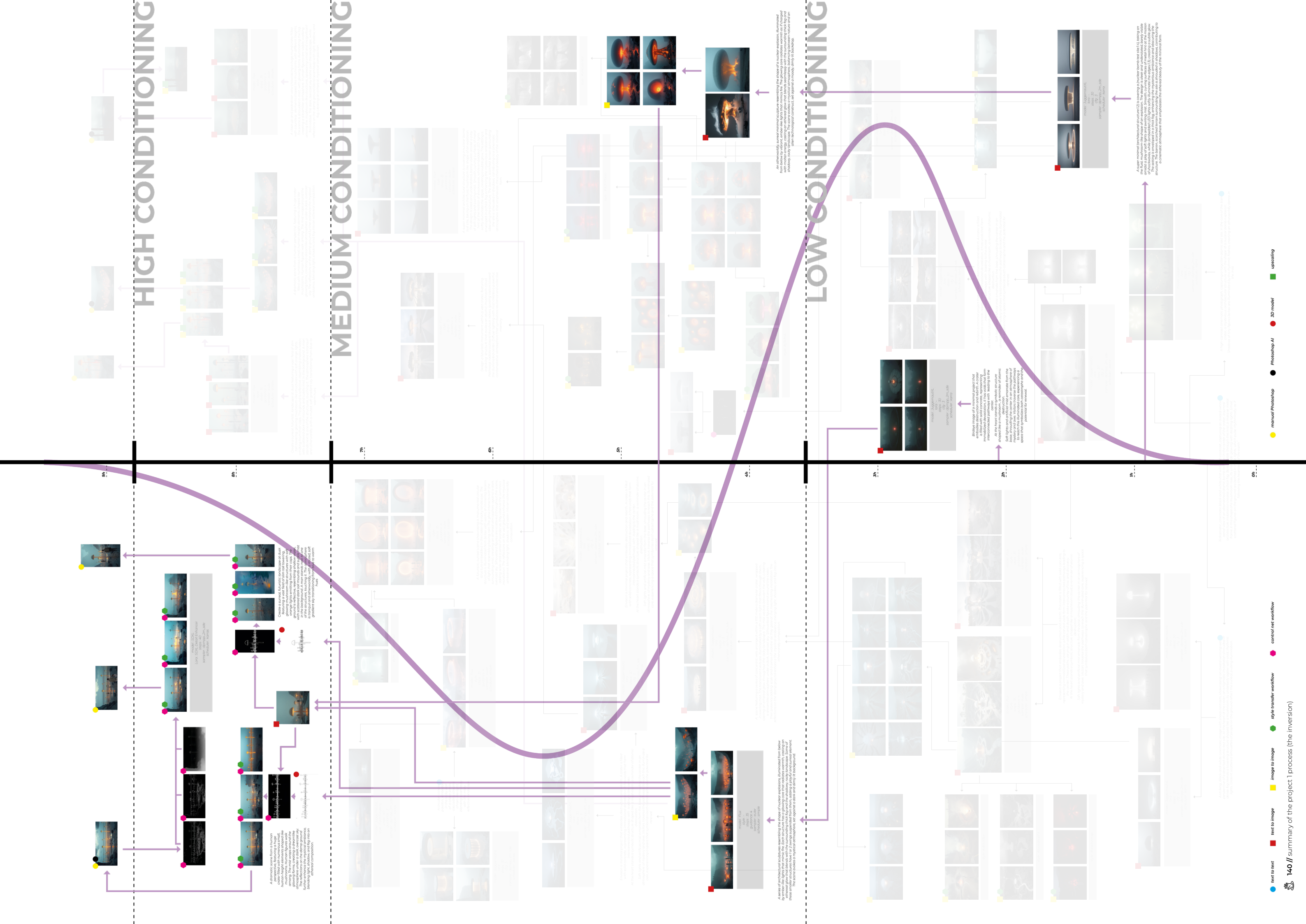
However, we can see how this new process has changed the traditional one by inverting the common working method, in a traditional process you follow a logical thread that after brainstorming and searching for the initial concept you start a process of modelling and searching for the suitable design to then arrive at the final image and consequently its elaborations.

With the integration of AI, in addition to the possibility of shortening timeframes, we see how points (1) and (2) are reversed, giving rise to a new process and production thread in which the final image is first created and then the 3D model.



138 // result of the process with AI integration project 1 process 2





HIGH CONDITIONING

MEDIUM CONDITIONING

LOW CONDITIONING

How can we customise artificial intelligence ?
What is a LoRA ?
How much influence can we exert on artificial intelligence ?

02 - Rethinking the Grattacielo della regione piemonte

Experimentation for Effectivness

Sitography

// <https://torino.repubblica.it/cronaca/2020/02/09/news/>
// <https://www.facebook.com/share/p/18NGTcPvbf/>
// <https://www.facebook.com/share/p/19VUEFtb3o/>
// https://www.regione.piemonte.it/web/sites/default/files/media/documenti/2022-08/torre_grattacielo
// https://it.wikipedia.org/wiki/Grattacielo_della_Region_Piemonte

141 // shot of 'Palazzo della Regione Piemonte' by Massimiliano Fuksas

Understanding the Skyscraper

Precondition

The Piedmont Region Building is the subject of much criticism and has been targeted by a great many citizens of Turin, the main accusations being the ruin of the Turin skyline and the aesthetics of the Skyscraper itself.

We have collected some of these criticisms and reported below to realize the general popular opinion, we believe that social opinion is of paramount importance since it is the everyday citizens who perceive and experience the skyscraper in their city.

Vittorio Sgarbi, a well-known Italian politician, bitterly criticized the skyscraper and Fuksas' work.

And Fuksas himself among the various problems and delays that the Skyscraper has had has declared that he no longer considers it his project³¹.

-Can we continue to call it the Fuksas skyscraper ?
-Call it Arturo. Do you like the name Arturo ?
-Yes.
-Then call it that.
(Fuksas, La Repubblica, 2020)

³¹Fuksas interview for La Repubblica (source : <https://torino.repubblica.it/cronaca/2020/02/09/news/>)
³²Facebook's posts about the skyscraper : <https://www.facebook.com/share/p/18NGTcPvBF/>
<https://www.facebook.com/share/p/19VUEFtb3o/>

Below are some comments made by the citizens of Turin in the Facebook pages³² and articles found about it, the comments are really many, and we feel to say that 95 percent are all negative.



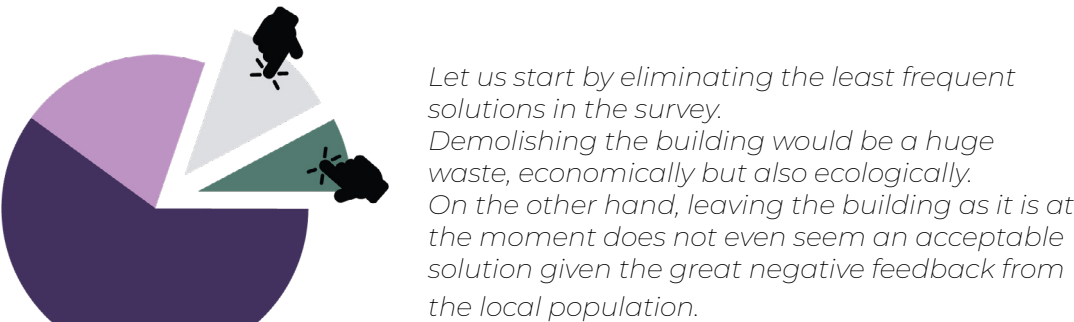
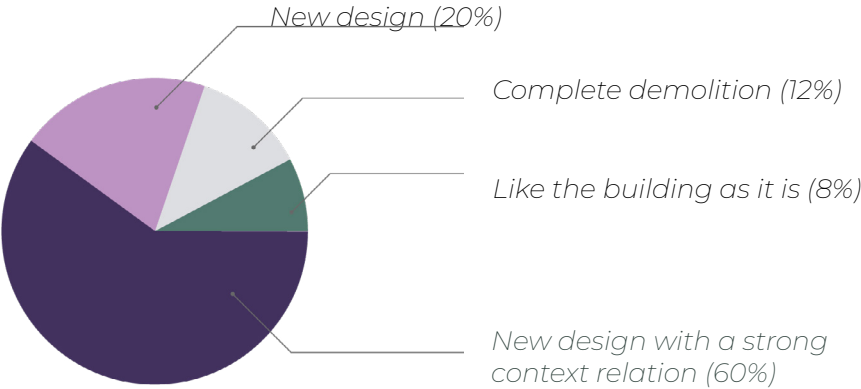
Survey

Since social opinion is quite important, we continued in this direction and asked the citizens themselves what they thought of the skyscraper and more specifically how they would like it if a redesign were to be considered.

We launched the challenge on some of the biggest Facebook groups that include citizens of Turin and Piedmont in general and asked them :

We have seen very negative opinions regarding the Palazzo della Regione Piemonte and wanted to ask the citizens who see and experience the building every day, how should the redesign be approached for a better vision ? We wanted to give some initial ideas, which could be: a redesign with clear references to the architecture of Turin and the Mole Antonelliana, or a redesign by another international architectural firm with a more iconic style for this particular project. Obviously any kind of opinion is accepted, and we invite you to write it down if you have any.

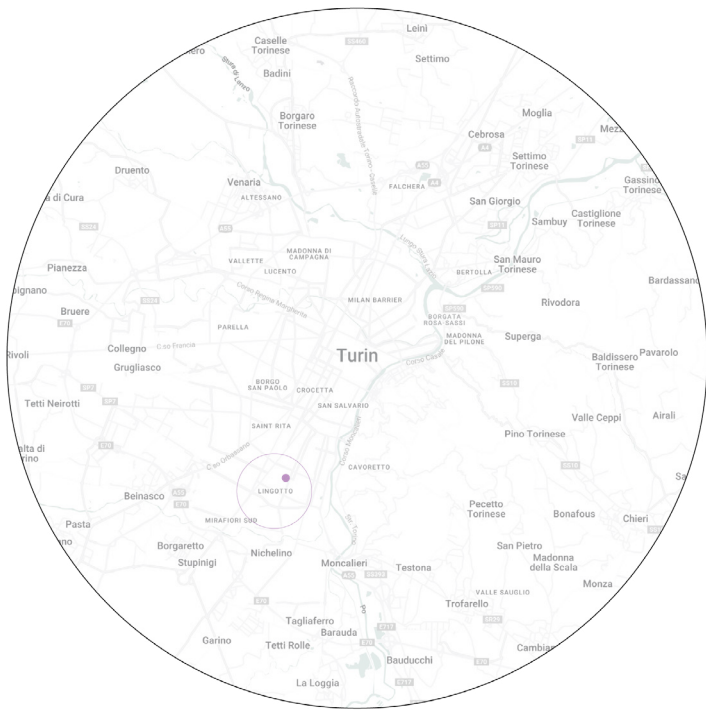
Thank you very much to those who reply!



The site

"The skyscraper of the Piedmont Region is an institutional building in Turin, and is the sole seat of the Region's central offices and organs, with the exception of the Regional Council, which has remained in the courtly Palazzo Lascaris building. Designed by Massimiliano Fuksas and presented to the Regional Council during a public hearing held in Turin on 22 November 2007, the building has 42 floors above ground, 41 of which are for civil use and the last one used as a hanging forest.

The land identified for the building's construction is an area formerly used by the former Fiat Avio in Turin's Nizza Millefonti district, close to the Oval Olympic Arena and not far from the Lingotto grounds. The skyscraper, which is served by the Italia 61 - Regione Piemonte metro station, is also connected to the Lingotto railway station by a cycle-pedestrian passage." (Wikipedia, 2024)



The skyscraper has become a landmark for the city of Turin, not being a city with a strong vertical development it immediately stands out among the Turin skyline as can be seen in the aerial view on the previous pages.

Together with the San Paolo intesa skyscraper realised in 2014 by Renzo Piano, it constitutes the tallest buildings in Turin, and with its height of 209 metres it ranks third among the tallest buildings in Italy.

The shape of the skyscraper is sleek and minimal, redefining the Turin skyline. The façade is completely glazed, maximising the amount of natural light that penetrates the building's interior spaces. The latter have been designed to meet precise operational requirements, the first floors are accessible to the public offering them various functions such as conference rooms, while the upper floors include administrative offices.

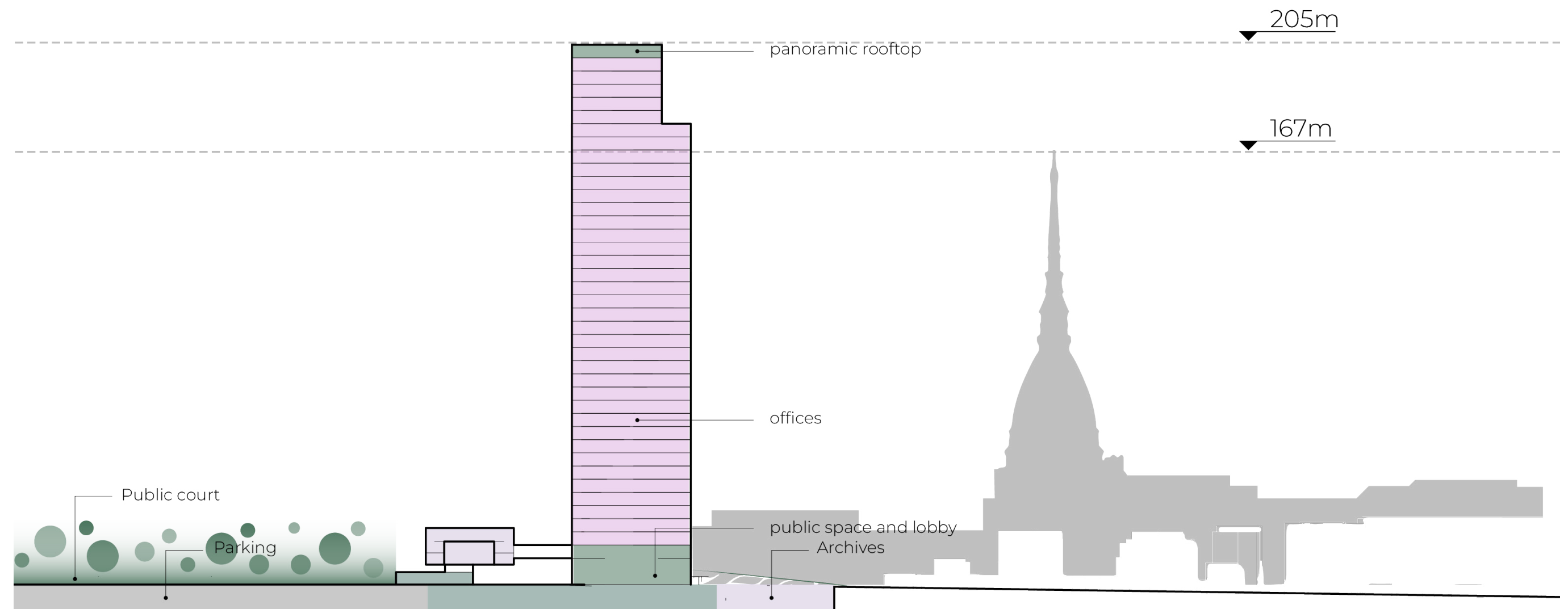
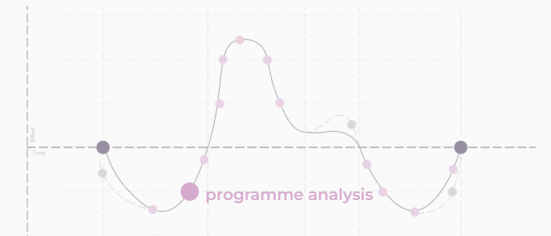
Although the aesthetics are very minimal, the building has more value from a sustainable point of view; the double façade system allows for reduced energy consumption together with advanced heating and cooling systems.

It is not only a place of administrative function, but a true symbol of Turin, obviously the incredible size compared to other Turin buildings makes it so, a reference point for citizens.

Programme Analysis

The Skyscraper of the Piedmont Region is a building mainly for offices, 41 floors above and two floors below ground for parking and public spaces.

On the ground floor, in the entrance area, there are mainly communal areas, while the subsequent floors are all fully used as offices for the region, which naturally also include conference and meeting rooms. On the top floor is a panoramic rooftop accessible to the public.

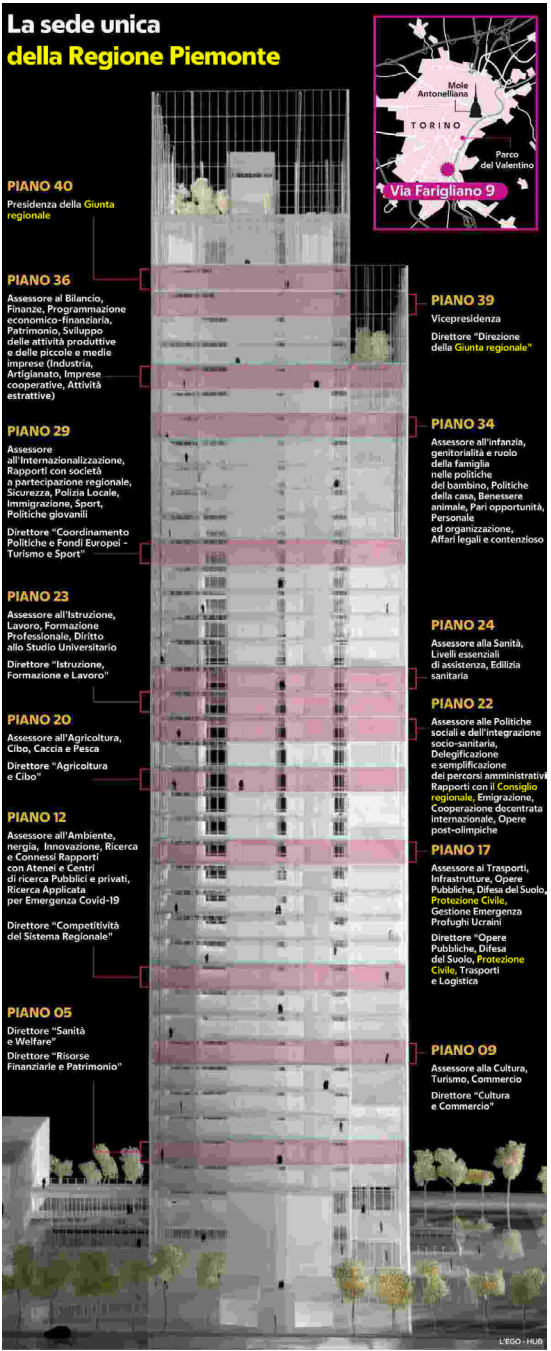


The tower has 40 levels above ground. The division into floors is made according to the management of the different departments. Therefore, lacking technical drawings, we were able to understand through the use of the official website of the Piedmont region, where the building is described, that it has 70,000 m2 of gross paved surface area, divided by the 40 levels we have a surface area of 1628m2 per floor.

Thanks to this estimated figure and the section found in an article in La Repubblica, published on 31/08/22, we were able to understand how many floors and thus square metres were dedicated to each department:

Direttore "sanità e walfare", 2 lvls, 3.256sqm	
Assessore alla sanità, 1 lvl, 1.628sqm	9.768m2
Assessore all'Ambiente, 3 lvls, 4.884sqm	
Assessore all'Agricoltura, 3 lvls, 4884sqm	
Assessore alle Politiche Sociali, 2 lvls, 3.256sqm	9.768m2
Assessore all'Istruzione, 1 lvl, 1.628sqm	
Assessore all'Infanzia, 5 lvls, 8.140sqm	
Assessore all'Internazionalizzazione, 5 lvls, 8.140sqm	21.164m2
Giunta regionale, 4 lvls, 6.512sqm	
Assessore al Bilancio, 2 lvls, 3.256sqm	
Assessore Cultura, Turismo e Commercio, 4 lvls, 6.512sqm	
Assessore ai Trasporti, 5 lvls, 8.140sqm	24.420m2

These clusters were organised from us, according to a thematic rigour that will serve as a basic rule for the arrangement of the program within the volume/volumes.

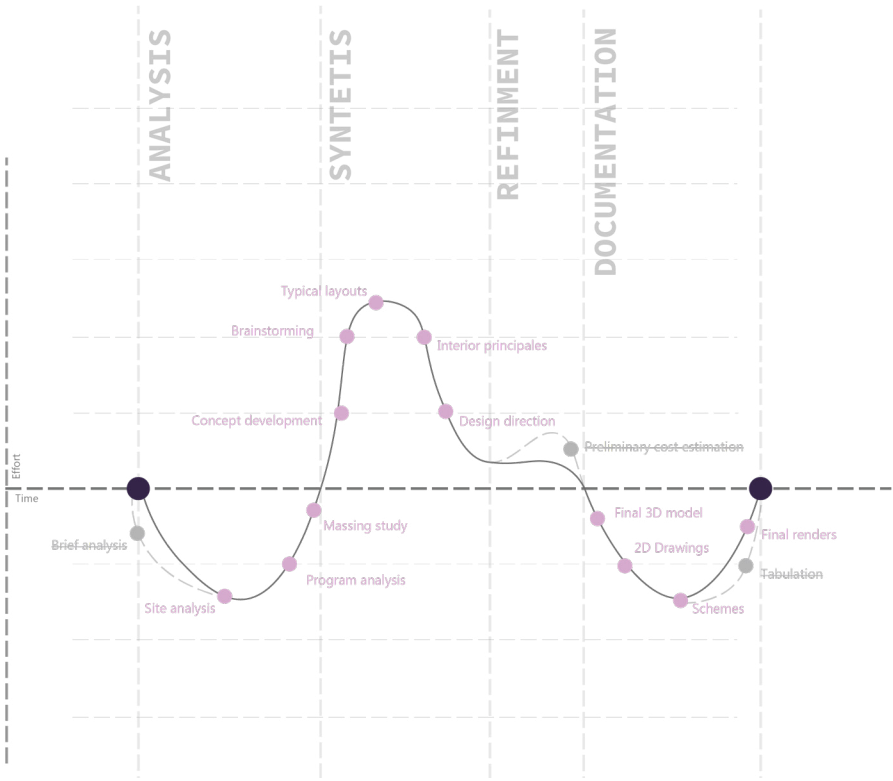
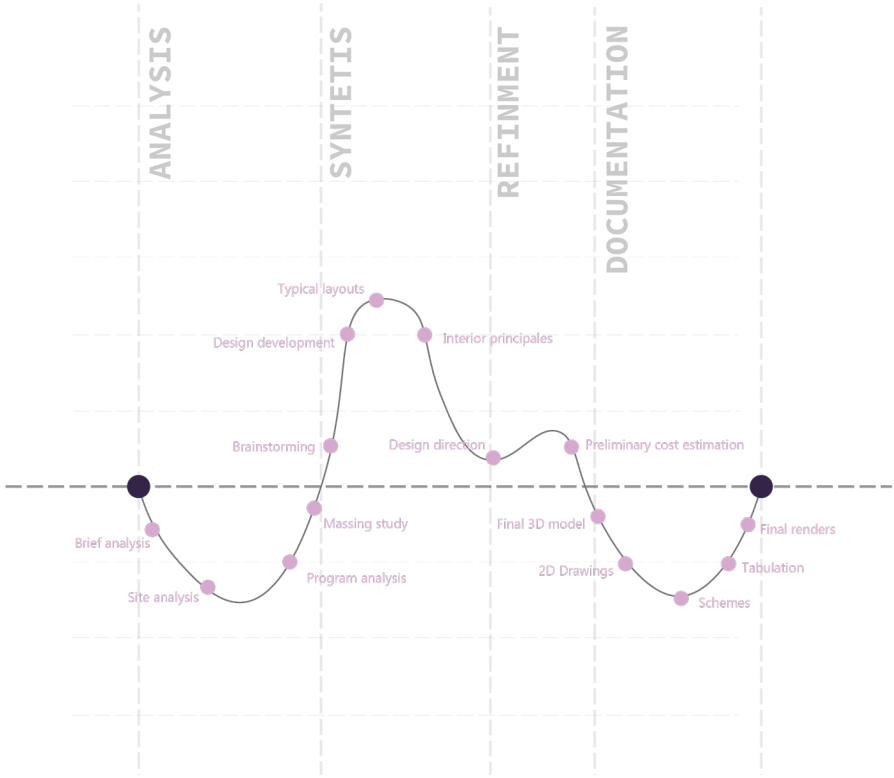


Lowering of the Curve

As we have done before, we carefully examined the early stages of architectural design, focusing on the important steps. Our careful approach enabled a complete analysis and comprehension of each key phase, revealing the connections between multiple activities and considerably improving the workflow.

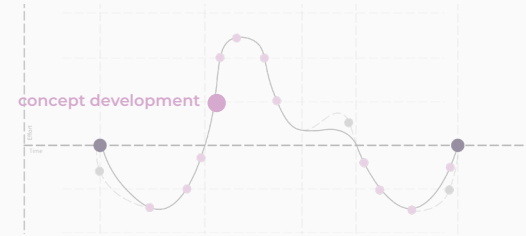
As the project we are going to tackle now is slightly different from a traditional project because it is a concept of ideas and the deliverables required are different, we have modified the scheme by adapting it to the requirements of this process.

In the final part we will see how this process has been influenced by artificial intelligence and how it has changed.



The Design Concept

Conceptualizer



The first part of the process, as also carried out on the previous small project, involves the creation of a large quantity of images through a process we have defined as 'Conceptualiser'.

To begin with, we decided to create simple images using a prompt which varied slightly during creation to explore new outputs. The objective is to have new concept ideas and immediately visualise the project, in this case we are looking for a skyscraper with modern characteristics that somehow integrates aspects and identity of the city of Turin and its symbols.

The first images were created without the help of any LoRA and we followed a process of addition and customisation.

The first 4 image files were created with JuggernautXL, a model based on SDXL, respectively with these prompts :

1 - a view of a modern skyscraper in Turin that embodies the city's innovation and dynamism while subtly referencing its identity. Incorporate elements such as sleek geometric forms, details inspired by the Savoy tradition, or nods to the city's industrial heritage, all within a bold and contemporary design

2 - a view of a modern skyscraper in Turin that embodies the city's innovation and dynamism while subtly referencing its identity. Incorporate elements such as sleek geometric forms, (details inspired by the Savoy tradition and Mole Antonelliana:1.5), or nods to the city's industrial heritage, all within a bold and contemporary design

3 - a view of a modern skyscraper in Turin that embodies the city's innovation and dynamism while referencing its identity

4 - a view of a modern skyscraper in Turin that embodies the city's innovation and dynamism while referencing the (identity and the style of the city of Turin:1.3)

While the remaining two files were created with Flux using prompts 3 and 4, obviously Flux has more power and this can be seen in the creativity and level of detail, the models we have trained are based on SDXL and not Flux because to do so requires enormous computing power which we do not have at the moment.



Fine-Tuning Methods

Different types of fine tuning

Modern diffusion models, like Stable Diffusion, create highly customized, precise images through fine-tuning. AI can now use several of these techniques to learn many new visual concepts, diverse styles and a variety of subjects without requiring complete retraining. Changing many pre-trained models offers considerably improved flexibility. This improvement eases substantially better personalization of generated content and artistic style refinement, all while optimizing computational resources.

LoRA, DreamBooth and Textual Inversion³³ are specialized techniques that modify how diffusion models, like Stable Diffusion, work. Using these methods, AI can create images containing multiple specific objects, multiple styles, or several concepts, eliminating the need for complete retraining.

DreamBooth improves the model by directly incorporating new subjects into its core, enabling the creation of highly realistic images of specific people or objects in diverse settings. This approach is effective, but requires important computing power and storage due to its wide-ranging base model modifications.

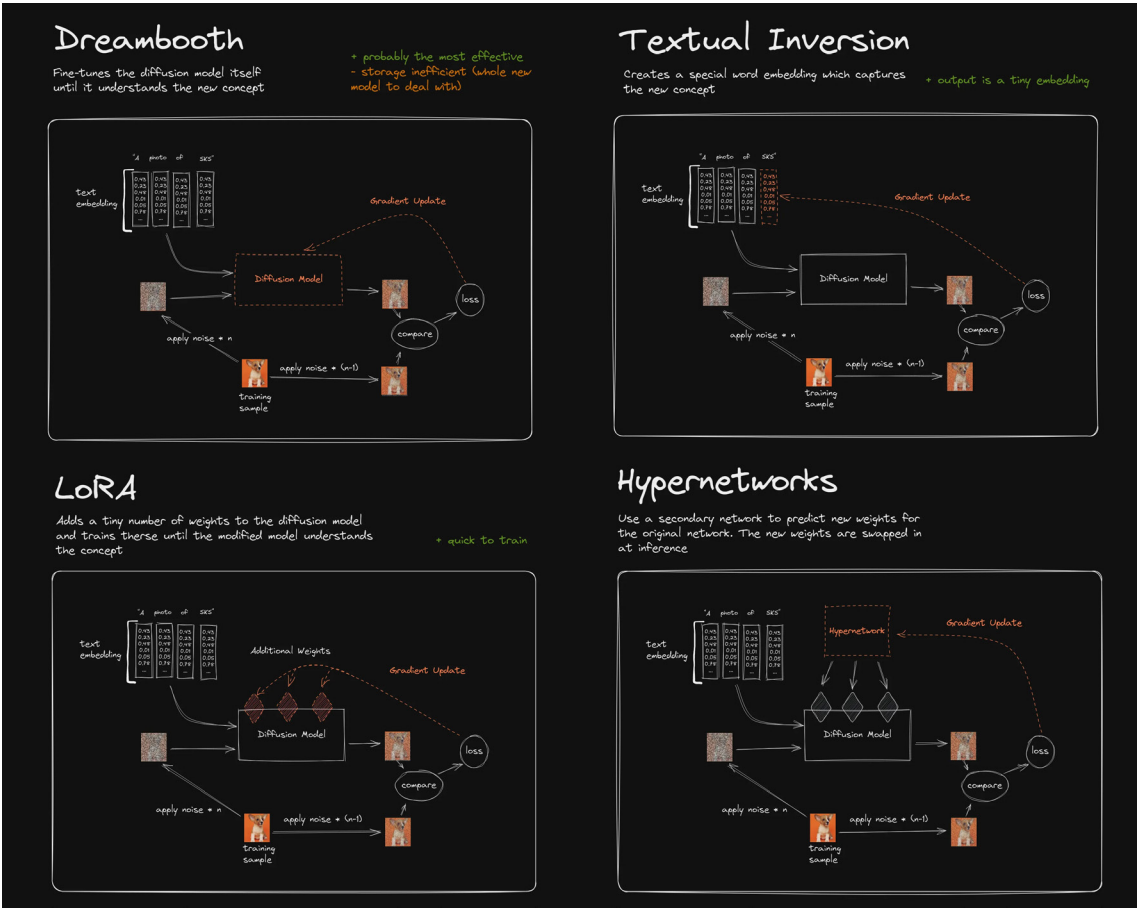
LoRA, or Low-Rank Adaptation, provides a more efficient training method. This method trains precisely a few additional parameters, leaving the original model unaltered. Because of its use of rank decomposition, LoRA exhibits considerably reduced memory requirements while preserving large adaptability, therefore proving ideally suited for environments with limited resources.

In contrast, Textual Inversion encodes new concepts into specific latent embeddings, rather than changing all model weights. Instead of learning a whole subject, it efficiently maps new information to a single special token in the model's text prompt, reducing computational costs.

Hypernetworks offer a new way to tweak diffusion models: a separate network steers the pretrained model without changing its core settings. Unlike DreamBooth, LoRA and Textual Inversion, which modify parts of the model, hypernetworks act as a separate system that dynamically controls the entire generation process.

The choice of technique depends on the application, customization needs and computational constraints.

We chose LoRA for its speed and efficiency. LoRA accelerates model fine-tuning by adding a small number of parameters and this process requires minimal computing resources while leaving the base model's weights unchanged. This solution offers highly efficient customization while remaining resource-efficient.



149 // explanation by Koiboi (source : <https://www.youtube.com/watch?v=dVjMIJsuR5o>)

³³source : <https://medium.com/@antalpha.ai/lor-dreambooth-textual-inversion-in-stable-diffusion-1-5-c6d561542873>

LoRa Training

The possibility of being able to tow one's own models allows one to have a greater influence on artificial intelligence.

Training also allows greater customisation, let's imagine an architectural study with a precise and defined style, we can think of training models on the basis of this style in order to obtain similar concepts and images.

In order to understand how it works, let's look at how it is done and the main parameters that influence its duration and result.

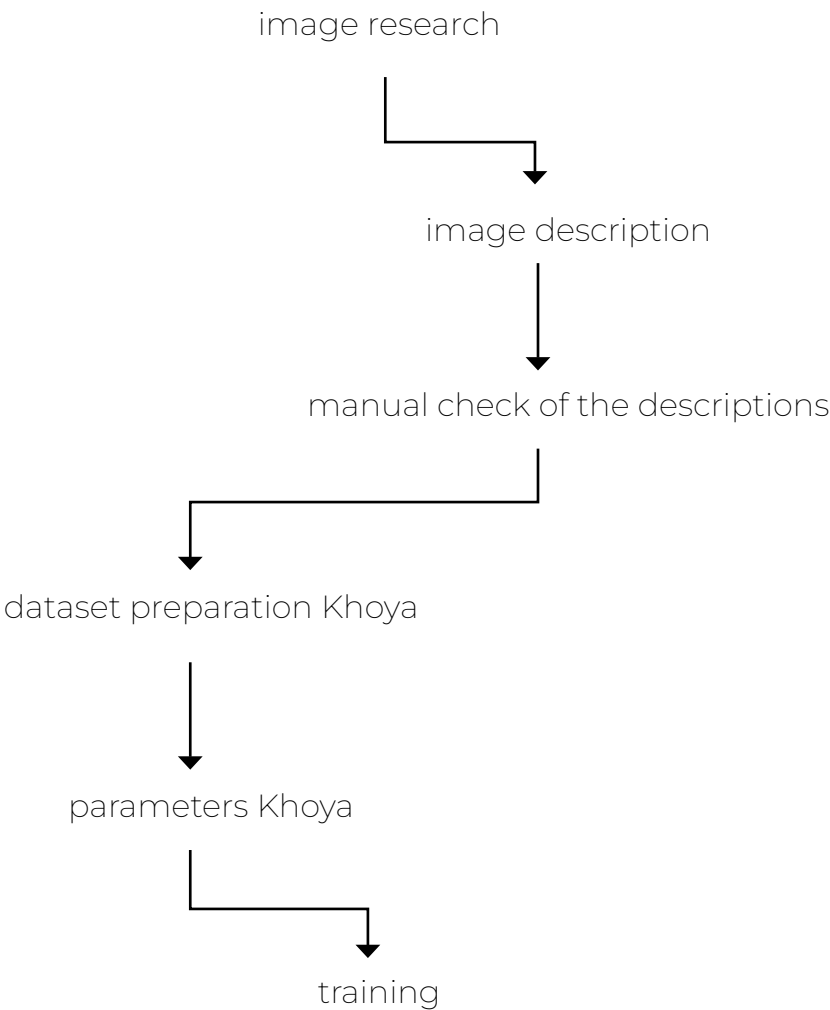
As mentioned earlier, the computational power is very important, we have trained several models on the basis of SDXL with an RTX 3060 with 12 GB of VRAM, with values lower than this it is not possible to train models on the basis of SDXL mainly due to the availability of VRAM.

Furthermore, we chose to train on SDXL because the ControlNet components have not yet been optimised for models such as Flux at the time we are writing this thesis.



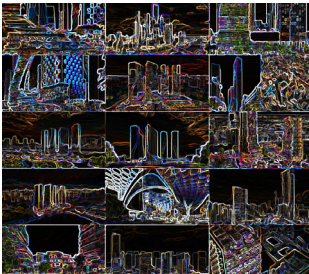
With hardware similar to ours and depending on other parameters that we will see later, the training time for a model based on 15-20 images is around 2 hours.

Next, a conceptual diagram is given to help us understand the various practical phases of a training and how they are connected:



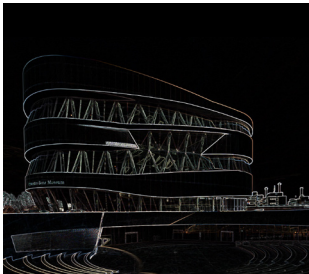
The LoRA's

UNS 01



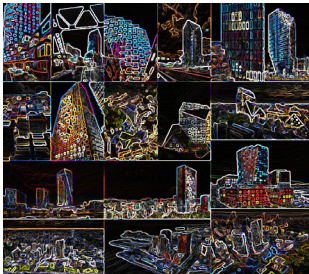
model trained on various projects carried out by UNStudio

UNS 02



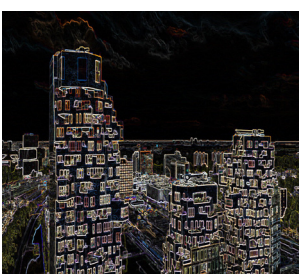
model trained on the 'Mercedes Benz Museum' project by UNStudio

MVRDV 01



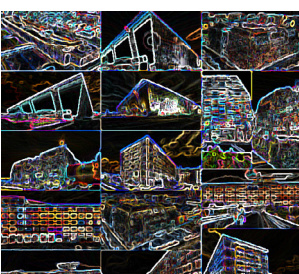
model trained on several projects carried out by MVRDV

MVRDV 02



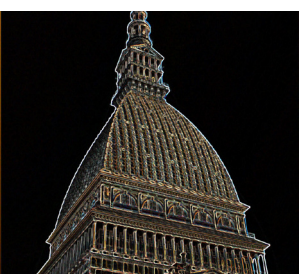
model trained on MVRDV's 'Valley' project

TO 01



model trained on architecture in the Lingotto district of Turin

TO 02



model trained on the Mole Antonelliana in Turin

The Used Dataset

One of the most important things during training is the selection of the images. In this case, we are carrying out a style training that requires a minimum of 15-20 images (obviously, the greater the number of images, the better the model, but this will also affect the training time, increasing it exponentially), while for a character training, for example, only 10 images or less are required.

The images must not be small and must be of good quality. It is not essential to have all the images of the same size, as they are automatically adjusted and processed during training.

Our image sets used for training comprise approximately 15 images.

UNS 02

1/15 - Mercedes-Benz Museum, UNStudio, 2006 (<https://www.archdaily.com/805962/mercedes-benz-museum-unstudio>)



153 // set of image UNS02

UNS 01



152 // set of image UNS01

- 1 - STH BNK by Beulah, UNStudio, 2018 (<https://www.unstudio.com/en/page/11738/sth-bnk-by-beulah>)
- 2 - STH BNK by Beulah, UNStudio, 2018 (<https://www.unstudio.com/en/page/11738/sth-bnk-by-beulah>)
- 3 - STH BNK by Beulah, UNStudio, 2018 (<https://www.unstudio.com/en/page/11738/sth-bnk-by-beulah>)
- 4 - STH BNK by Beulah, UNStudio, 2018 (<https://www.unstudio.com/en/page/11738/sth-bnk-by-beulah>)
- 5 - Hiwell Amber Centre, UNStudio, 2020 (<https://www.unstudio.com/en/page/17259/hiwell-amber-centre>)
- 6 - STH BNK by Beulah, UNStudio, 2018 (<https://www.unstudio.com/en/page/11738/sth-bnk-by-beulah>)
- 7 - STH BNK by Beulah, UNStudio, 2018 (<https://www.unstudio.com/en/page/11738/sth-bnk-by-beulah>)
- 8 - Hiwell Amber Centre, UNStudio, 2020 (<https://www.unstudio.com/en/page/17259/hiwell-amber-centre>)
- 9 - Hiwell Amber Centre, UNStudio, 2020 (<https://www.unstudio.com/en/page/17259/hiwell-amber-centre>)
- 10 - New Heart, UNStudio, 2023 (<https://www.unstudio.com/en/page/17281/new-heart>)
- 11 - Hiwell Amber Centre, UNStudio, 2020 (<https://www.unstudio.com/en/page/17259/hiwell-amber-centre>)
- 12 - Hiwell Amber Centre, UNStudio, 2020 (<https://www.unstudio.com/en/page/17259/hiwell-amber-centre>)
- 13 - K.Wah G72 Mixed-use Development, UNStudio, 2020 (<https://www.unstudio.com/en/page/17056/k.wah-g72-mixed-use-development>)
- 14 - New Heart, UNStudio, 2023 (<https://www.unstudio.com/en/page/17281/new-heart>)
- 15 - K.Wah G72 Mixed-use Development, UNStudio, 2020 (<https://www.unstudio.com/en/page/17056/k.wah-g72-mixed-use-development>)
- 16 - STH BNK by Beulah, UNStudio, 2018 (<https://www.unstudio.com/en/page/11738/sth-bnk-by-beulah>)

MVRDV 01



154 // set of image MVRDV01

- 1 - The Canyon, MVRDV, 2023 (<https://www.mrvd.com/projects/423/the-canyon>)
- 2 - The Canyon, MVRDV, 2023 (<https://www.mrvd.com/projects/423/the-canyon>)
- 3 - The Canyon, MVRDV, 2023 (<https://www.mrvd.com/projects/423/the-canyon>)
- 4 - Baltyk, MVRDV, 2017 (<https://www.mrvd.com/projects/51/baltyk>)
- 5 - Baltyk, MVRDV, 2017 (<https://www.mrvd.com/projects/51/baltyk>)
- 6 - Baltyk, MVRDV, 2017 (<https://www.mrvd.com/projects/51/baltyk>)
- 7 - Downtown One Tirana, MVRDV, 2016 (<https://www.mrvd.com/projects/388/downtown-one-tirana>)
- 8 - Downtown One Tirana, MVRDV, 2016 (<https://www.mrvd.com/projects/388/downtown-one-tirana>)
- 9 - Nieuw Bergen, MVRDV, 2019 (<https://www.mrvd.com/projects/290/nieuw-bergen>)
- 10 - Nieuw Bergen, MVRDV, 2019 (<https://www.mrvd.com/projects/290/nieuw-bergen>)
- 11 - Nieuw Bergen, MVRDV, 2019 (<https://www.mrvd.com/projects/290/nieuw-bergen>)
- 12 - Chengdu Tianfu Software Park, MVRDV, 2022 (<https://www.mrvd.com/projects/922/chengdu-tianfu-software-park>)
- 13 - Downtown One Tirana, MVRDV, 2016 (<https://www.mrvd.com/projects/388/downtown-one-tirana>)
- 14 - The Canyon, MVRDV, 2023 (<https://www.mrvd.com/projects/423/the-canyon>)
- 15 - Chengdu Tianfu Software Park, MVRDV, 2022 (<https://www.mrvd.com/projects/922/chengdu-tianfu-software-park>)
- 16 - The Canyon, MVRDV, 2023 (<https://www.mrvd.com/projects/423/the-canyon>)
- 17 - Chengdu Tianfu Software Park, MVRDV, 2022 (<https://www.mrvd.com/projects/922/chengdu-tianfu-software-park>)



155 // set of image MVRDV02

MVRDV 02

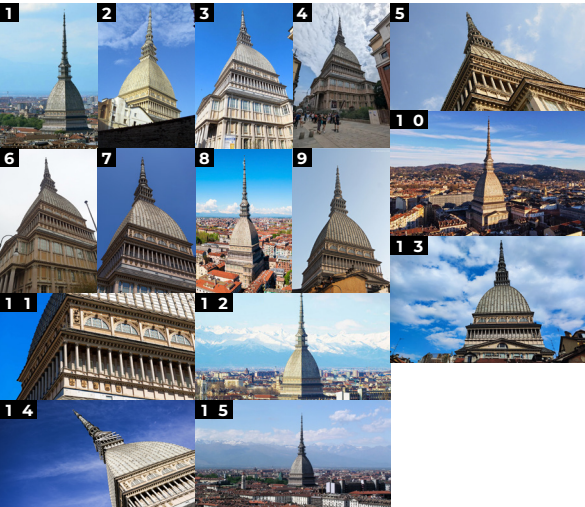
1/17 - Valley, MVRDV, 2022 (<https://www.mvrdv.com/projects/233/valley>)



157 // set of image TO02

TO 02

- 1 - Lingotto Shopping Centre (<https://www.teknoring.com/news/progettazione/doppia-vita-lingotto-torino/>)
- 2 - Oval Olympic Arena, Zoppini Associati/HOK Sport, 2005 (<https://www.studiocattivelli.it/realizzazioni/pala-oval-torino/>)
- 3 - Green Pea, ACC/Negozio Blu, 2020 (<https://www.infobuildenergia.it/torino-il-green-pea-di-farnetti/>)
- 4 - Oval Olympic Arena, Zoppini Associati/HOK Sport, 2005 (<https://www.studiocattivelli.it/realizzazioni/pala-oval-torino/>)
- 5 - Oval Olympic Arena, Zoppini Associati/HOK Sport, 2005 (<https://www.studiocattivelli.it/realizzazioni/pala-oval-torino/>)
- 6 - Green Pea, ACC/Negozio Blu, 2020 (<https://www.pambianconews.com/2020/12/09/green-pea-apre-a-torino-il-retail-park-sostenibile-306342/>)
- 7 - Green Pea, ACC/Negozio Blu, 2020 (<https://beauty2business.com/2022/12/02/bilancio-del-green-pea-di-torino/>)
- 8 - Green Pea, ACC/Negozio Blu, 2020 (<https://www.negoziobluarchitetti.it/it/in-primo-piano/21-torino-inaugura-green-pea.html>)
- 9 - Lingotto Shopping Centre (<https://www.expedia.it/Torino-Lingotto.dv6055016>)
- 10 - Lingotto Shopping Centre (<https://www.expedia.it/Torino-Lingotto.dv6055016>)
- 11 - Green Pea, ACC/Negozio Blu, 2020 (<https://www.turismodelgusto.com/tuttofood/istoranti-a-torino-green-pea/>)
- 12 - Lingotto Shopping Centre (<https://www.expedia.it/Torino-Lingotto.dv6055016>)
- 13 - Lingotto Shopping Centre (<https://www.expedia.it/Torino-Lingotto.dv6055016>)
- 14 - Lingotto Shopping Centre (<https://www.expedia.it/Torino-Lingotto.dv6055016>)
- 15 - Lingotto Shopping Centre (<https://www.expedia.it/Torino-Lingotto.dv6055016>)
- 16 - Oval Olympic Arena, Zoppini Associati/HOK Sport, 2005 (<https://www.expedia.it/Torino-Lingotto.dv6055016>)
- 17 - National Automobile Museum, Amedeo Albertini, 1960 (<https://www.museotorino.it/view/5cc514a27847e4a57a6f7910736517522>)
- 18 - National Automobile Museum, Amedeo Albertini, 1960 (<https://www.museotorino.it/view/5cc514a27847e4a57a6f7910736517522>)



156 // set of image TO01

TO 01

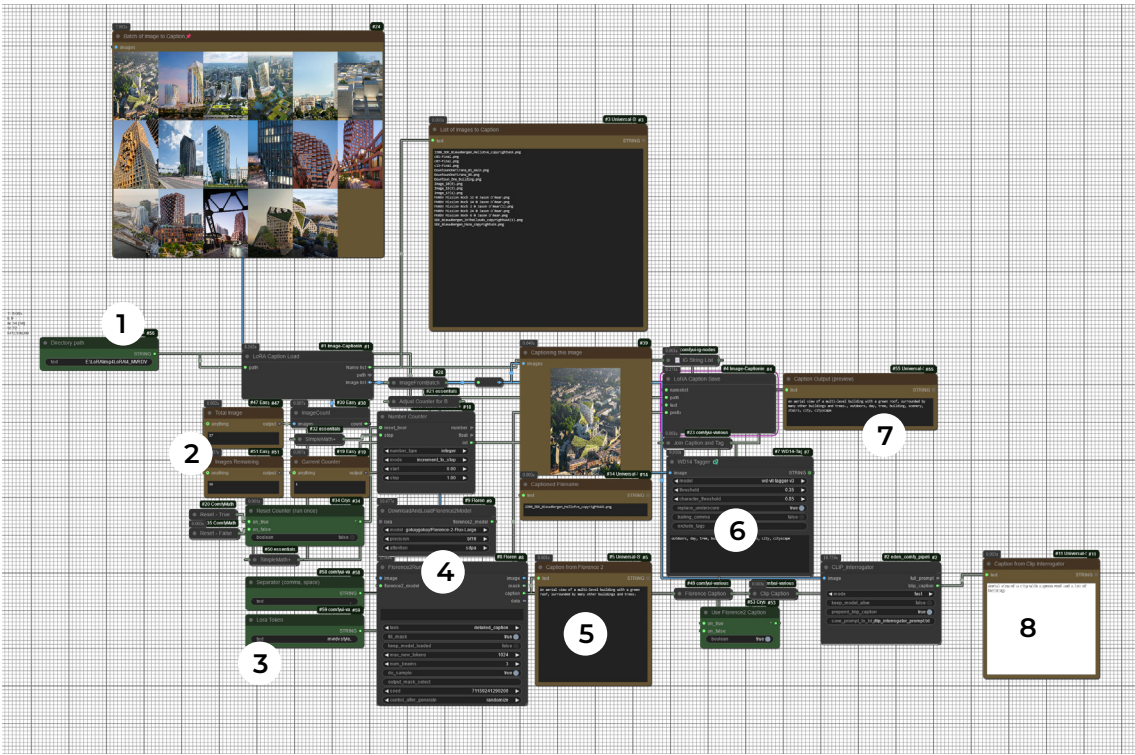
- 1 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.museocinema.it/itmuseo-e-fondazione-ma-prola/mole-antonelliana>)
- 2 - Mole Antonelliana, Alessandro Antonelli, 1863 (https://it.m.wikipedia.org/wiki/Mole_Antonelliana)
- 3 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.museocinema.it/itmuseo-e-fondazione-ma-prola/mole-antonelliana>)
- 4 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.museocinema.it/itmuseo-e-fondazione-ma-prola/mole-antonelliana>)
- 5 - Mole Antonelliana, Alessandro Antonelli, 1863 (https://commons.wikimedia.org/wiki/File:Mole_antonelliana_2006.jpg)
- 6 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.villedecor.com/it/architettura/41461524/mole-antonelliana-il-simbolo-di-torino-custode-della-settima-arte/>)
- 7 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.villedecor.com/it/architettura/41461524/mole-antonelliana-il-simbolo-di-torino-custode-della-settima-arte/>)
- 8 - Mole Antonelliana, Alessandro Antonelli, 1863 (https://it.m.wikipedia.org/wiki/Mole_Antonelliana)
- 9 - Mole Antonelliana, Alessandro Antonelli, 1863 (https://it.m.wikipedia.org/wiki/Mole_Antonelliana)
- 10 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.guidatorino.com/la-mole-antonelliana/>)
- 11 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.museotorino.it/view/53b1a0b1906c64a22b777cfab5df1d54d>)
- 12 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.museotorino.it/view/53b1a0b1906c64a22b777cfab5df1d54d>)
- 13 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.guidatorino.com/la-mole-antonelliana/>)
- 14 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://mole24.it/2021/08/30/mole-antonelliana-monumento-simbolo-di-torino/>)
- 15 - Mole Antonelliana, Alessandro Antonelli, 1863 (<https://www.museotorino.it/view/53b1a0b1906c64a22b777cfab5df1d54d>)

Image Description

After all the images that are needed for training have been chosen, there is the next step of describing the images.

This is one of the most important and time-consuming parts of the process, as each image must be described in a certain level of detail - in practice, the more detail and description, the better the output from the trained model.

To save time, we automated this process by creating a workflow in ComfyUI that automatically described the images within a chosen folder by creating text notes associated with the images.



158 // comfyUI workflow

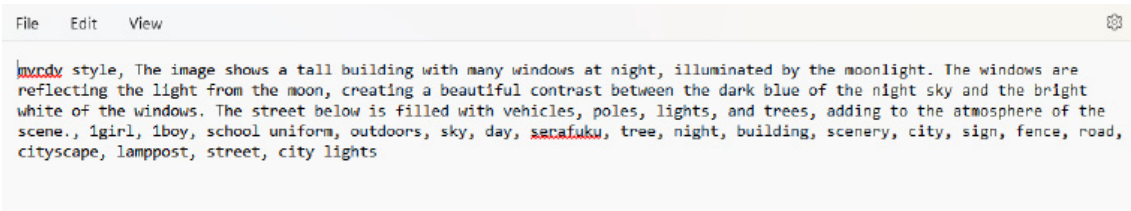
This workflow has allowed us to save an enormous amount of time, as mentioned above we have to give as input our folder with the images (1), the nodes (2) allow us to see how many images are present and how many still have to be processed while in the node (3), the LoRA token we can insert parts of text that will be added automatically in the description of each image.

Node (4) is the heart of the process, from here we choose the model we use to process the images and describe them, in this case we have used the Florence 2 Flux Large model, but there are several models available that can be tested.

In part (5) we display the output of Florence2. This workflow sees the output of two different processes being merged to describe an image, the second process is the WD14 Tagger (6) and in node (7) we see the merged outputs, thus Florence2 + WD14 Tagger. Another tagger process is also provided, which is the CLIP Interrogator (8) but we preferred to use Florence2 as we found it more precise and descriptive.

Once the process has started, text notes are created in the folder with the name of the processed image.

Before moving on to the next step these text notes must be checked one by one, the AI can add words that are not appropriate for the image (in our case in some text notes we found the word 'tokyo' which is not part of the image).

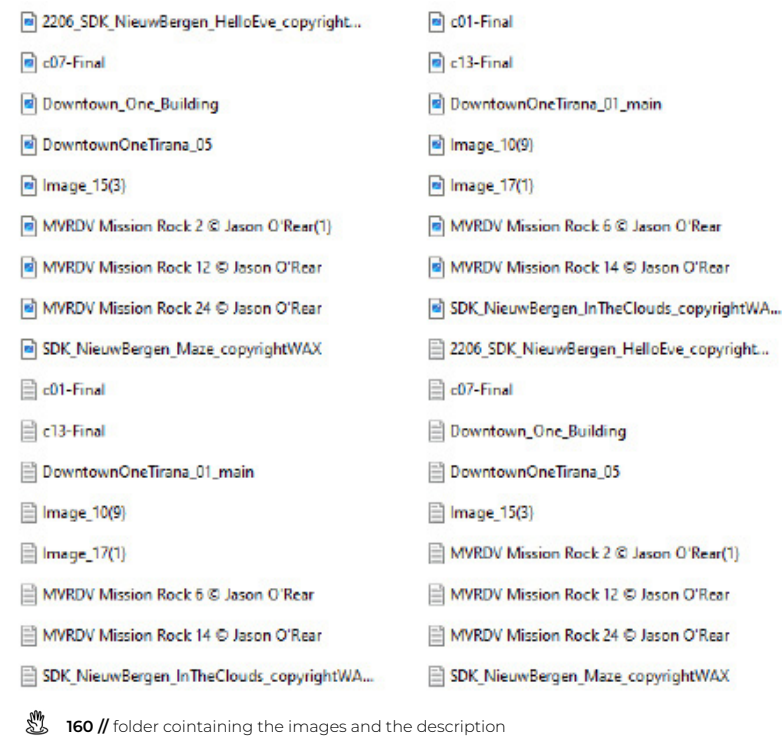


159 // example of the text description output

Kohya - Dataset Preparation

Most of the work is done once the images have been chosen and described; it is a fairly time-consuming process, but it must be done well and carefully otherwise the results will not be as desired.

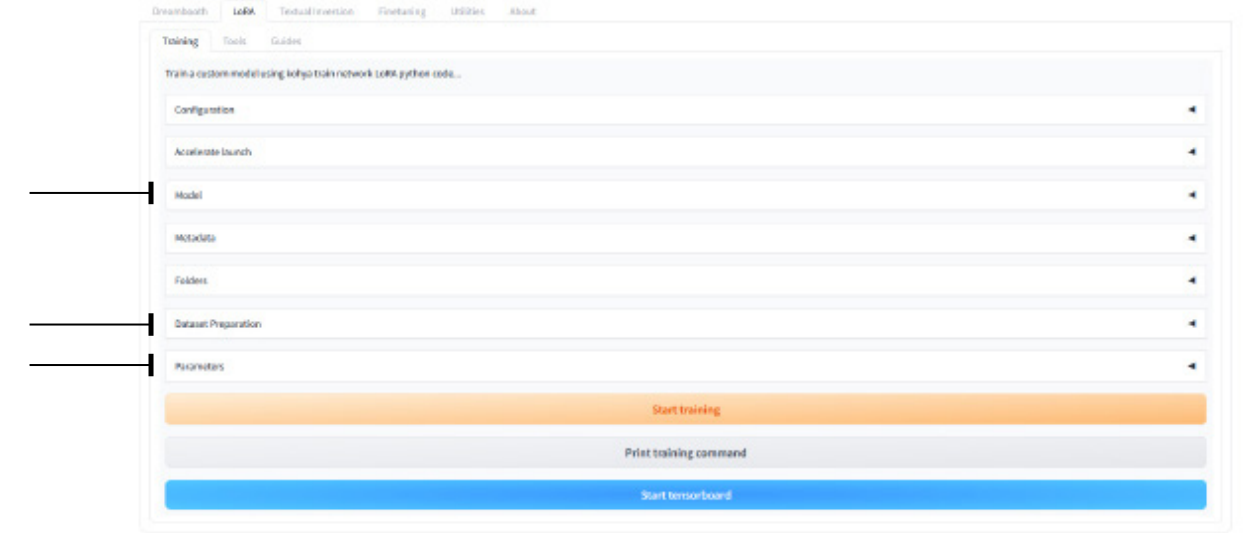
Once this part is finished, we will end up with a folder containing our images and text files with descriptions.



160 // folder containing the images and the description

To proceed with the training, we will rely on an external process, Kohya³⁴.

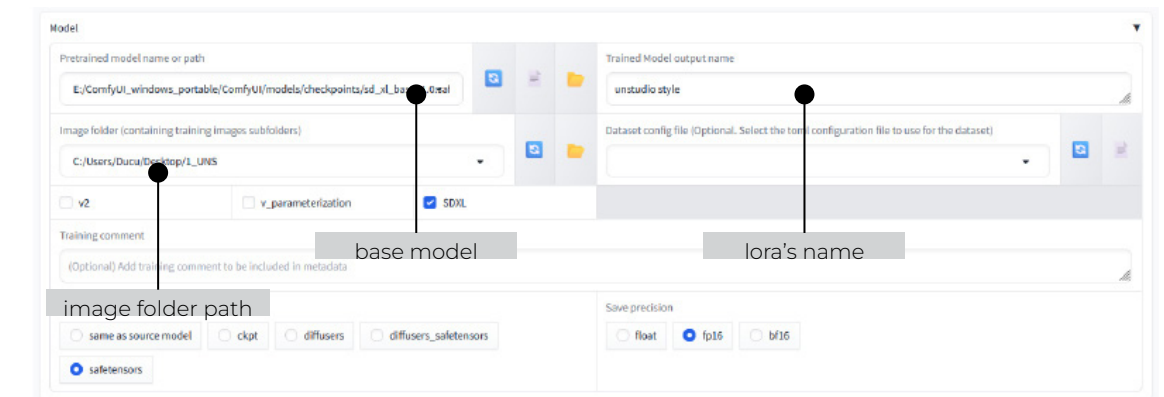
³⁴repository containing training, generation and utility scripts for Stable Diffusion



161 // Kohya interface

There are a number of parameters and sections that can be filled in and modified, our aim is not to get too much into the technicalities of a process but still it is quite important to understand what influences the training process.

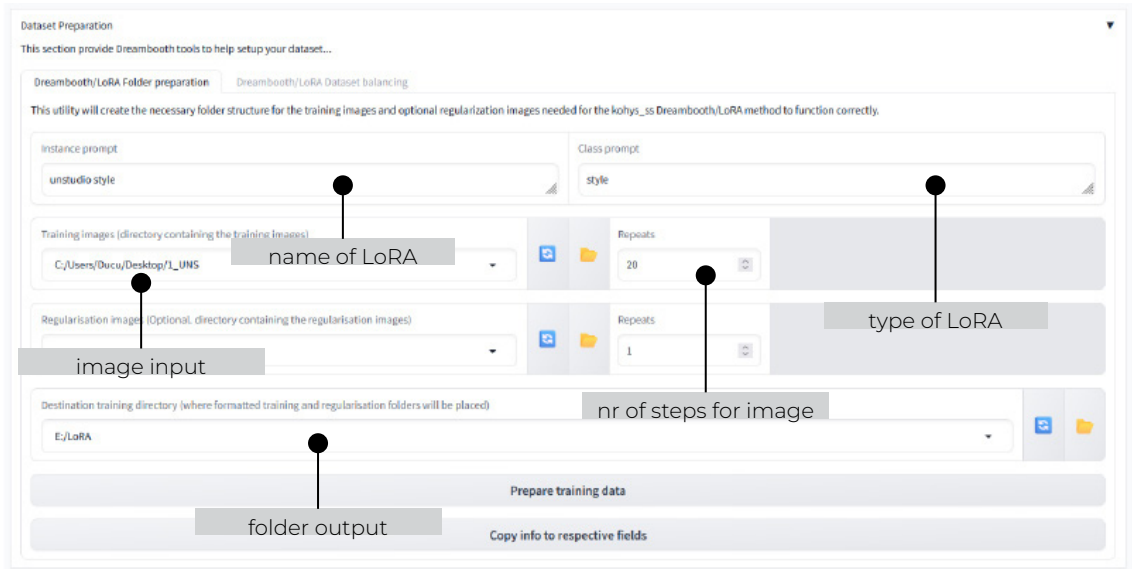
The model section, as the name suggests, allows us to choose which basic model to train our LoRA on, name it and select the folder path with our images and descriptions.



162 // model parameters input

Kohya - Parameters

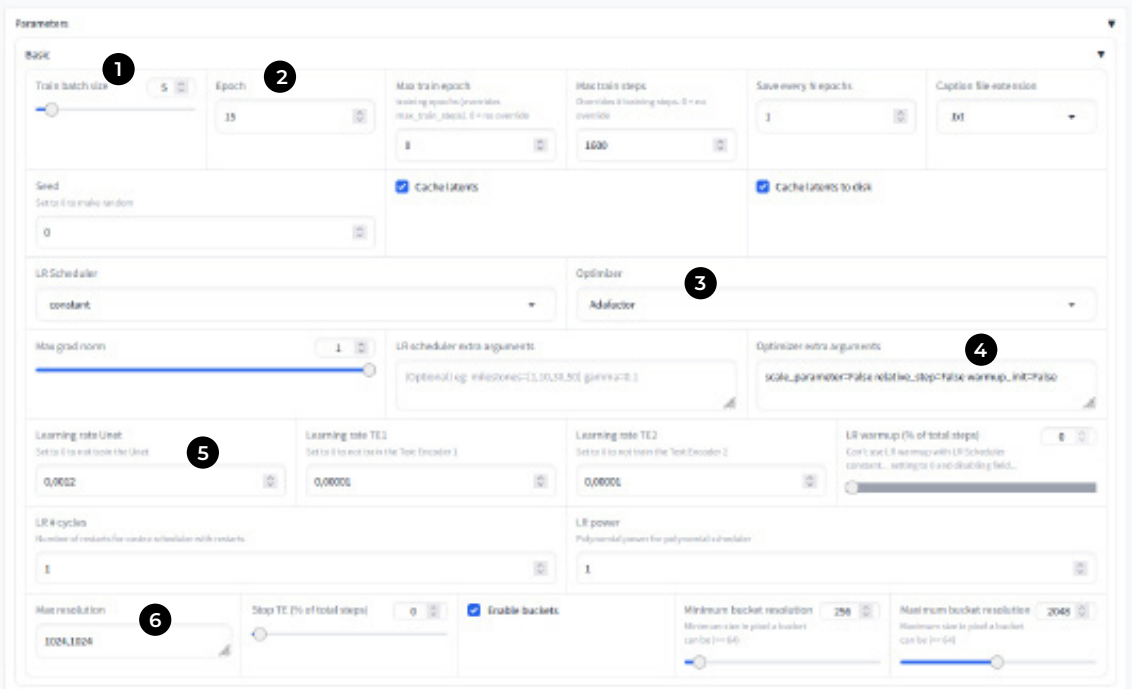
The Dataset Preparation part can be defined as the logistical input part. Here we define, again, the path of the folder with our images, some nomenclature and the path of where our LoRA will be created, and we can also specify how many times our images will be processed, this value affecting the time required for training.



163 // dataset preparation input

Next comes the parameter management part, the data that can be changed are indeed many, but we will only look at the main ones:

- train batch size 1
- epoch 2
- optimizer 3
- optimizer extra arguments 4
- learning rate 5
- max resolution 6



164 // parameters input

Train Batch Size (1) refers to the number of samples per iteration, in simple terms how many images are processed at the same time, we have kept it at 1 due to our computing power, if you have more computing power this value is preferable to keep it high as it decreases the training time.

The epoch (2) corresponds to a complete cycle in which all input data is processed at least once.
If a dataset comprises 100 samples with a batch size of 5, $100/5 = 50$ iterations will be required to complete an epoch.
We can also define an epoch as a checkpoint, at the end of an epoch a checkpoint of the model is saved which can be used, we will see the differences in a later study, and for each epoch generated during training the model gradually improves by learning the patterns of the data better.

The optimiser (3) is the component that allows the model to ‘learn’ data by adapting the model’s weights to minimise data loss (loss function). We also used extra optimiser arguments (4) to customise the behaviour of the optimiser.

Another fundamental parameter is the learning rate (5) which, as the word itself says, determines how fast the model learns:

- In general, a learning rate of $1e-4$ / $1e-5$ is used for general tasks and $1e-3$ for models with smaller datasets.

The last parameter we see is the max resolution (6), which refers to the maximum value for image resolution during training.

During the training process, there are various parameters to check, but as mentioned above, we have only analysed the main ones to get a rough idea of what influences the output and especially its timing.

On the side are all the parameters used for these trainings.

training parameters

```
save precision : bf16
repeats image : 20
train batch size : 1
epoch : 15
max train steps : 1600
cache latents : true
cache latents disk : true
LR Scheduler : constant
optimizer : Adafactor
optimizer extra arguments : scale_parameter=False relative_
step=False warmup_init=False
learning rate : 0.0012
max resolution : 1024,1024
enable buckets : true
minimum bucket resolution : 256
maximum bucket resolution : 2048
clip skip : 1
max token lenght : 75
bucket resolution steps : 64
sample sampler : euler_a
```


LoRA's Iterations

We tested the new LoRA's created with the same parameters to understand the substantial differences and how the latter affects the final output of the generation.

prompt : view of an house with a pool in front



model : JuggernautXL
seed : 2025
steps : 40
cfg : 6.0
sampler : dpmpp_2m_sde
scheduler : karras



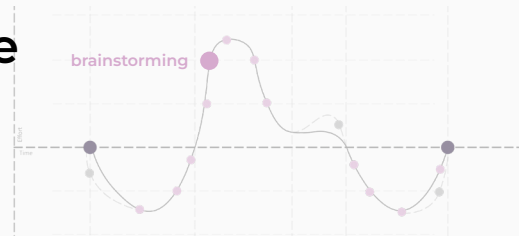
The study above shows how a LoRA can completely change the output, and the same parameters have always been used, and we have simply changed the weight the LoRA has in order to understand the substantial differences, but remember that in ComfyUI we can control multiple parameters that alter the final result.

One can also combine the various LoRAs to obtain even more different outputs, as we will see later.

The trained LoRAs are consistent with the inputs used in the training, similarities can be seen in the various images created.

AI Process - The Dialogue

Customisation



In the previous images, despite the fact that we sought through the prompt, even giving it more power, a somewhat Torinese touch, we did not get the desired results, and even as far as the modern look is concerned, there was not much difference between the creation, of course you can change the prompt endlessly and generate completely different results, adding or not adding detail to our prompt.

We realised that it is difficult to obtain a Turin recall, so from now on we have integrated the models trained by us into our workflows, for now we have tested them by giving them more context and directing them towards the type of architecture we are looking for.

The models tested are the six introduced previously, they were tested by integrating them into the same workflow used previously with different weights to adjust their influence and consequently modify the output.

On the right-hand side we have in the first two rows the models dedicated to UNStudio architectures, then in rows three and four the models dedicated to MVRDV architectures, and in the last two the models dedicated to Turin architectures.

The customisation component that we are able to achieve through the LoRAs is extremely important, in this way we are able firstly to search for a precise style, but secondly we make the artificial intelligence personal. By not using our personal models, in a certain sense anyone could arrive at a similar result, but by doing so we make the architecture extremely personal.



Advanced Customisation

In order to contextualise even more and to have a result that we can define as real and reliable, we have to work directly on the original building.

We have chosen some images and extrapolated masks on which we are going to apply the workflow integrated with the models we have trained.

In addition, for some tests several LoRAs were used at the same time to attempt a blending between the models.

Abbiamo scelto due immagini reali e per ognuno abbiamo testato i diversi set di LoRA : 1 (UNS), 2 (MVRDV), 3 (TO).



167 // view of Grattacielo della Regione Piemonte



168 // mask input

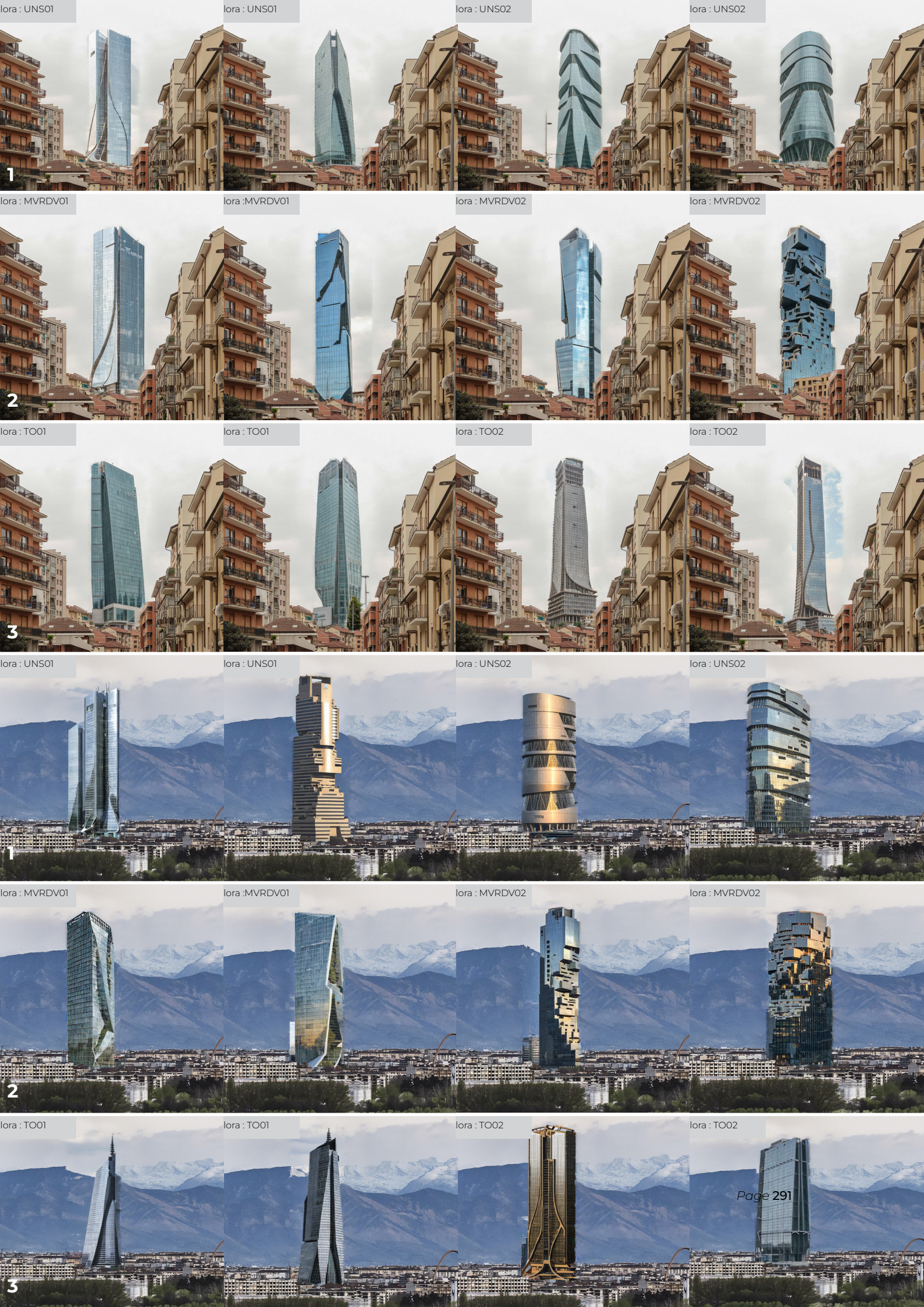
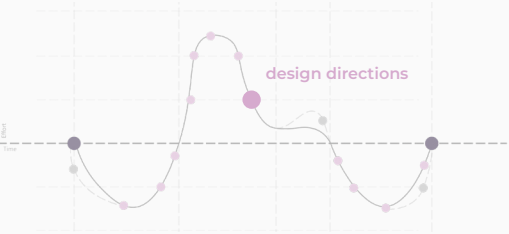


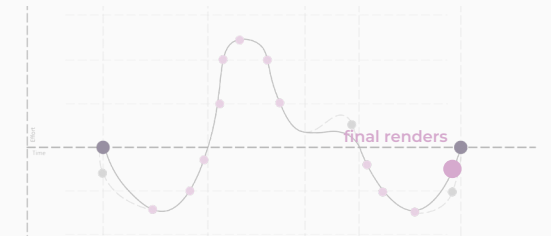
169 // view of Grattacielo della Regione Piemonte



170 // mask input

171 // on the right : different images created with AI





As in the first project, here too we can proceed through the same workflow to retouch the images in order to correct them where they are inconsistent.

In this way we go directly to the Final Renders phase, in case there is no need for this, the first outputs may already be fine.



In this process we have completed all the steps we have seen so far with the sole use of artificial intelligence, but as we shall see later on, this is not good enough here as we need the production part.

Furthermore, for a more complex architecture, it is more difficult to create different views of the same composition.



Generating the 3D Model

We leveraged artificial intelligence to generate a 3D model, subsequently refining its detail and accuracy with Rhino's capabilities. To speed up computer processing, we first used the ReduceMesh command to simplify the mesh geometry, from more than 20000 meshes in the 1st 3d model to 100 in the 2th, reducing the number of faces. We converted the mesh to NURBS in the 3th model for easier manipulation. Rhino's NURBS surfaces then allowed us to precisely and flexibly manage complex details, producing smoother, more editable geometry.

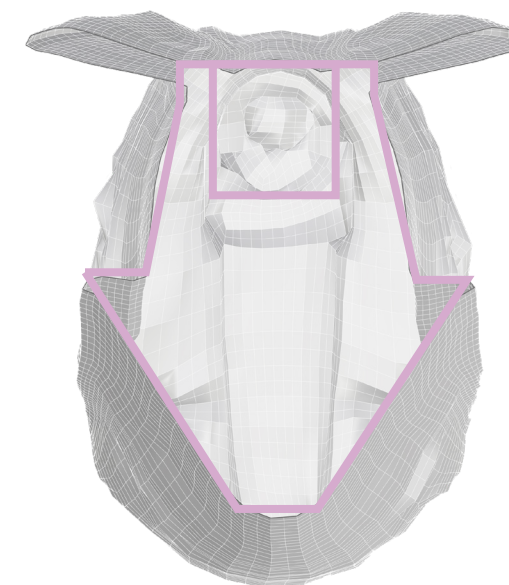
Once we processed the model, its natural design weaknesses became glaringly obvious, considerably affecting the ensuing analysis. The model has exceptionally important potential. While imperfect, the plans suggest a central core and show some spatial organization.



172 // 3d view of the generated model and his derivatives



173 // plan view of the generated model and his derivatives



174 // plan suggestion from AI 3d model

The model remains imprecise and despite large improvements in our understanding and large efforts to refine it, meaningful progress continues to be elusive. Nevertheless, as previously noted, the AI's attempts to replicate reality, even in its initial planning phases, are of large interest, especially its generation of an important sense of spatiality. The model generates a plan outline and offers several helpful suggestions. This provides an useful foundation for future development, despite its inaccuracies.

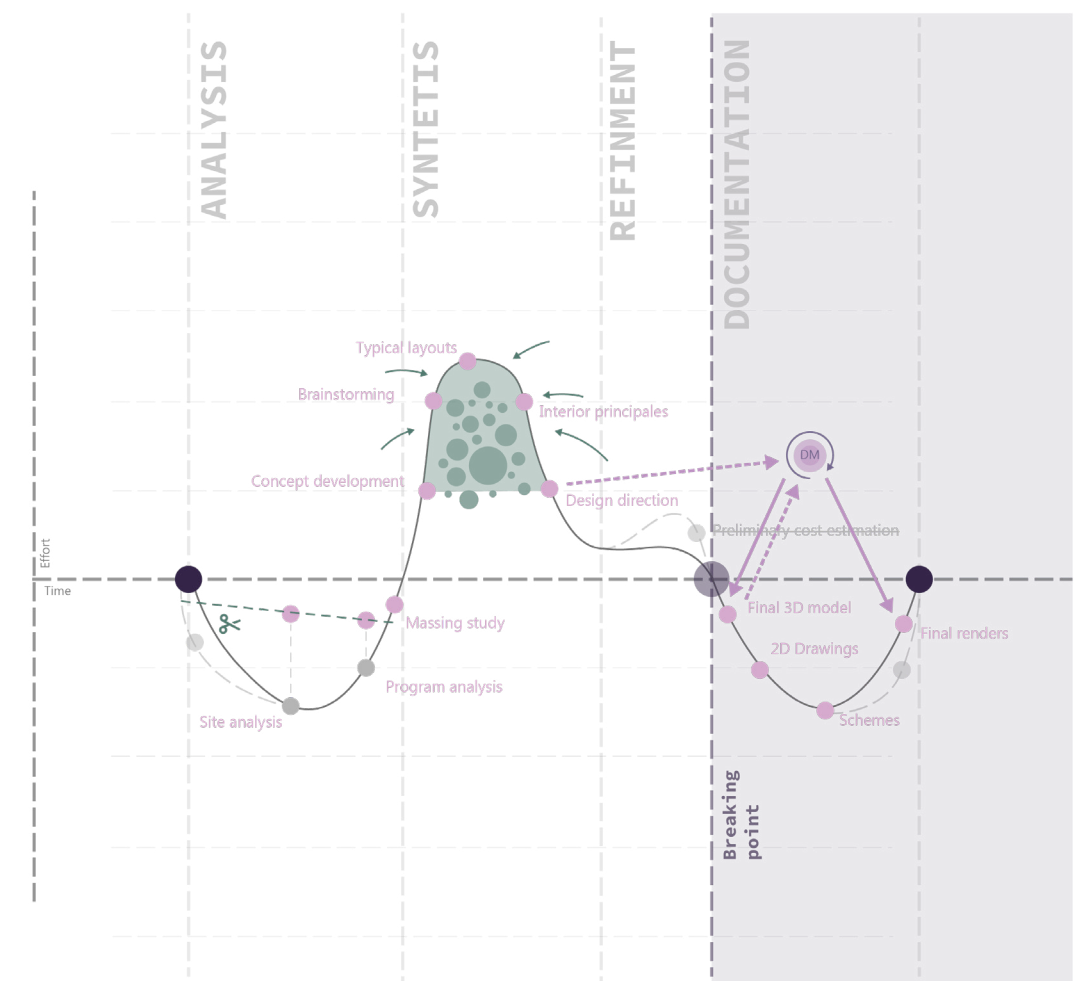
The Dialogue Conclusion

In the diagram on the right, we see how the dialogue process has changed the traditional process.

As we have already seen, the first part of the analysis and concept could lead to a creative increase in the proposed solutions and a shortening of timeframes, but we see how following a completely AI-driven process we arrive at a breaking point that we cannot bridge using artificial intelligence alone, this point is exactly before the production phase, at the time we are writing our thesis we are not yet able to generate drawings and technical drawings that meet the precision and requirements required using AI alone, we need to enter into the traditional, manual process of modelling the chosen design in order to be able to continue into the next stages of the project.

In this case we are only able to fulfil one step of the production phase, which is the render phase, the AI in this case allows almost instantaneous delivery of the required output.

The Dialogue process has limits that are reached at its breaking point, we must point out that there are tools that given different inputs can proceed to the automatic generation of different layouts and project plans, but it is not our aim in this thesis to go into this subject in depth as we focus mainly on the diffusion models.



AI Process - The Inversion

Massing Study

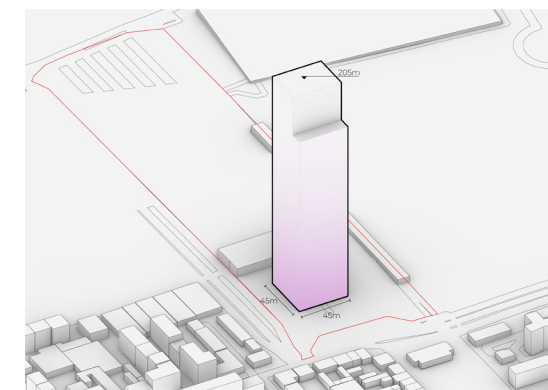
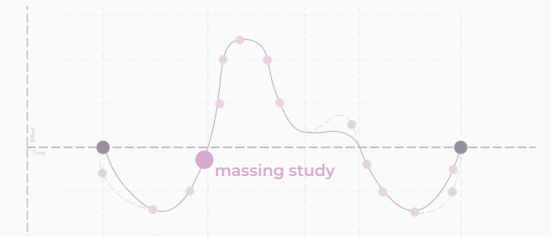
To approach the project, we considered 4 different massings with the same volume of 400 000 m², i.e. the original volume of the building.

The analysed massings change in composition and building size in order to explore different design solutions.

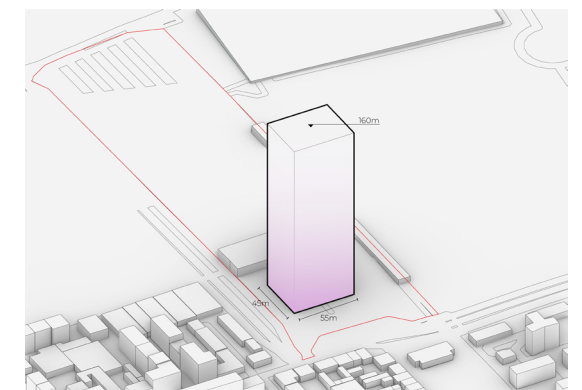
In the first case we kept the massing the same as the original in order to have a true comparison with the real building.

In the second case we have lowered the building to a height of 160 metres, increasing its width to reach the required volume, the height has been lowered to that point so that it does not exceed that of the Mole Antonelliana of 167 metres, being the Turin icon we wanted to explore this option which does not exceed it in terms of tonnage.

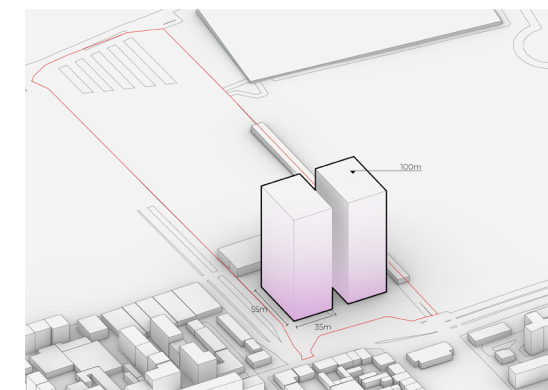
The third and fourth options see a decomposition of the volume to further break up the size of the original building, which has a major impact on Turin's skyline.



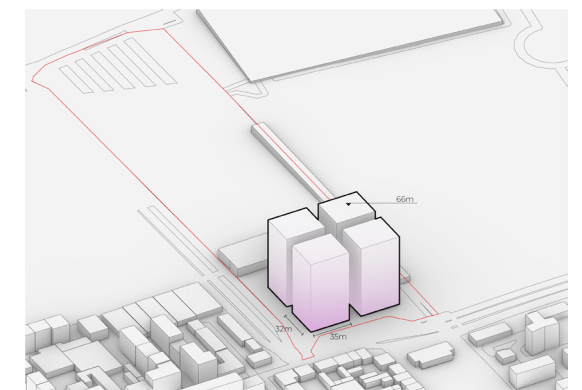
Actual Design (205 meters height)



Option 01 (160 meters height)



Option 02 (100 meters height)



Option 04 (66 meters height)

Step-story diagram

After the above experience developed which featured the fail of the process “the dialogue”, we started to introduce a more classic, canonical design approach.

The project is to be located in an awesomely large site, with a number of services now underground concentrating the mass of the program in a single monolithic element. Being to be a loud statement of identity, creating an important landmark, the monolithic dimension is not appreciated at the level it had hoped for.

The Turin skyline is not characterised by the presence of buildings of such heights, especially if it is among the tallest in Italy. Additionally, the façade itself contributes to the building’s presence. Moreover, hiding a large portion of the programme underground leads to accentuating this great majesty of the building.

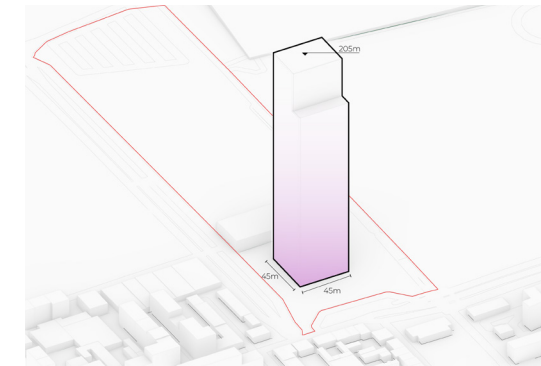
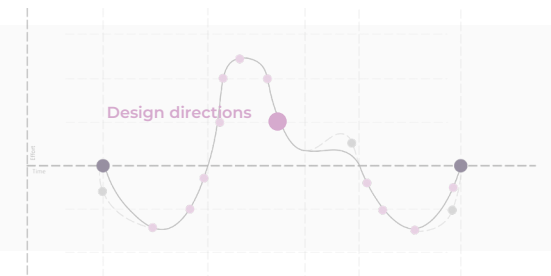
Therefore this leads to clash with the background of the very eclectic neighbourhoods.

In the previous process the aim was to redesign just the façade alone, taking advantage of the residual space built in the main façade characterised by diagonal elements and a space running through the whole building. It was therefore conceivable to redesign the façade, having in mind a strategy that was purely focused on breaking up the volume itself by dividing the façade, using different materialities and also playing with different levels of transparency.

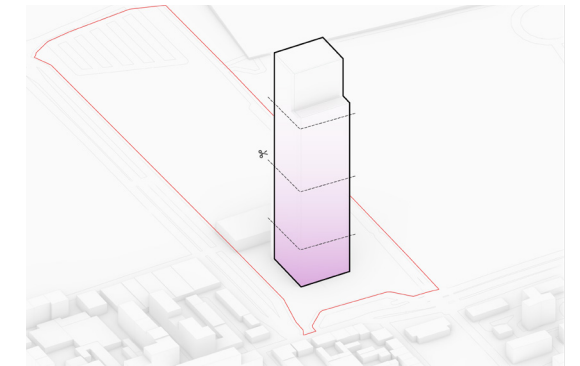
Instead, it was decided to develop this second process of wanting to challenge in a more concrete way the use of artificial intelligence and the help it can bring to the process itself.

Therefore, We started a whole redesign of the site based on strategic development of a cluster of volumes. Having identified, thanks to the study of the programme, that different plans were articulated in a thematic way, we tried to imagine clusters of professions that could face the same issues. We made up 4 different clusters as explained in the following pages.

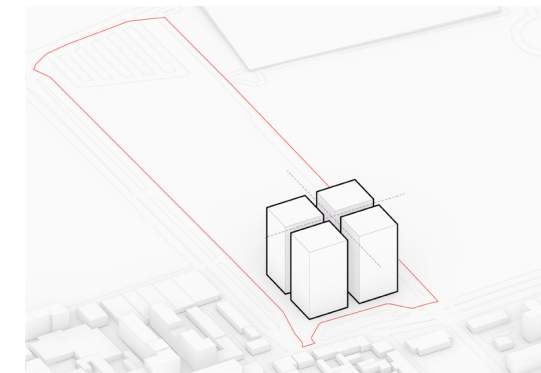
Working with the basic premises and adding concepts from one classical design method, let’s try to insert AI tools.



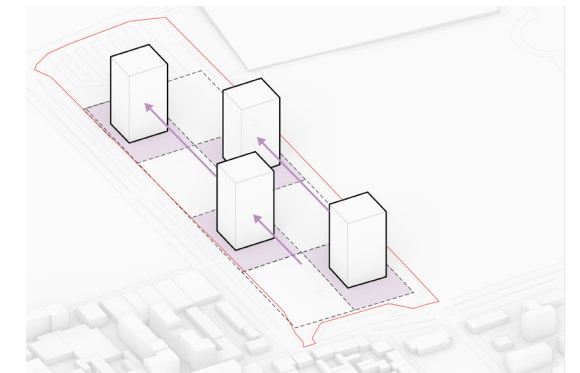
STEP 00 : the original tower



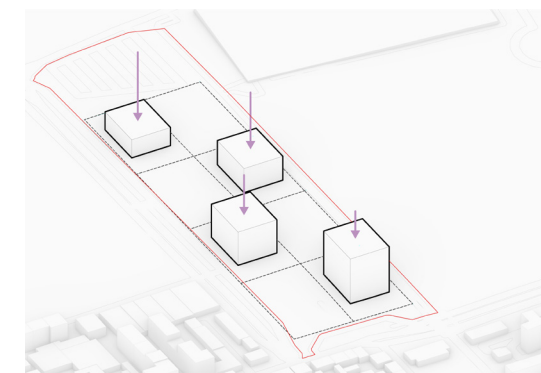
STEP 01 : Braking the volume



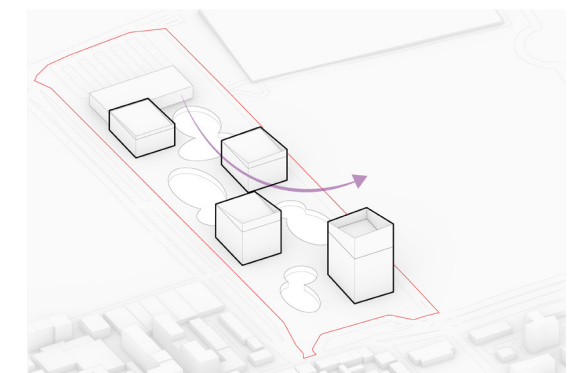
STEP 03 : Diving the program



STEP 04 : populating the site



STEP 05 : adapting the program



STEP 06 : shape identity

The Final Reference



178 // different images created with AI



179 // starting image inversion project 2 process 2

Genereting the 3D Model

In previous chapters we tested a platform that allows you to produce a 3D model from a 2D image, the results were very promising but lacked precision and the possibility of manipulation. Despite this, the image input in the first input test was a rather simple and clear image with non-complex geometries.

We tested this tool on the main reference image of this project. As can be seen below, from a more compositionally complex image it was difficult to get a model, as it recognised an element in the foreground and processed that. After this first test, we resized the image and cropped only the building, i.e. the part most affected.



180 // input 01 for 3d and his output



With the resized image, the process does indeed work, but the generated 3D model, although reflecting some characteristics of the image, is in principle not usable, as it lacks precision and above all, as already mentioned, possible manipulation and modification.



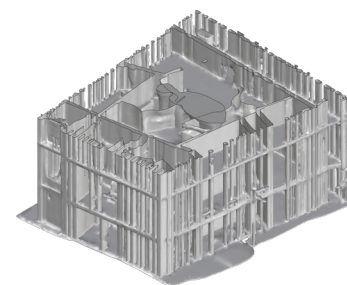
181 // input 02 for 3d

182 // output from image input 02

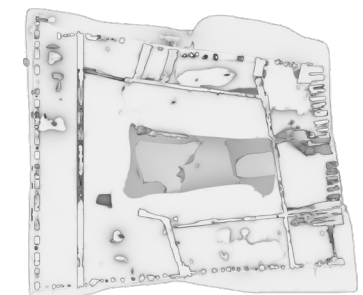
183 // textured output from image input 02

It is interesting to see, in the sectioned model, how the AI somehow recognises the type of building and tries to recreate some internal partitions, despite the fact that there are very few references in the image to an internal distribution.

In spite of this, research and development continues, and it could be that in a few months or years' time it will be possible to generate 3D models with far greater precision.

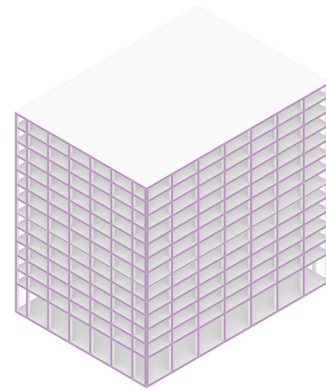


184 // sectioned 3d model from input image 02



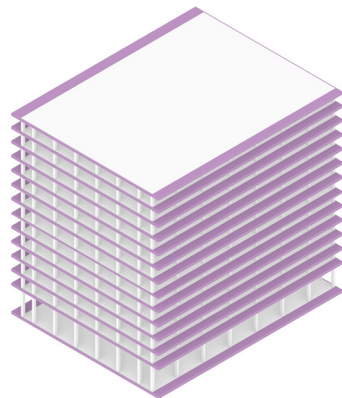
185 // plan of the 3d model

The Production



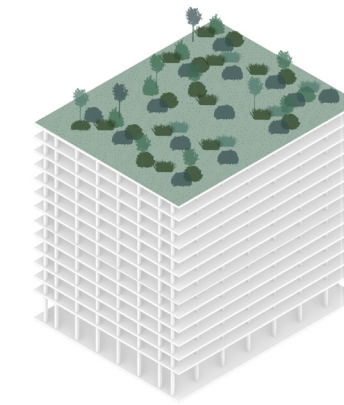
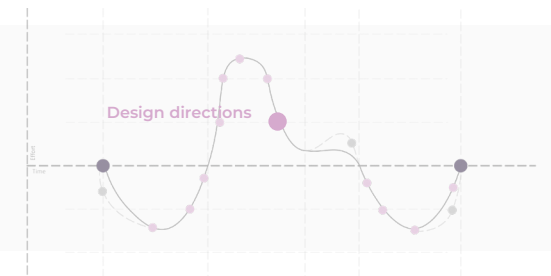
1. Contextual reference :

At the basis of the design proposal we have a strong rigour given by a marked structure dividing the façade with a very rational grid. This base is very important to maintain a strong link with the context by directly recalling the façade of the Lingotto.



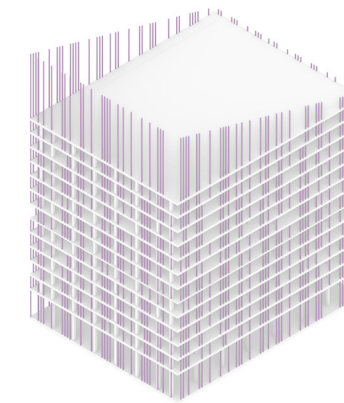
2. Residential feeling :

The programme is spread over the entirety of the large available site, leading to the creation of a true low-density neighbourhood. This is accentuated by the presence of balconies on the façades, making the basic grid itself lighter and giving it a residential feel. This further entices the public, as well as civil servants, to make more use of the landscape.



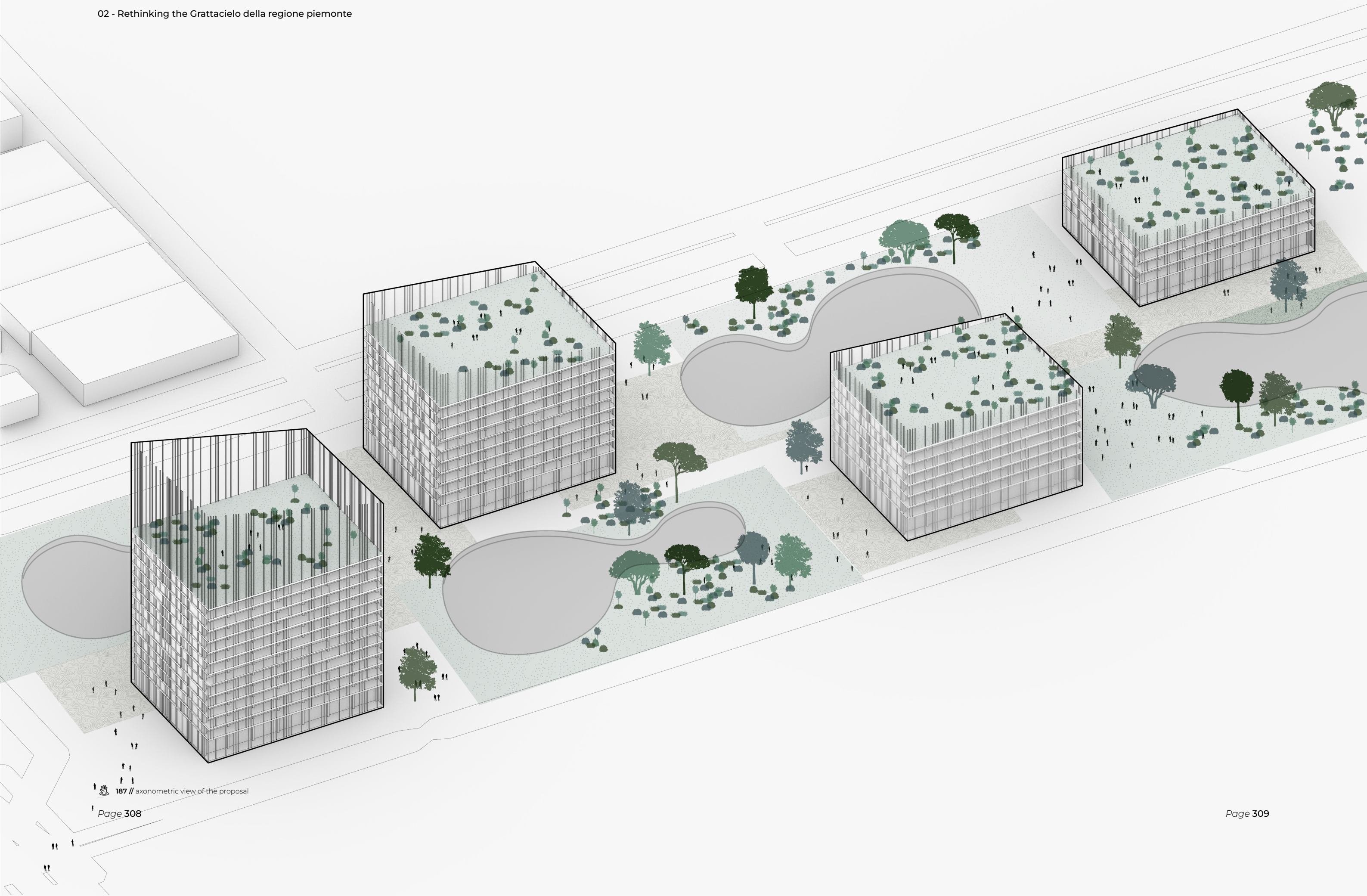
3. Rooftop garden :

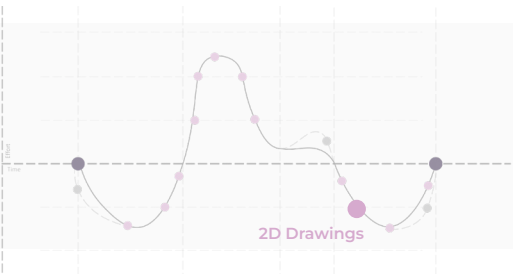
Again taking its cue from the Lingotto building, the design proposal populates the roof with vegetation and recreational spaces for employees.



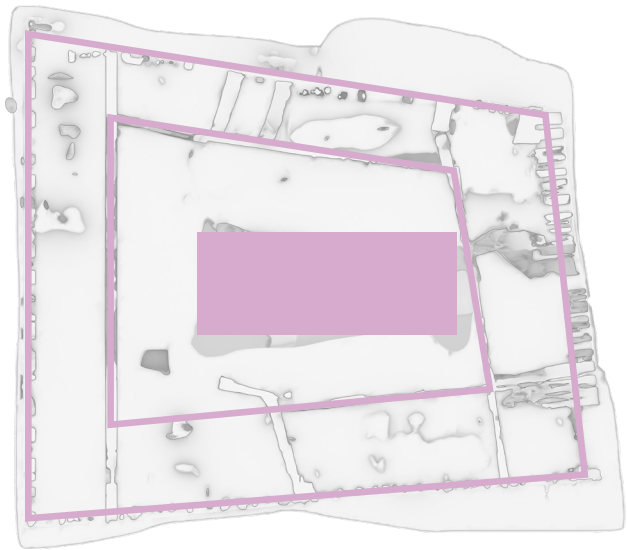
4. Ornamental twist :

In order to further break the strong geometric rigour given by the basic grid, it was decided to use shading elements giving the elements verticality, thus also giving them slenderness and lightness. They not only serve the function of privatising and shielding the interior spaces, but by extending well beyond the volume they allow the rooftop to be comfortably used.

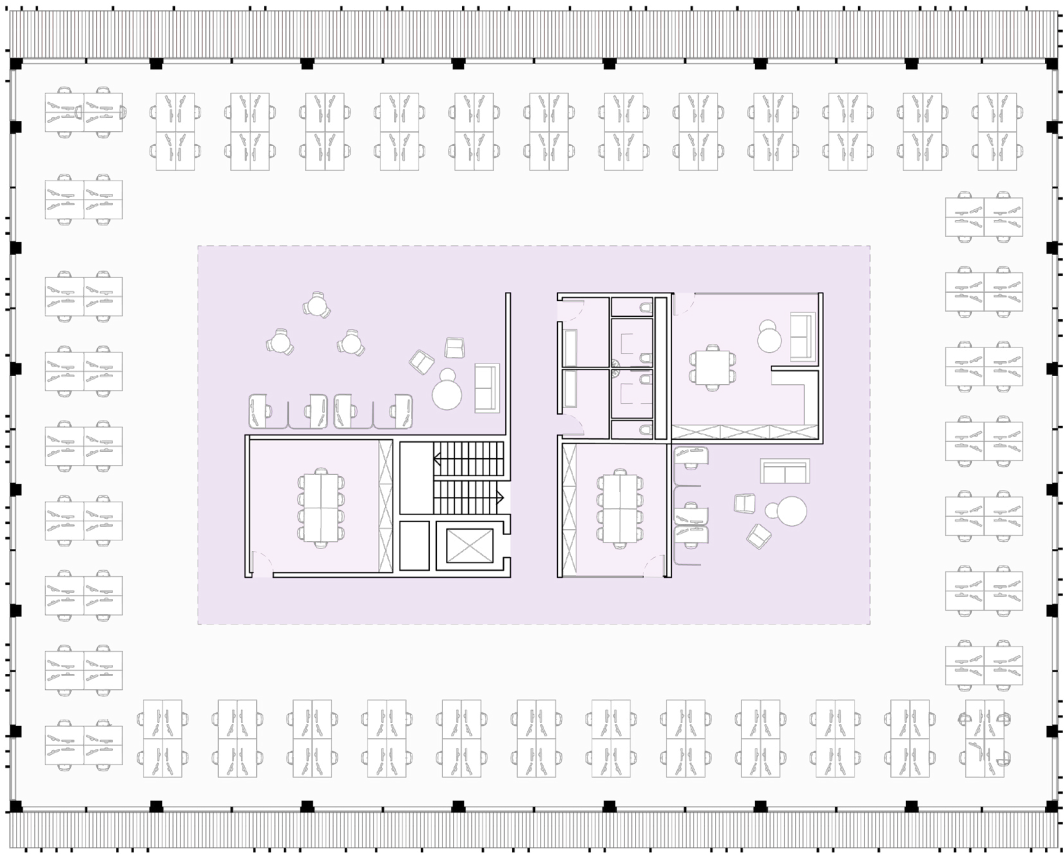




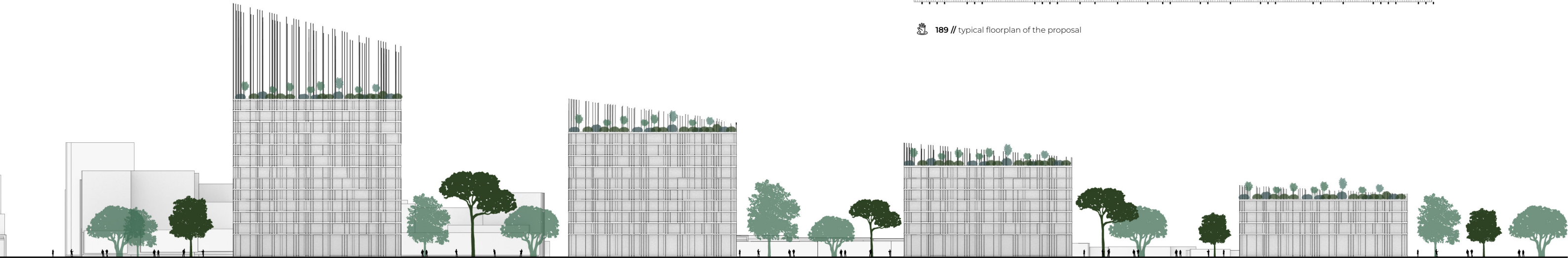
For the definition of the plan, we have broken down what the AI has done and tried to make logical sense of it, although the shapes are imprecise and not regular, we have identified certain characteristics that we wanted to bring back: a central core and a distribution of spaces around it, although simple, it could be that the AI has tried to suggest something like this.



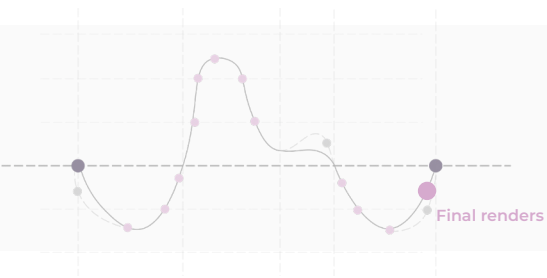
188 // interested features highlated in the ai generated plan



189 // typical floorplan of the proposal



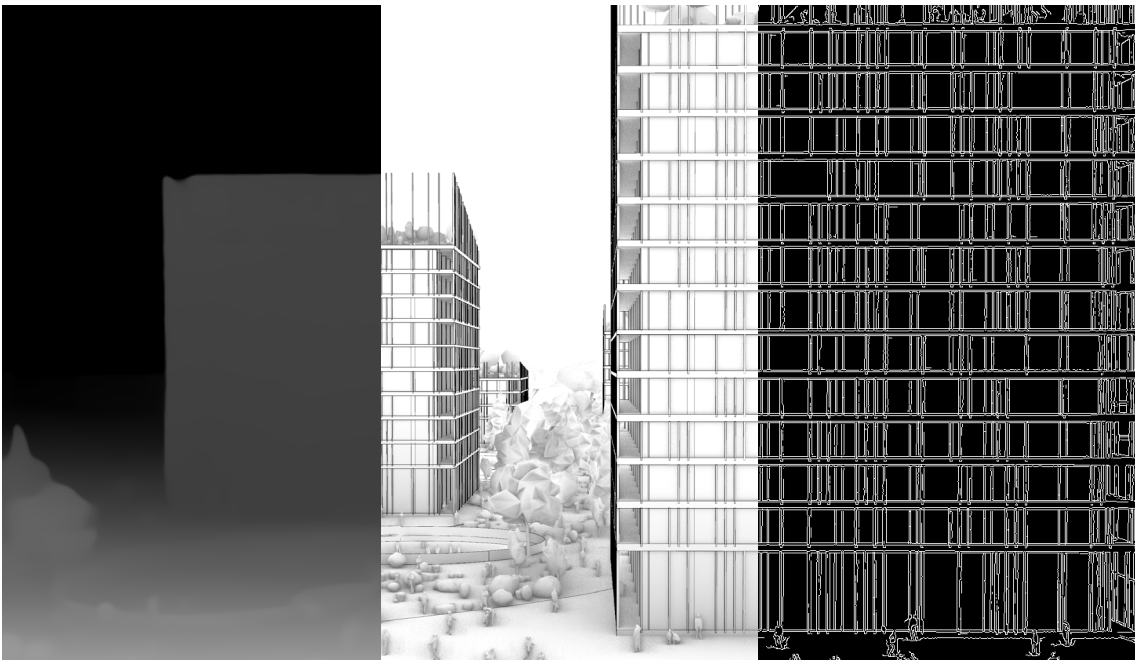
183 // elevation of the proposal



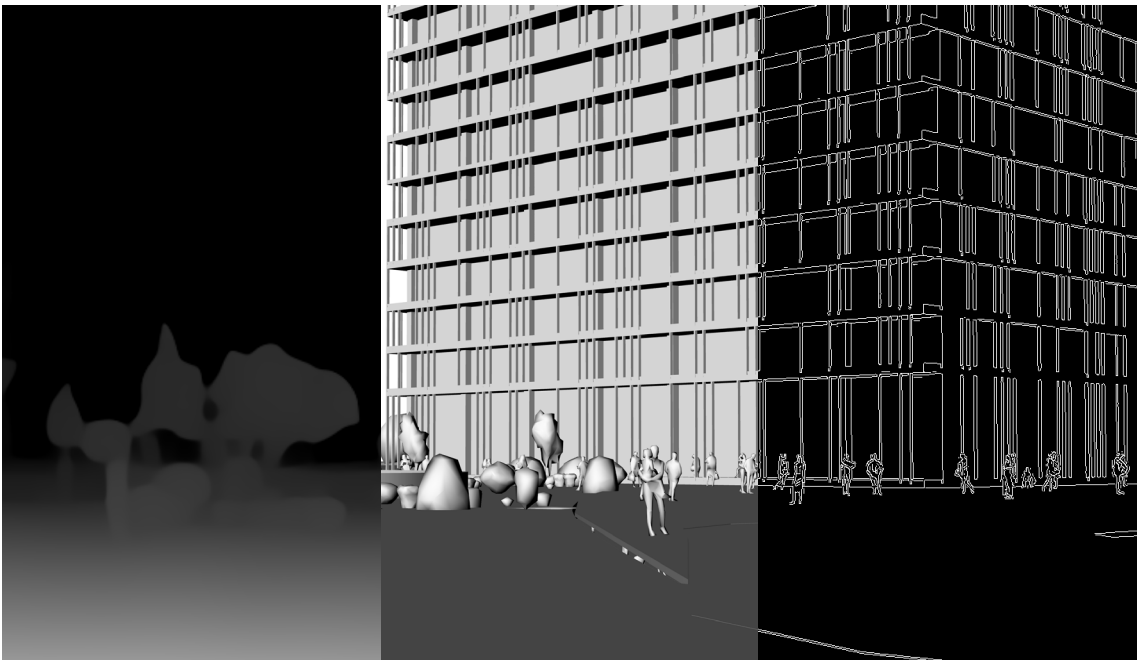
To process the final renders once we had composed a quick, more or less in-depth model, we applied the same process as previously used (reference page 226-227).

Extrapolating some masks such as canny and depth masks, we processed the images through artificial intelligence with a style reference to obtain similar results in all views.

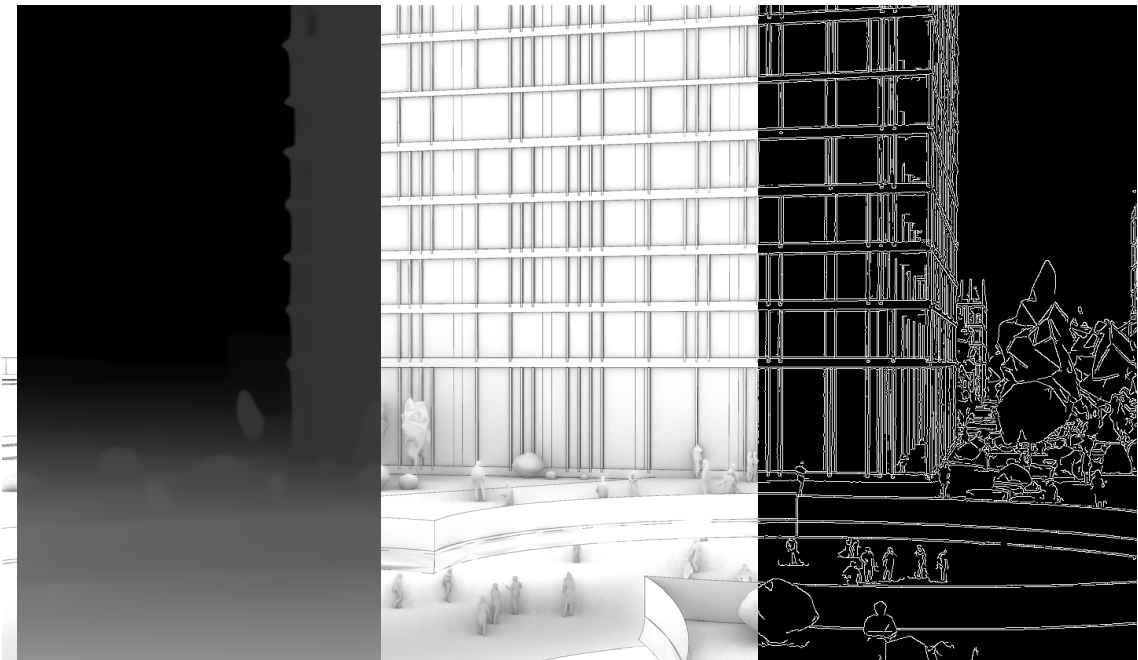
STYLE
TRANSFER
INPUT IMAGE



190 // input image 01 and masks extrapolated with AI



191 // input image 02 and masks extrapolated with AI



192 // input image 03 and masks extrapolated with AI

model : JuggernautXL
prompt : a modern building with a sleek glass facade and white vertical lamellas interrupted at different heights, creating a dynamic interplay of transparency and structure. The surrounding landscape has lush greenery with trees, shrubs, and soft vegetation enhancing the architectural setting. People are scattered around, engaging with the space—some walking, others sitting or interacting—bringing life to the scene. The atmosphere is bright and inviting, with natural light accentuating the textures and materials

steps : 35, cfg : 6, dpmpp_2m_sde, karras



193 // final output image 01

model : JuggernautXL
prompt : a pair of modern office buildings with white vertical lamellas and glass behind, creating a dynamic interplay of transparency and structure. The surrounding landscape incorporates a void on the left that adds depth and contrast between the two structures. Lush greenery is integrated with trees, shrubs, and soft vegetation enhancing the architectural setting. People are scattered around, engaging with the space—some walking, others sitting or interacting—bringing life to the scene. The atmosphere is bright and inviting, with natural light accentuating the textures and materials

steps : 35, cfg : 6, dpmpp_2m_sde, karras



194 // final output image 02

model : JuggernautXL
prompt : a modern building with a sleek glass facade and white vertical lamellas, creating a dynamic interplay of transparency and structure. The surrounding landscape is designed on multiple levels, incorporating a central void that adds depth and contrast. Lush greenery is integrated throughout, with trees, shrubs, and soft vegetation enhancing the architectural setting. People are scattered around, engaging with the space—some walking, others sitting or interacting—bringing life to the scene. The atmosphere is bright and inviting, with natural light accentuating the textures and materials

steps : 35, cfg : 6, dpmpp_2m_sde, karras



195 // final output image 03

model : Flux
prompt : a bright and modern interior view of an office building with an open-space layout. Natural daylight streams through large windows, illuminating the vibrant workspace. Employees are engaged in various activities—some focused at their desks, typing on computers or writing notes, while others are standing and chatting in small groups. The office is designed with sleek, contemporary furniture, indoor plants, and a mix of warm and neutral tones and the outside facade has white vertical lamellas. The atmosphere is dynamic and collaborative, with a sense of energy and productivity

steps : 30, guidance : 3, dpmpp_2m_sde, karras



196 // final output image 04

The Inversion Conclusions

This process, unlike the previous, proved to be particularly efficient. Integrating artificial intelligence at different stages, it not only modified the approach to the project itself but also significantly improved efficiency, especially during the production phase.

Compared to the other process, and linking back to our first exercise, we can see how reversing the workflow, starting from the final image, is a recurring and highly effective strategy.

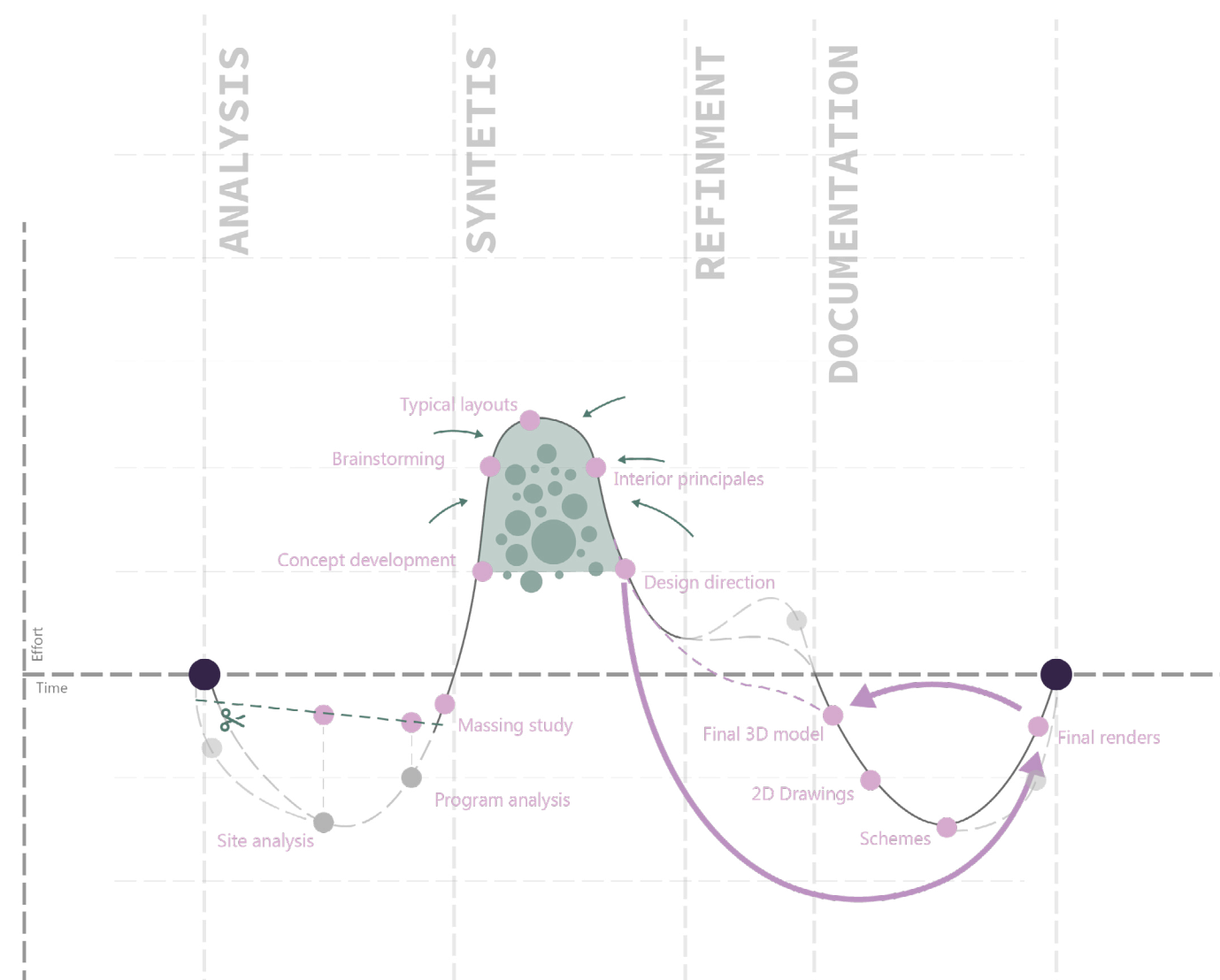
This approach can be applied to many different phases of the design process, completely reshaping the workflow without radically changing it. Instead, it adds to the process while simultaneously mitigating the effort required to reach a specific goal.

We find this to be an ideal process. However, as we have previously explained, every workflow will differ depending on the client's needs and the project's specific requirements.

Despite this, the strength of this non-AI-driven process is its flexibility. With the right expertise, the tool can be adapted in real time to fit different demands.

It allows for countless iterations, exploring façade variations, materiality, volumetrics, public and interior spatial configurations. Nevertheless, we concluded this process once all fundamental objectives, as defined by the AIA, were met, believing that further iterations would have been excessive and an unnecessary effort.

We also realized that the methodology would have remained unchanged, leading to an endless loop of reversing the process until reaching the final stage of design.



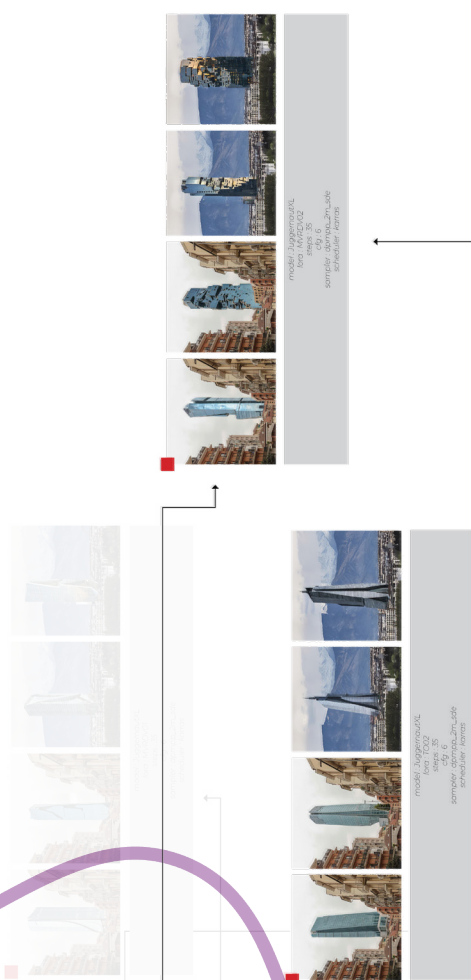
197 // result of the process with AI integration project 2 process 2

HIGH CONDITIONING

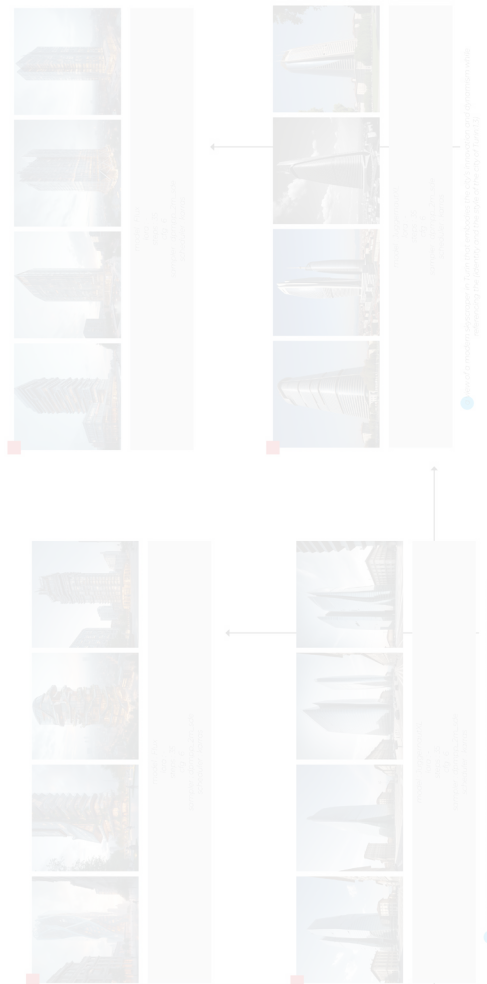
02 - Rethinking the Grattacielo della Regione Piemonte



MEDIUM CONDITIONING



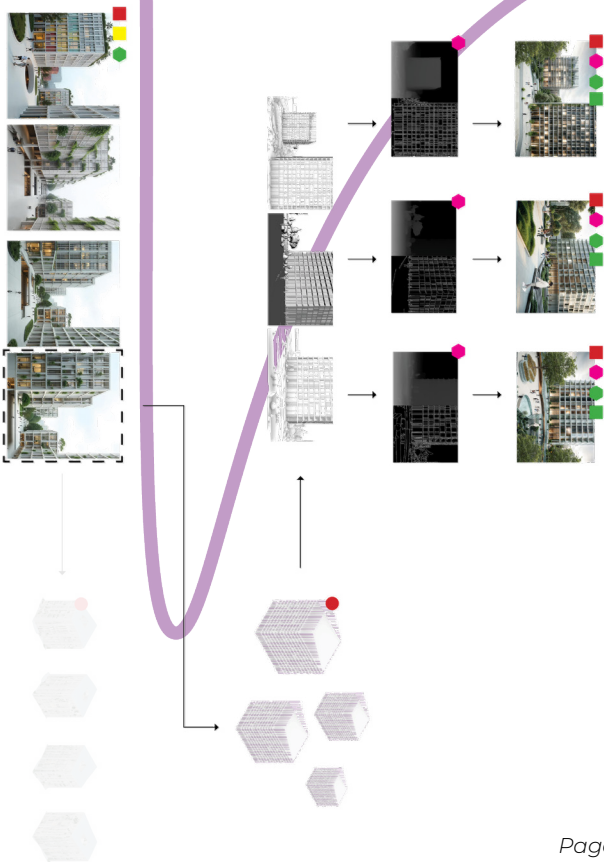
LOW CONDITIONING



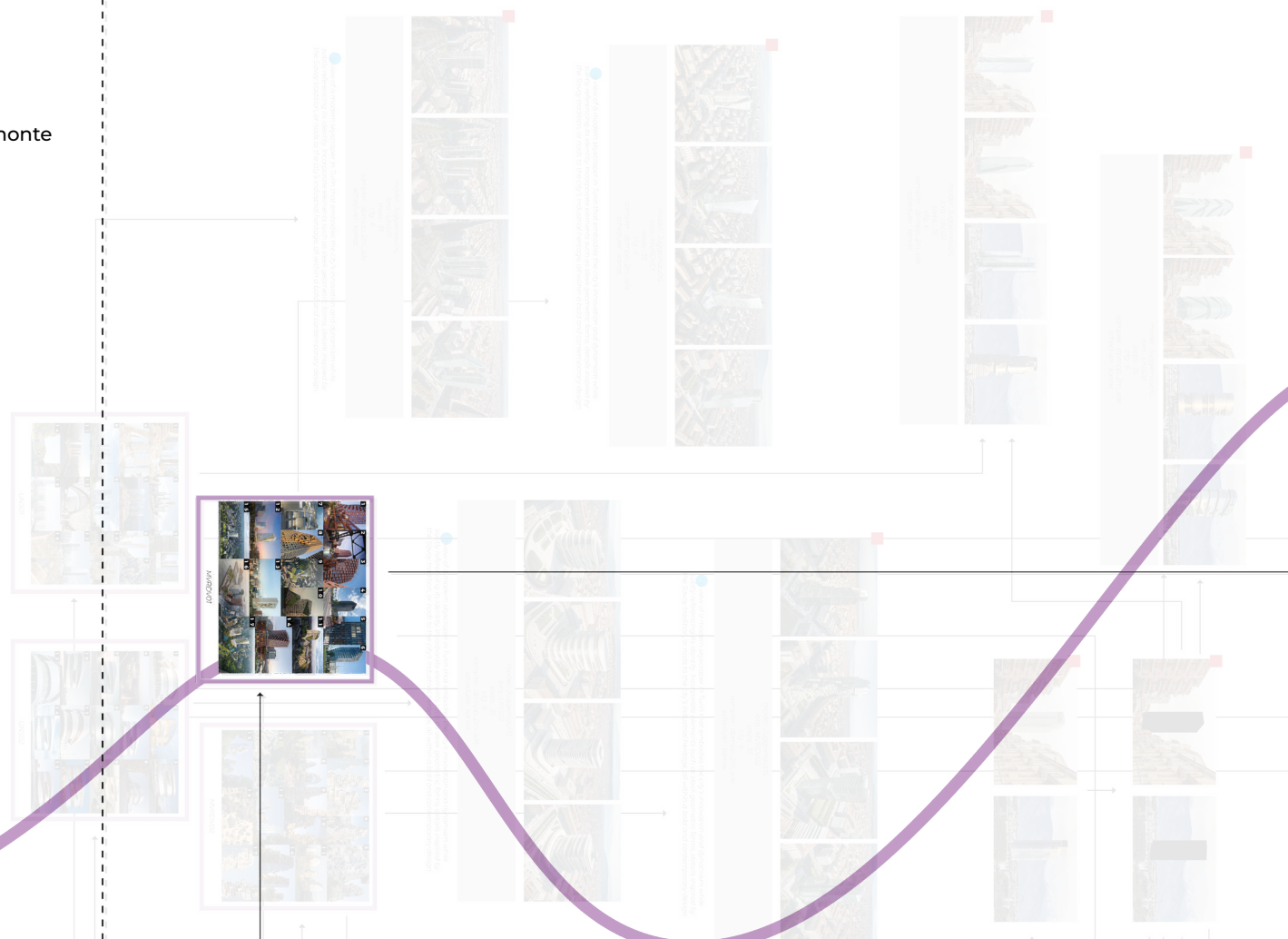
3D model
upscaling
LoRA model



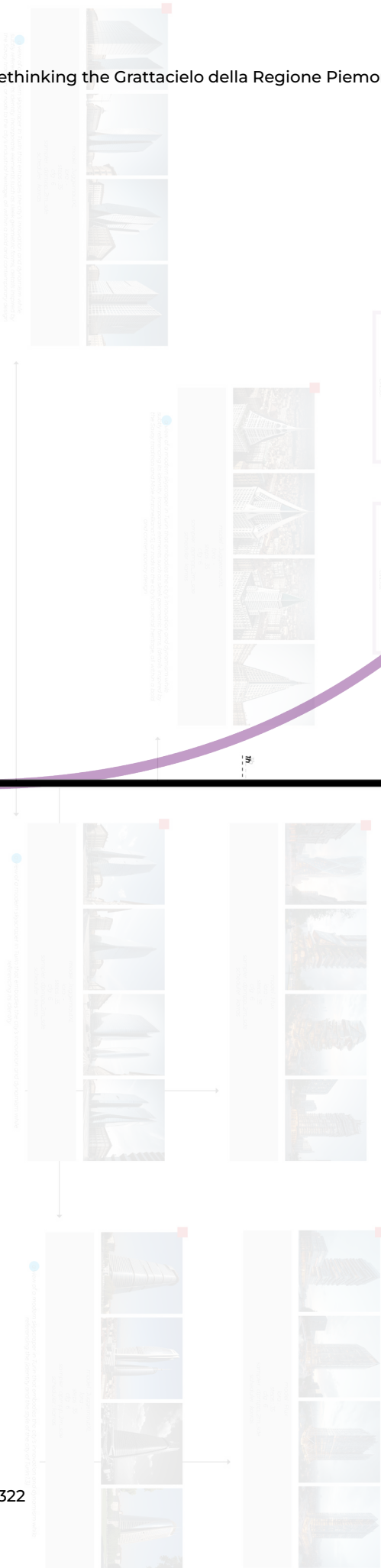
HIGH CONDITIONING



MEDIUM CONDITIONING



LOW CONDITIONING



How can we integrate artificial intelligence into the architectural process ?

Conclusions

AI-Driven vs AI-Assisted

AI in the canonic process ? How ?

During the writing of this thesis, our research focused on the current professional landscape of international offices, revealing that Diffusion Models can influence the traditional design process to varying degrees. Currently, every office is experimenting with these tools using the resources available to them, with teams ranging from highly specialized to less experienced. This highlights how young the field still is and how much it continues to require further attention, research, and study.

As we demonstrated during the design phase of this thesis, using Diffusion Models requires careful consideration. Engaging with this evolving sector demands knowledge and can be complex, but these skills are essential for building an effective relationship with the tool and understanding where it can provide the most value in the design process.

Moreover, mastering the tool distinguishes work where artificial intelligence is not just a means of improving efficiency by accelerating production but also a way to enhance design quality, enabling the exploration and creation of entirely new ideas.

Efficiency, as confirmed by all our interviewees, is the most reliable outcome. We have demonstrated how it can drastically reduce time, particularly during the production phase, by bypassing several steps that would typically be required in a traditional workflow.

Effectiveness, on the other hand, is more challenging to study and prove. Nevertheless, we have worked to develop a methodology that is as objective as possible to address this issue. Our goal was to demonstrate not only how the design process can be accelerated but also how it can positively impact the most creative phases.

These elements have been examined through two hypothesized macro-categories of processes that emerge with the introduction of artificial intelligence into the workflow:

AI-driven process

This approach, in particular, was of great interest. Through a process of dialogue, we experimented with what would happen if the project were not solely designed by the architect but instead introduced an additional actor capable of designing. However, we realized that, at this moment, such a process remains utopian and can only be applied in projects where its implementation is feasible. Moreover, the process loses strength when it cannot be carried out entirely under the strong influence of artificial intelligence.

AI-assisted process

This approach envisions artificial intelligence as a supporting agent for individual tasks throughout the workflow. Unlike the AI-driven process, this is the more widely accepted and currently feasible approach given the available tools. Viewing AI as an addition to the architect's vast toolset appears to be the most suitable, controllable, and effective strategy for now. With the right knowledge, AI can provide advantages across the entire design process, integrating seamlessly throughout its duration. In doing so, it shifts AI away from being the guiding thread and instead ensures that the architect retains full control.

Tameing the Creativity Flow with Image Generation

What is the impact on the crativity process ?

Images are one of the most powerful tools in an architect's profession, almost like a language of their own. Architects often rely on images or visual representations to communicate, especially with clients. This is because an image can convey so much about a project: its materiality, architectural composition, overall scale, and even technical details.

And yet, in the conventional design process, images are still considered secondary. Other elements, such as technical documents, regulations, and bureaucracy, tend to dominate the process.

Given their potential, however, we believe images could play an even more central role in design. Over time, their communicative power has made them increasingly valuable. Today, some clients are willing to pay significant amounts to firms dedicated solely to project visualization.

To truly harness this tool and elevate the quality of images, not just in terms of resolution, but in how they contribute to the design itself, we need to understand how to master them. The biggest challenge lies in the high level of control required to generate images. This isn't just about the user's skill level; it's also about the limitations of the software. As Cosimo Scotucci said during our interview: "...it simply can't read my mind, at least not yet...." But that unpredictability is also what makes this tool so powerful.

As time goes on, resources grow, our knowledge deepens, and we discover new ways to use AI, making it both more versatile and more specialized.

One of the techniques we explored in this thesis was the use of LoRAs (Fine-Tuners). More than just a tool for gaining greater control over image generation, it forced us to deeply understand the role of images in the creative phase of the design process. By working with it, we had to break down images into their fundamental roles, what they were meant to represent, how they should be represented, and what architectural qualities they could reveal.

This process made us realize that an image is not just a final output but an active part of design thinking. It shapes how ideas take form, influences creative decisions, and communicates spatial and material qualities in ways that traditional documentation cannot. The tool didn't just give us control over the visuals; it made us more aware of how and why images matter in architectural design. In doing so, it reinforced the idea that mastering visual representation isn't just about aesthetics, it's about enhancing the entire design process itself.

Another key factor is prompting. More than just a way of interacting with design tools, it forced us to rethink how we communicate ideas. Unlike traditional workflows, where architects directly manipulate forms and drawings, prompting requires us to translate design intentions into structured language. This shift made us more aware of how we conceptualize and articulate a project before it even takes visual form. Through this process, we realized that prompting isn't just about generating images, it's about directing the creative flow, defining intent, and refining outcomes. Mastering it means gaining control over AI as a design partner rather than just a tool, ensuring that the results align with architectural thinking rather than being driven solely by technology.

The Role of AI in the Creative Process

Between Potential and Limitations

The objective of this thesis was to demonstrate how artificial intelligence tools, particularly generative models, can be integrated into the creative design process. Through analysis and experimentation, we have shown their applicability in various forms, highlighting their notable flexibility depending on the context.

One of the most significant aspects that emerged is the layering of specialization in these tools. Some are highly generic and, due to their few internal limitations, are only useful for generic tasks. On the other hand, multimodal tools like ComfyUI, when used with a deeper knowledge of their functions, can be precisely adapted, allowing the transition from a simple text-to-image workflow to more complex systems that combine different levels of conditioning, such as ControlNet, Style Transfer, and Segmentation.

However, our research has also highlighted the current limitations of these tools. In the experimental projects, we found that while AI can serve as valuable support, it remains challenging to control. This is particularly evident when the designer has a precise idea of the desired outcome; in such cases, AI tools often prove to be slower and less efficient than traditional methods. In other words, while AI is extremely useful for stimulating new ideas and visual possibilities, it is less effective in accurately reproducing an already-defined concept.

A key aspect discussed is that AI-generated images are not mere visual representations, but real communication tools for architects. These images contain information on materials, lighting, spatial quality, architectural typologies, and technical details. This understanding is crucial: AI-generated images should not be perceived as mere aesthetic outputs but as integral components of the design process.

Lastly, through experimentation, we have developed a predictive understanding of how AI tools behave. This awareness has allowed us to optimize workflow efficiency, avoiding unnecessary delays in tasks that are not well-suited for AI and improving our ability to guide the tool with precise instructions. In this sense, mastering this new language becomes a fundamental skill for the architect of the future.

Integration, Challenges, and Future Prospects

As previously stated, this thesis deliberately omitted an exhaustive study along with replication of the entire design process, choosing to focus on its most energetic as well as conceptual stages. As projects move forward, technical details and exact specifications become more important than creativity. The research would have been larger and additionally would have moved it quite far from its original intention.

The study of AI-generated 3D modeling is important, but it has not been explored in depth. Several research limitations explain this decision: the thesis focused on AI's original potential in image generation in addition to its direct application to three-dimensional modeling. A different approach would have been needed to investigate this field and attention would have been greatly diverted from the main aims.

We tested an available tool and the results appeared promising. Precise three-dimensional models of only simple, linear geometries were generated by the software. The second project demonstrated adequate usefulness. It maintained specific proportions within the architectural composition. However, many of these tools are markedly poorly integrated into professional workflows and are not yet entirely smoothly compatible. Image generation is now more user-friendly; however, truly effective AI-driven 3D modeling still needs a lot of manual work. A number of generated models are accurate; however, those models sometimes need adjustments to integrate smoothly into professional software, such as Rhinoceros 3D, so they fit into the design process.

Generative design's use in 3D modeling can greatly improve AI capabilities. Despite current limitations, a major improvement is expected in this field. Integrating AI into architecture could change how designs are made, letting generative techniques do more than create images and giving more control over architectural forms.

Beyond Image Generation

Responsible Use and Sustainability

Another important concern, demanding mindful deliberation, is the natural potential for these tools to be misused. Their large adaptability allows for many highly personalized options, a meaningful advantage. A defined design intent is necessary; otherwise, a large risk of producing work lacking both quality and a consistent stylistic identity will arise. Artificial intelligence functions as a tool; it cannot replace all important thinking and design expertise. Its success depends on the user's knowledge and ability.

Very important is to address concerns about ecological sustainability. Although not a central component of the thesis, this topic raised some consideration. Generative AI's biggest problem is its high energy consumption, which results from the creation of many iterations, frequently lacking a clear purpose, thus wasting important computing resources. On the other hand, AI used carefully and tactically in architecture, improves workflow efficiency, reduces waste and optimizes building performance starting with the design.

The ease of use of these tools, particularly those creating images from text, demands greater care and responsibility from us. Excessive and thoughtless image generation can be dangerous. Experimentation helps understanding the tool, but this should be confined in the learning process. In design, AI is most important for improving and perfecting current processes, not for replacing them.

The future of AI into the architectural process is contingent on designers' complete comprehension of both its potential and limitations. AI use accelerates design processes, improves design quality and encourages sustainable architectural practices. Designers and AI should learn together; human expertise will improve some tools and strategies as the AI's design capabilities grow.

Bibliography

// Mario Carpo. A short but believable history of the digital turn in architecture, 2023.

// Bernstein Phil. Machine Learning: Architecture in the age of Artificial Intelligence , 2022.

// Jiaqi Wang, Zhengliang Liu, Lin Zhao, Zihao Wu, Chong Ma, Sigang Yu, Haixing Dai, Qiushi Yang, Yiheng Liu, Songyao Zhang, Enze Shi, Yi Pan, Tuo Zhang, Dajiang Zhu, Xiang Li, Xi Jiang, Bao Ge, Yixuan Yuan, Dinggang Shen, Tianming Liu, Shu Zhang. Review of large vision models and visual prompt engineering, Meta Radiology, Volume 1, Issue 3, 2023.

// Chaillou, Stanislas. Artificial intelligence and architecture: from research to practice. Birkhäuser, 2022.

// Hetrick, Suzanne H., Roy S. Lilly, and Philip R. Merrifield. "Figural creativity, intelligence, and personality in children." Multivariate Behavioral Research 3.2 (1968): 173-187.

// Brownlee, Jason. Generative adversarial networks with python: deep learning generative models for image synthesis and image translation. Machine Learning Mastery, 2019.

// Li, Pengzhi, Baijuan Li, and Zhiheng Li. "Sketch-to-architecture: Generative ai-aided architectural design." arXiv preprint arXiv:2403.20186 (2024).

// Buonamici, Francesco, et al. "Generative design: an explorative study." Computer-Aided Design and Applications 18.1 (2020): 144-155.

// Almaz, Amira Fawzy, et al. "The Future Role of Artificial Intelligence (AI) Design's Integration into Architectural and Interior Design Education is to Improve Efficiency, Sustainability, and Creativity." Sustainability, and Creativity 3.12 (2024): 1749-1772.

// Jalia, Aftab, Ron Bakker, and Michael Ramage. The edge, Amsterdam: showcasing an exemplary IoT building. Technical report, Centre for Digital Built Britain, University of Cambridge, UK, 2019.

// Caetano, I., & Leitão, A. (2019). Architecture meets computation: an overview of the evolution of computational design approaches in architecture. Architectural Science Review, 63(2), 165–174.

// Desislavov Georgiev, Radosvet. "Analysis of Deep Learning Inference Compute and Energy Consumption Trends." (2021), 11-13.

// Li, Chengyuan, et al. "Generative AI for Architectural Design: A Literature Review." (2024).

// Bernstein, Phil. Machine learning: architecture in the age of artificial intelligence.

RIBA Publishing, 2022.

// Tan, L., & Luhrs, M. (2024). Using Generative AI Midjourney to enhance divergent and convergent thinking in an architect's creative design process. The Design Journal, 27(4), 677–699.

// Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "Image style transfer using convolutional neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

// As, Imdat, and Prithwish Basu, eds. The Routledge companion to artificial intelligence in architecture. Abingdon, UK: Routledge, 2021.

// Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun. Deep Residual Learning for Image Recognition. 2015.

// Gropius, Walter. Scope of Total Architecture. G. Allen & Unwin, 1956.

// Brookshear, J. Glenn, and Dennis Brylow. Computer Science : An Overview. 12th edition, Pearson, 2015.

// Hanna, Robert. "The Non-Conceptuality of Creative Imagination and The Impossibility of Strong AI."

// López de Mántaras, Ramon. "Artificial intelligence and the arts: Toward computational creativity." (2016).

// Hui, Yuk. "On the limit of artificial intelligence." Philosophy Today 65.2 (2021): 339-357.

// Frantz, Roger. "Herbert Simon. Artificial intelligence as a framework for understanding intuition." Journal of Economic Psychology 24.2 (2003): 265-277.

// Leach, Neil. "In the mirror of AI: what is creativity?." Architectural Intelligence 1.1 (2022): 15.

// Gero, John S., and Mary Lou Maher. Modeling creativity and knowledge-based creative design. Psychology Press, 2013.

// Casakin, Hernan, and Andrew Wodehouse. "A systematic review of design creativity in the architectural design studio." Buildings 11.1 (2021): 31.

// Milošević, Jelena, et al. "Automated compositions: artificial intelligence aided conceptual design explorations in architecture." Proceedings: Geometry, graphics and design in the digital age, The 9th International Scientific Conference on Geometry and Graphics MoNGeometrija 2023. Faculty of Technical Sciences, University of Novi Sad, 2023.

Bibliography

// Beyan, Eleonora Vilgia Putri, and Anastasya Gisela Cinintya Rossy. "A review of AI image generator: influences, challenges, and future prospects for architectural field." *Journal of Artificial Intelligence in Architecture* 2.1 (2023): 53-65.

// Jaruga-Rozdolska, Anna. "Artificial intelligence as part of future practices in the architect's work: MidJourney generative tool as part of a process of creating an architectural form." *Architectus* 3 (71 (2022): 95-104.

// Agkathidis, Asterios, Yang Song, and Symeonidou Ioanna. "AI-Assisted Design: Utilising artificial intelligence as a generative form-finding tool in architectural design studio teaching." (2024).

// Almaz, Amira Fawzy, et al. "The Future Role of Artificial Intelligence (AI) Design's Integration into Architectural and Interior Design Education is to Improve Efficiency, Sustainability, and Creativity." *Sustainability* 3.12 (2024): 1749-1772.

// Fitriawijaya, Adam, and Taysheng Jeng. "Integrating multimodal generative AI and blockchain for enhancing generative design in the early phase of architectural design process." *Buildings* 14.8 (2024): 2533.

// Tan, Linus, and Max Luhrs. "Using Generative AI Midjourney to Enhance Divergent and Convergent Thinking in an Architect's Creative Design Process." *The Design Journal* (2024): 1-23.

// Hetrick, S. H., R. S. Lilly, and P. R. Merrifield. 1968. "Figural Creativity, Intelligence, and Personality in Children." *Multivariate Behavioral Research* 3 (2): 173–187.

// u Osijeku, Sveucilište Josipa Jurja Strossmayera. "Text-to-image Stable Diffusion model."

// Ho, Jonathan, Ajay Jain, and Pieter Abbeel. "Denoising diffusion probabilistic models." *Advances in neural information processing systems* 33 (2020): 6840-6851.

// Enjellina, Eleonora Vilgia Putri Beyan, and Anastasya Gisela Cinintya Rossy. "A Review of AI Image Generator: Influences, Challenges, and Future Prospects for Architectural Field." Department of Architecture, Atma Jaya Yogyakarta University.

// Jaruga-Rozdolska, Anna. "Artificial intelligence as part of future practices in the architect's work: MidJourney generative tool as part of a process of creating an architectural form." ResearchGate.

// Agkathidis, Asterios, Yang Song, and Symeonidou Ioanna. "AI-Assisted Design: Utilising artificial intelligence as a generative form-finding tool in architectural design studio teaching." ResearchGate.

// Almaz, Amira Fawzy, et al. "The Future Role of Artificial Intelligence (AI) Design's Integration into Architectural and Interior Design Education is to Improve Efficiency, Sustainability, and Creativity."

// Fitriawijaya, Adam, and Jeng Taysheng. "Integrating Multimodal Generative AI and Blockchain for Enhancing Generative Design in the Early Phase of Architectural Design Process." National Cheng Kung University.

// Tan, Linus, and Max Luhrs. "Using Generative AI Midjourney to Enhance Divergent and Convergent Thinking in an Architect's Creative Design Process." **The Design Journal**, 2024.

Sitography

// <https://rarehistoricalphotos.com/life-before-autocad-1950-1980/>

// <https://ourworldindata.org/brief-history-of-ai>

// <https://www.technologyreview.com/2018/12/19/138508/mighty-mouse/>

// <https://www.oracle.com/it/artificial-intelligence/what-is-ai/ai-vs-machine-learning/>

// <https://www.archdaily.com/785967/the-edge-plp-architecture>

// <https://www.ladybug.tools>

// <https://space10.com/projects/regenerative-futures>

// <https://www.unstudio.com/team>

// <https://www.mvrdv.com/about/team>

// <https://arch-hive.com/competitions>

// <https://www.youngarchitectscompetitions.com/past-competitions/weihai-shandong-art-museum>

// <https://github.com/comfyanonymous/ComfyUI>

// <https://stable-diffusion-art.com/how-stable-diffusion-work/>

// <https://github.com/tencent-ailab/IP-Adapter>

// <https://www.labellerr.com/blog/unveiling-florence-2-microsofts-cutting-edge-vision-model-for-diverse-ai-applications>

// <https://www.tripo3d.ai/>

// <https://civitai.com/>

// <https://architecturecompetitions.com/?open-competitions>

// <https://torino.repubblica.it/cronaca/2020/02/09/news/>

// <https://www.facebook.com/share/p/18NGTcPvBF/>

// <https://www.facebook.com/share/p/19VUEFtb3o/>

// https://www.regione.piemonte.it/web/sites/default/files/media/documenti/2022-08/torre_grattacielo

// https://it.wikipedia.org/wiki/Grattacielo_della_Regione_Piemonte

// <https://blenderneko.github.io/ComfyUI-docs/Core%20Nodes/Sampling/KSampler/>

// <https://comfyui-wiki.com/comfyui-nodes/sampling/k-sampler>

// <https://medium.com/@shahram.seifcar/how-to-use-ai-in-architecture-and-managing-the-design-process-phases-1116286190ac>

// <https://wc-studio.com/journal/2019/7/5/working-with-an-architect-understanding-phases-of-design-construction>

Appendix A – Interviews with industry professionals

MVRDV Answering

Lorenzo discusses how AI tools like MidJourney, Stable Diffusion, and ComfyUI are enhancing architectural processes, offering real-time modifications and new creative possibilities. While he values these tools for inspiration, he emphasizes that AI won't replace the core aspects of architecture, such as sociological, regulatory, and ethical considerations. AI will undoubtedly increase efficiency, but human expertise in materials, structure, and design details will remain crucial.

Lorenzo Mattozzi

Architect | Associate



Mattia : Let's start with contextual understanding. How do you perceive the current landscape of AI in architecture? Is it truly transformative, or is it being overhyped in some respects?

Lorenzo Well, I think it is being overhyped for sure, but in a way, rightly so. It's a natural process. You need this overhype, and then you finally settle into a more rational approach. The overhype is more of an emotional response to the possibilities that AI entails. But mainly, the overhype comes from people who approach it for the first time. As soon as you dig in a little deeper, you start seeing its limitations as well, and you begin to rationalize your approach and understanding of it. It is a transformative tool for sure, and to me, there is no way back ... no returning to a time without AI. But it's not the kind of revolution some people make it out to be.

Mattia : What do you see as the most critical research area or question regarding generative AI in architecture today?

Lorenzo : So, as you mentioned, you're looking into text-to-image diffusion models. If we focus only on that aspect of generative AI, I think the critical research area really depends on which offices we're talking about. For some offices, the most important research is about creating models that can replicate a certain language or style. Many large offices work with a precise style in mind, and AI models tend to respond better to that. If an office has a distinctive architectural language, it's easier to mimic and replicate. For smaller offices, or those without a widely recognized style, the critical research question is more about how to integrate generative AI into the workflow in a way that is efficient and not just another way of wasting time. An established office with a recognizable style will likely ask, "How can I adapt this process to produce my kind of architecture?" For everyone else, the question will be more like, "How can I use this process efficiently?" Otherwise, AI becomes just another distraction.

Mattia : Yes, I agree. Let's move on to the next chapter, which is about the role of AI in the design process itself. Do you see AI as a continuation of technological evolution in architecture, or does it represent a completely new paradigm?

Lorenzo : I think it's a technological evolution. We had AI before, and now we have generative AI. So, I don't see it as a complete shift. Talking about the overhype again, I don't think AI is going to replace people, at least not in architecture. But, of course, you need to get used to these new possibilities and work with them to stay competitive. You can't just ignore AI and act as if it doesn't exist. In that sense, for people unfamiliar with it, it might feel like a new paradigm. But if you've followed the evolution of technological tools in architecture closely, AI is just one more tool in the toolbox. Computer-aided design didn't replace architects; it just made them work in a different way. The same goes for 3D modeling and BIM software. Generative AI is part of this same progression. It doesn't replace anyone or erase people from the process; it just forces them to adapt and use it. That's all.

Mattia : Let's go deeper into this topic. What are the potential and limitations of AI in generating design ideas? Can AI ever be a true partner in the creative process, or is it just about efficiency?

Lorenzo : I think AI can be a partner. You can build a sort of relationship with the engines that generate images, refining prompts, understanding the tool, or developing a better workflow by. It's really a relationship you need to develop. You have to know the software well, speak its language, and understand what it can provide for you. Sometimes, it's easier to be surprised by the software than to get exactly what you have in mind. If you already have a clear idea, it's often faster to execute it the traditional way. To me, AI is a brainstorming tool; it's not the end goal. If I have a very clear design solution in mind, I can develop it in many other ways that are more efficient than using AI. Then, I might use AI again to refine it multiple times. That's where its real power lies. Right now, there is a tendency for AI to homogenize language and style. You could even call it "MidJourney architecture" because you can recognize when something has been generated by MidJourney. But as AI improves, this may change. Still, these programs work by learning from existing data. They need past references to generate new outputs, so some degree of homogenization is built into their learning process.

Mattia : That's a very interesting insight, especially how you describe this duality. MidJourney can help conceptualize new ideas and perspectives, but at the same time, it's a great tool for refining existing work.

Lorenzo : Yes, I actually think AI works even better for refining rather than for starting from scratch. When you already have an idea expressed graphically or visually, AI can help refine and enhance it. The times when I've had only an idea in mind and tried to get generative AI to produce exactly what I wanted, I was often the most disappointed. It never quite matched my expectations, and trying to force it to do so became a huge waste of time. In those cases, I found that modeling it myself, applying materials, and rendering it manually was actually faster.

Mattia : Let's talk about the different phases of a project. How do you see AI being integrated across different project stages, such as preliminary design, concept design, and schematic design?

Lorenzo : I think AI is particularly useful in two key moments. First, during the concept design phase, it can serve as a brainstorming device when you're open to exploring many possibilities. Then, later in the process, especially for interiors, it becomes incredibly powerful. Some programs allow you to input a simple white model with basic furniture and introduce just a mood board, without even needing a reference image. The AI then applies colors, materials, and atmosphere based on that mood board. This is really valuable, especially for communicating design options quickly and clearly to clients. So, I see AI being most effective in those two phases: at the very beginning, when you're generating an overall concept, and near the end, when finalizing elements like façade details or interior atmospheres. In the middle phases, when things need to be more precisely controlled, AI tends to be less efficient. It works best for broad ideation at the start and detailed refinement at the end.

Mattia : So we can agree that during the production process, it's hard to introduce the use of AI. It can be very useful, as we said before, to fine-tune designs, but the production of material and documentation has different needs.

Lorenzo : There are specific moments where you need to present and sell an image, but a big part of our work is not about selling an image anymore. It's about developing details, producing accurate floor plans, or analyzing constraints, rules, and regulations. So, a significant portion of our work has nothing to do with text-to-image generative AI. However, AI can still be used in other areas, such as sorting regulations or analyzing climate conditions. We can use AI to calculate climate conditions, such as noise and wind forces, through generative and predictive models. In essence, AI can support technical and analytical processes beyond just generating imagery.

Mattia : Alright. Then, do you think AI will primarily serve to enhance efficiency in architectural practice, or do you see it as a tool that could foster entirely new forms of creativity?

Lorenzo : Right now, it's not really enhancing efficiency. I don't think so. Not yet. That's the conversation I'm having with our colleagues: "How can we find a workflow where this is really efficient?" Because the way we've used it so far has not been very efficient. Finding an efficient workflow for text-to-image generative AI is a whole research topic in itself. As for finding new forms of creativity and design, I'm not sure either, but it is a secondary topic to me. What interests me right now is how we use AI as a part of our workflow in an efficient manner. That, to me, is the most interesting topic. In terms of new forms of creativity and design, I don't think AI will necessarily lead to that. It's difficult to design something entirely new without finding some precedent or similar concept that has been explored before. I don't believe AI will create entirely new creative paradigms, but I do believe in making it a part of everyday life and using it efficiently.

Mattia : My colleague and I are working on a thesis focused on the practical integration of AI into architectural workflows. We are analyzing standard architectural practices and how processes shift due to deadlines, client feedback, and design changes. Our goal is to explore how AI can help manage and control these shifts more effectively. That's our question: How can we do that? At the same time, we want to raise awareness about the correct use of AI in design, as creativity can be compromised if AI is not used properly.

Lorenzo : I think if you work with AI software and generate something, what you are generating is a form of creativity, and it is personal. You and I are not going to use the tools in the same way, and we're not going to land on the same results. I think there is still authorship in the way we use AI tools and software. Just because you use AI doesn't mean there is no authorship. So, the issue with generative AI is not about authorship or what comes out of it, because I'm totally in favor of crediting authorship to the person who used the tool to develop an image, for instance. The issue we usually have with copyrights is how the model was trained, but that's not really a concern for the end user. It's something we can question, of course, but that's a responsibility of whoever developed the tool or software: how they trained it, whether they used transparent methods, or if they sourced images ethically. The result of using the tool to generate something, to me, belongs to the person who created it. I have no doubts in that regard. Some critics compare using AI to cheating. To me, it's just like drawing a floor plan in AutoCAD or Revit instead of by hand. Just because I'm using a computer or a technological tool doesn't make the drawing any less mine. AI is just another tool, and as such, we use it in different ways and achieve different results, for which we are responsible. We're both the users and the authors.

Mattia : Then we can move on to the more technical part. What AI tools are you currently using, and how do you assess their capabilities and limitations in practical terms?

Lorenzo : Well, I'm using several AI tools, either directly or indirectly through software that incorporates AI in the background. A direct tool would be ChatGPT, for instance, which I use to organize documents and draft texts. An example of an indirect tool would be Forma, which calculates sun exposure, noise, wind loads, etc., using predictive AI. However, if we limit the discussion to text-to-image generative AI, I mostly use MidJourney and Stable Diffusion. Recently, I've started experimenting with MagnificAI and Runway, though I am most familiar with Stable Diffusion and MidJourney.

As for capabilities, well, they keep improving. Some of these tools are updated regularly, often introducing significant enhancements. Of course, as capabilities grow, limitations tend to diminish. However, limitations still exist, and they will remain until AI can fully interpret human intent. Even with all the advancements, I still struggle to get exactly what I want. When I have a clear vision of the result I expect, that's when I struggle the most with text-to-image generative AI. I hope it will eventually help me get there faster, and it's frustrating when that doesn't happen.

To me, that is the biggest limitation, but it's also a limitation on my part, on how I use the tool. I sometimes realize that the issue is not just the tool itself but also how I communicate with it. Nevertheless, in some cases, AI truly fails to understand what you are asking. I've tried in the past to generate basketball courts on mountain landscapes or in unusual settings, and for some reason, the courts are never depicted correctly. You would think such a simple request would be easy to achieve, yet it proves to be extremely difficult. I believe AI will improve, but in practical terms, there are still plenty of limitations, which makes it not as efficient as it could be. As we progress, I expect these tools to become more accurate and significantly enhance efficiency.

Mattia : Specifically, what are your thoughts on tools like ComfyUI and MidJourney? Have they changed the way you approach architecture?

Lorenzo : No. The short answer is no, they have not changed my approach to architecture. That's because architecture is not just about generating images. It's far deeper than that. There are sociological aspects, construction constraints, cost considerations, urban integration, sustainability, and ethical issues. Architecture encompasses so many dimensions that reducing it to just the production of images would be an oversimplification.

AI tools have not changed my approach to architecture, but they have certainly inspired me and broadened my creative process. Sometimes, I use them for inspiration, and they can be quite effective. ComfyUI, in particular, is fantastic in how it operates, allowing for real-time interactions. That, to me, is the next level: using AI in real time.

For example, imagine working on a physical model: you modify it, and on a screen, you instantly see those changes translated into a fully developed building with different material options. This combination of a highly analog approach, where you physically manipulate a model, and digital augmentation through ComfyUI is incredibly fascinating. That's something I'd like to explore further. I think it could become a standard part of presentations, meetings, or workshops. It could be a great way to foster discussions about design directions.

However, architecture is much more than that. We still need to read regulations, understand the client's vision, align with the city's ambitions, and fulfill a range of practical requirements ... things that no single image can resolve. That's why we still need people behind these software tools.

Mattia : Looking ahead, what do you believe will be the most significant paradigm shift in architecture due to AI over the coming decades? Do you foresee the emergence of new roles or a restructuring of the architectural process as a result?

Lorenzo : I think the main paradigm shift is that everyone will need to engage with AI in some capacity. Text-to-image AI may remain more relevant to the design and visualization stages, but every aspect of architecture will use some form of AI. People, willingly or unwillingly, will have to acknowledge and integrate AI into their workflows. I do see new roles emerging. We might have AI specialists in architectural offices. Depending on the team and the type of work, there could be dedicated experts in AI-assisted visualization or even AI-driven structural optimization. You may have AI experts focusing on material applications throughout the design process. However, this will not fundamentally restructure the entire architectural workflow. Instead, AI will be integrated as an additional tool, expanding possibilities, but not replacing the fundamental principles of architecture. For example, we might soon be able to change the entire atmosphere of an interior with just a few quick AI-assisted steps, making the process faster and more efficient. But that does not alter the overall architectural methodology. It's similar to the way scripting has become a staple in architecture. Today, most teams have at least one person proficient in scripting because it streamlines work. But not everyone needs to script, someone in the team does it when needed. I believe AI will follow a similar path.

Mattia : Looking further ahead, do you think AI will enable more personalized and responsive architectural design, or will it standardize and limit creative expression?

Lorenzo : Again, AI is just a tool. it's like a knife: you can use it to prepare a great dish, or you can misuse it entirely. If you don't develop the right skills, you'll likely end up producing something generic and repetitive. It's the same with AI: if you use it lazily, you'll just generate uninspired copies of existing styles. I mean, how many Zaha Hadid-style projects can you look at in one day before they all start to feel repetitive? However, in the right hands, AI can enhance creativity rather than restrict it. It all depends on how people use it. For some, AI will expand their creative potential; for others, it will become a shortcut that leads to mediocrity.

Mattia : Could AI lead to new methods of communication between architects and clients?

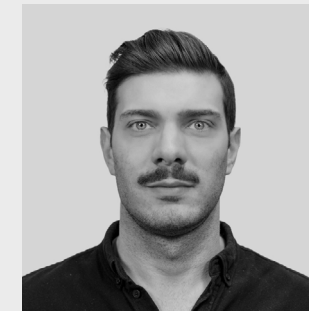
Lorenzo : I hope so, and I think I touched on this earlier. For instance, imagine bringing a physical model to a client meeting, along with a camera, ComfyUI, and a screen. The digital model could update in real time as the client interacts with the physical model. That would be a fantastic communication tool. However, adoption is often slow. Take virtual reality (VR) as an example. We've had it for years, and it's fully capable of immersing clients in a design. We even assumed VR would become standard in meetings, allowing clients to "walk" through a virtual building. Technically, it's easier than ever, yet we rarely use it. Why? Because it takes time to set up, and there are practical challenges, such as only one person can experience VR at a time while others in the meeting are left watching. For similar reasons, I think AI-assisted presentations will take time to integrate into standard workflows. At the end of the day, clients appreciate fast and simple tools that won't introduce technical issues. Bringing a physical model to a meeting remains an effective communication strategy because it allows for direct interaction. If adding AI to the process requires setting up a camera, a high-performance laptop, launching software, and troubleshooting potential glitches, it risks becoming a liability rather than an advantage. On paper, AI has tremendous potential to revolutionize how we communicate with clients. But whether we actually take the necessary steps to integrate it into everyday practice remains to be seen. It will take effort to make AI an essential part of architectural workflows and client presentations. For me, that's one of the most fascinating topics to explore.

MVRDV Answering

AI is transforming architecture by streamlining workflows and supporting creative exploration. Cosimo sees tools like ChatGPT and Stable Diffusion aiding in research and idea generation but notes that AI struggles with tasks requiring creativity and abstract thinking. While it can optimize efficiency, human input is essential for interpretation and deep design expertise. In the future, AI will help automate routine tasks, but the architect's role in refining and applying ideas remains crucial.

Cosimo Scotucci

Senior Project Leader | Architect CNAPPC



Mattia : let's start with the contextual understanding so from the more open picture than we go further in detail, how do you perceive the current landscape of AI in architecture? Is it true transformative or is it being overhyped in some respect?

Cosimo : Well, I believe AI changed our lives. Life as renew both professional architectural life and general life. in a sense that the impact of ChatGPT as being insane, nothing that was seen before, so in that respect has been transformative. I cannot name a project or a file that get out of an architecture office that has not been double check , created or simply fine tuned by ChatGPT. So in that sense, the landscape of Architecture and the landscape of life has been changed radically, for the rest, all the other AI are being introduced and used somewhat in the in the architectural projects, but definitely it's not the demanding of change is not comparable to the one that ChatGPT had.

Mattia : Okay. Uh, then for the next question, what do you see as the most critical research area or question regarding generate AI in architecture today?

Cosimo : Well, it will be amazing to get to architecture, so at the moment we can create a text, we can create images, but we're far from generating architecture. We can create plants, but it will be nice to incorporate this. I've done a little bit of research on um 3D. I mean there are some like AI assistant you can incorporate in Rhino, you can do a little bit of volumes and things. But yeah, we're very far from a certain type of use for AI in offices like MVRDV. I do believe that if you are a builder, you know, in somewhat near future, you'll be able to generate buildings where the AI it's completely different story to generate architecture with the capital A, you know, something that has been through from different people with different background, analyzed the contextual social economical factors with the complex modeling, if you know what I mean, you know. That's that's probably the most critical part of the research, integrating these kind of models into the best practice, not on the standard practice.

Mattia : Yeah, then so in some aspect, the thesis itself addresses this problem. I's say, uh okay, we have the AI, it's a very good, as we said, you can for his nature of recognizing patterns and repeating them, is very logical, okay? And it's very um useful in that way. So, for the normal profession is could be amazing and uh also for someone that is more in a in the practice of constructing itself that works a lot with data. It's amazing. It's an amazing tool. But right now we're trying to challenge the AI to see, if it helps also with the creativity and how it can help the creativity of the the mind of the architecture profession.

Cosimo : Yeah, I'm also surprised that paradoxically what ChatGPT to do is if you ask the tool to name how many sea or I don't know, how many "r" are in the word Ramarro? It fails hundred times, you know, in a row. It cannot get it straight because it's not logical. It's like an illegal in a way, so it does analyzed data things, but they obviously nobody has written anything about these things, right? So there is no data where AI do to to generate the correct answer, right? So that's that's the thing, allucination of ChatGPT. I mean, we're not talking about AGI, right? So that's far from where we are now.

Mattia : Absolutely. I agree. Um, then we can move to the next uh, let's say, uh topic. AI and the role in the design. You are already talking about it. And uh do you see AI as a continuation of the technological evolution architecture or does it represent a completely new pardigm?

Cosimo : Well, at the state of the art is indeed a continuation. I mean, the way I see and we use it is no more nor less than, you know, generating some references and stuff like that. So in a way, it's I mean, you can generate your own references, you can generate images that it was hard and long before, but yet again, it's at least in the top practice in the world, there yet it's a completely new paradigm in the profession. I'm sure it will become, but I see that like at the beginning, there was gigantic hype on the on AI. Then little by little somewhat and then like fewer and fewer people are using it on a daily basis. So we all know that time to time we can get back to AI and do things with their AI when we need to do a certain type of job. But yet again, I don't see people using it on a daily basis, and I do believe because the besides ChatGPT, everybody is using it. But nobody's using it yet on the daily basis, all the other forms of behind, because it's not ready yet to be used a certain level. Again, if you design white boxes, possibly you can through, I don't know, four or five different AIs, generate floorplans, get an image and stuff like that. But like when you somewhat raise the bar, then it's where things a little bit falls apart.

Mattia : Yeah, also, it's not a very user friendly. So you need to um spend a lot of time and effort to learn something that probably is not going to give you like the uh the final result, then you need to tweak it anyway. And probably you're gonna lose time.

Cosimo : Indeed. Yeah, it's also different kind of indeed interface, you know, as architects we're not that used to text. I mean, I don't know who's used to uh prompt things, right? But that a paradigmatic change in the field. So we need to start thinking about prompt which is a word that probably two, three years ago, nobody knew, right? Oh, and everybody now it's on, it appears daily million of times, right? So that's something that is changing. We need to see what how to do our prompt. I don't know if there exist any sort of prompt engineering, but you know, that's something that we need to look up into.

Mattia : Then, next question, how do you see AI integrated across different projects phases such as Prelimeter Design, Concept Design and Schematic Design?

Cosimo : Common factories again, ChatGPT because in the PD, you look for the information about the context you're working in, you generate the text of the projects, stuff like that. Then obviously, later on in the face, you keep using ChatGPT to do all sort of research and, other things. For the rest, different AI can serve different process, so you can use it, but then again, it's really personal. So if you like AI if you or at least that's the people I will confronted with. So if you know a certain AI, you tend to use it on certain specific parts of the process and and stuff like that. So we all say that Midjourney it's very conceptual while Stable Diffusion, maybe you use it later on, this because one is very creative, the other one is maybe not, right? So in that sense, I think there are different kinds that you can use throughout the whole process depending on what you do.

Mattia : These three phases, let's say they are much more creative than the other ones. So that's why we're challenging them and not the other ones because if you go on the other ones, it's more databases, right?

Cosimo : Yeah, but they still in Schematic Design, you do a lot of facade iterations, right? So something that used to be tedious and and very long is something that now you do it in blink of an eye. And that's what I like about the this, because, you don't need to put three people on a task that was a test, and then they need to spend eight hours on it, but then you can take a person doing like how many tests you want on Stable Diffusion and then spend the other two for other task. So that's amazing. So I surely making our time more efficient is making our daily life more efficient. yet again, I mean, with ChatGPT when you need to do research, you go straight to the point, you ask for sources, you double check the sources, right? So you don't need to spend like insane amount of hours into searching for one for something that might not exist or might not even be true. That's amazing. Or with other tools, you can generate references that are much more specific for what you whatever you have in mind. So in that sense, our AI is making our life easier in all these phases, I mean, I've seen people doing floor plans in a blink of an eye, right? So in a way, I'm sure you can use them very efficiently in all these phases,

yet again is what can you do with it? So what I've seen is that as soon as the architecture do has a certain complexity, then AI struggle to understand that complexity, you know, if again, if you do a box, three story high, shoe box building, I'm pretty sure you can do it all of it in AI. But like as soon as you do complex stuff, then obviously it's hard.

Mattia : And that's a basically you responded already to the next question, because it's: Do you think AI will primarily serve to enhance efficiency in architectural practice, or do you see it as a tool that could foster entire new forms of creativity in design? So are we confronting ourselves to a tool that wants just to build more stuff in quantity or do we confront over searching something that it's more qualitative?

Cosimo : the way we use it at the moment is at least the way I use it at the moment it's a tool to enhance my efficiency, for sure. But yet I know cases like this ARUP, years ago, they had to develop like a node for like a multi angular steel connection. So there was this very complex shape they made, aggregating pieces, and then they asked, hey, can you generate the most optimized shape for it? And AI came with something that is like a blob, strange looking object that we as humans will never come up with, right? but then it was like they were saving a certain percentage of steel, it was much more straightforward to the task that he was supposed to do. So I think both answers are right in a way. So you can use it to enhance your efficiency, but obviously can help you generate unforeseen design options in a way.

Mattia : Then we can pass to the next chapter, let's say, technological evaluation. Which AI tools are you currently utilizing and how do you assess the capabilities limitation in practical terms?

Cosimo : Yeah, so again, one is what became that my best friend from July 2022 is ChatGPT. Again, that's something I'm using on a daily basis for like a variety of scopes. Then on top of that there are like some other AI like Midjourney, Stable Diffusion or Magnific that we use on a daily base, but mostly when needed, right? So ChatGPT, you use it like from day zero because you wanna know climate or specific area or specific region and then you start using all the other AIs as soon as you going into architecture, right? The the best thing about ChatGPT is that it's generic. You can use it like in every form of field while all the others you use it as certain stage of the design when you need an image, right? The problem with architecture at the moment is that all these tools or most of these tools are not told to be architecture specific, so if you ask to Midjourney to generate the image of a ballerina it will get it will be perfect, right? Or the face of a, I don't know, oversized Chinese woman, right? It will be perfect. It will be, you know, and then you can generate millions of these iterations, right? And it will be beautiful. But then, as soon as you put it into architecture, because it's not been created for architecture, you start to create having lacking and creating problems and stuff like that. You know, or as soon as you try to create something that is not basic architecture, that's when things falls a little bit. We can go draw in it like yet again, I do believe that's what you saying, that not everybody's in you is using it because you acquire the knowledge to use it properly, it takes time and it's a different completely framework to operate within. I believe it, we will get there and also when REVIT came up, it was a different interface, it was a different kind of working, right? But then like at the beginning, we all thought, it was weird a new kind of interface. We didn't want to interface with it, but then we all ended up using it. I think it's just a matter of time.

Mattia : We can skip also the next question, because you already responded to it. Looking forward instead, What do you believe will be the most significant paradigm shifts in architecture due to AI over the coming decade? Do you foresee the emergence of new roles or a restructuring of the architectural process as a result?

Cosimo : In a way, It will enable us to look for the unimaginable, or to create this kind of weird combination or oxymorons, right? So when you combine things that seem to not go along with each other, sometimes we struggle to do so, but then with AI I do believe we'll be able to combine these things and make them work and functional and beautiful. So, the believe that the future is gonna be is bright. I mean, we will get more fusion, we will get more creative, we will get like the most accurate data and clients will be able to verify the data we give them. In that sense, the whole game will elevate, I'm very positive about architecture. AI in general, and in life. The node thing from ARUP I was talking about. That's amazing! I mean, if you only think on a kind of an environmental fog, I mean, you're saving within. If we apply this on every scale from the little node between two elements of the building, to the whole building, to the whole master plan of the entire planet. So it will enable us to compute data that, you know, it impossible for us to compute. And uh so we can um, you know, design life uh a an unprecedentedape.

Mattia : I think the same, but the thing that concerns me the most and this also concerns Freddy, for example, is the amount of work, we're going to reach some sort of a efficiency, much higher than now. The client itself is going to expect even more and even more and even more that the time passes. What is going to happen? are we going to lose quality on the work? Are we going to be able to keep up or also insainability levels? Is it seriously sustainable, as a thing? Because we are using a lot of energy to do this stuff.

Cosimo : To respond to the first part of your statement. Yes, we used to do things with pencil and paper. So when we start doing things with computers, in a way, we became faster, better, more accurate, etc. So do clients shape the process? I don't think so. I mean, the technological abandonment somewhat dictated what we could do in what time, you know? So I don't think it will be the client dictating what we can do, but it will be the technological advancement together along with the purity of what you can generate with that. If in our far, far, far future, you take a photo of a site "I want a building here" and then you can do it with your phone connected to your brain, changing, blah, blah, blah, and obviously things will get but this sci-fi is not reality in a way. Second one, yes, we consuming a lot of energy, but we know that our life in general needs energy, right? All living things, they somewhat use energy, right? We as a human being, we eat to move around and do what we do, right? Animal plants, they all do the same. So, and we know that the biggest shift in uh human life, has been done when we introduced like gas oil, we elevated the energy level in our society that's when the the lifespan got increasingly longer, when we could, like everybody had like a significant step up into the game. So population is still growing will be, ten billion soon. We need tools that migrant zoom certain energy to make sure we can increase the energy level of the people of society. At the moment we have a very like a big disparity, you know, between countries, we have the north and south of the world. I think we will need to somewhat balance it these two things. For instance, I do believe that the way we will find unconventional kind of solution that will help us saving energy here and there, making things more efficient, making things circular, right? In a way that everybody can rise up their bar of life, as we know, it can progress. I'm an optimistic person, in general, but that's I do believe the the common goal of humanity as a whole, right? I do believe that a third party, something that is not human in a way can help us, you know, getting there. If we compared with like super smart being, that we can make things faster and faster. The beauty of history, is that you build upon what has been done, right before. So the problem is that at the moment we don't have a circular economy in a way that we reach we do products that reaches an end. The byproduct of these end piles up piles up till becomes problems. We need to understand that someone's trash could be someone else's treasure. I think this kind of comb computational power, you know, this kind of uh different perspective on things it can be eventually transforming to some one else good through a different perspective. I can look into these things and easily figure out what to do with our byproducts, in a way. This is already happening, so that's where I think we're going.

Mattia : Uh, always looking ahead, do you believe where I will enable more personalize and uh responsive architecture design or would it standardize and limit creativity expression, clear to a new methods of communication between architects, client and public?

Cosimo : That's a funny thing because, as soon as you start working with MidJourney, if you don't know how to prompt properly, the first results you get are often these half-cooked, fake Zaha Hadid-style buildings. With earlier versions of the tool, it had a tendency to generate very organic designs by default. If your prompt was vague, you'd likely end up with something blobby essentially "Zaha-ish". There is a risk of that happening, but I'm optimistic. I think AI will become more personalized over time. For example, we're already seeing AI systems that tailor advertisements to individual preferences. There are tools, like ChatGPT and similar systems, that position themselves as personal assistants. These AIs can be trained on specific data your research, your interests, or even your workflow. I think we're moving toward more personalized, responsive AI systems, but perhaps not specifically for architecture. Instead, they could connect your search history, the books you're reading, and the architecture you're exploring to generate tailored outputs. This process mirrors what we do as humans: we gather information, experience it, and metabolize it. Somehow, in our minds, connections happen, and we reuse the data we've collected in meaningful ways. I believe AI, especially if systems become more interconnected, could enable us to have not just a few architects or designers but billions of them.

Mattia : The last one instead, could we lead a new methods of communication between architects client and the public. Is it going to transform the relationships?

Cosimo : In a way, yes, because briefs are written in the AI, right? The texts we send are written in AI. So, in a way, it's already the middleman between these two parties and the public. I think editors, in general, already work this way. For example, you'll get the transcript of this interview through AI you'll likely put it through some kind of text extractor, right? So, in a way, we go through you and the public when you present this lecture.

I wonder if one day we'll all use digital versions of ourselves. You could be giving a presentation in China while you're physically in the US, relaxing. Another version of you could be attending a meeting in Europe simultaneously.

Mattia : But I said, good or is it bad?

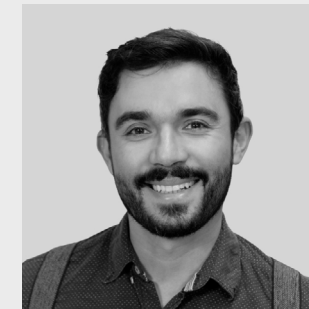
Cosimo : I don't know. I really don't know. I think only time will tell. I don't know. I mean, yeah, of course, it would be nice to give the presentation in China and then come back to Europe to attend a meeting and then eventually go to the US to chill out. But that's life as we know it, yeah? God, it's really impossible to say what the future will look like now because, you know, all these, like, the Moore's law and all that—it's saying that we're progressing at a certain speed and so on, et cetera, et cetera. And, you know, it's really impossible to tell where we're going. Like, I don't know, five years ago, if you told anybody about a tool like ChatGPT, they would laugh at you. And now, all of a sudden, everybody's using it. And so on and so forth. I think I believe that technology is something that never goes backward. So in a way, life uses technology, adopts it, and, in a certain sense, also adapts to technology. So I guess we will see. We will just adapt. I think we will adopt technology, but we will also use it for our own good. And when I say our own good, I mean as architects, obviously, but most importantly, as a species on this planet. So, yeah, I don't know. I mean, really, this is the first time in history when sci-fi feels like the present. You know, you see all these things people do on video, and every day it still amazes me to see what people can do. You know, I've seen a person being interviewed by their digital twin. It was amazing to see that the digital twin remembered things much more clearly than the person. The guy asked, like, I don't know, what did I write in that book? Can you explain that chapter to my grandma? And the digital twin did it smoothly, and I understood it. It was a financial thing, and it was very hard to grasp. At first, it presented what was written in the book, which was impossible to grasp for someone not trained in that area. Then he said, explain it as if for a two- or five-year-old boy, and then he did it again for my grandma. It was amazing how it enabled communication in a way. So that communication was targeted to a specific audience. The digital twin managed to communicate a very difficult concept to a non-trained group of people. And ideally, it could have done it in Chinese, Hindi, and English, one after the other. That, I found amazing. So, in that sense, I don't know if it's bad that the digital twin is presenting a project in China, in Chinese, following strictly the input I gave. So it's somewhat representing me my beliefs, my goals, my boundaries. It won't derail from it in any way.

MVRDV Answering

Fredy discusses AI's potential to boost efficiency in architecture, from quick renders to data-driven design. While it may allow specialized architects to produce more, he expresses concerns about its impact on creativity and the environment. Despite AI's capabilities, the role of the architect in critical thinking and refining ideas remains essential. AI will enhance productivity, but the human touch in design is irreplaceable.

Fredy Fortich

Architect | Technical specialism



Mattia :How do you perceive the current landscape of AI in architecture? Is it truly transformative, or is it something that's being overhyped in some respects?

Fredy : Well, personally, I have seen, and I got into this field because I saw it was going to be transformative for me. I saw it from the beginning, even when the quality of the images wasn't as convincing to other people. I saw that there was a transformation that was going to happen, and it will continue to happen as the tools become better and better.

This, not only for architecture. I think it's going to transform everywhere; it's going to affect all disciplines, and as the technology progresses, it'll become even more transformative. I'm not talking about just Diffusion Models, but also LLMs will play an important role, because then you will be able to do these sorts of connections between human language and software.

Usually, we have to learn software to be able to produce things, but if the tool is good enough that we can actually just speak the tool learns our language, rather than we having to learn the tool's language. So, that's the transformative part, where you don't need to learn to code; you need to learn to give instructions to produce the code or to produce the Dynamo script or to produce whatever.

So, it's transformative in the way that it changes our connection to the tools. I think we'll use it in all aspects of creation, from ideation to improving, and it will be basically as common as electricity. You will not even see that you're using AI, but already, you somehow are using AI in social media, and all this is run by algorithms that are tailored for you.

I feel that it is transforming. But of course, there is this sort of hype cycle (Gartner hype cycle) that I have in my presentation, which is basically the hype graph, where you have this peak of inflated expectations. Then people think, "No, this doesn't really work for anything," and they kind of go into the valley of disillusionment.

But then, after, as you start to use it, and you become better, you go through the slope of enlightenment, and you achieve this sort of extra boost in productivity that was not as big as the original hype, but you will anyway have it.

That's how I see it. It'll be less, for now, than what we expected, but it will continue to improve and improve.

Mattia : Then the next question, what do you see as the most critical research area? Or question regarding generative AI in architecture today.

Fredy : I find that the most important thing to research is control. There's this sort of lack of control. Getting your idea is very good for brainstorming, but that has its limitations. What I find most important is having a sort of control, being able to guide it through text, through parameters. That can be achieved with different approaches. It's not just an idea that came out, but an idea that's also informed by your approach to it. So, this sort of control will be very important.

But there's also critical research about text-to-3D or video-to-video. For example, text-to-video, I don't find it very interesting. I find that it's a bit gimmicky, at least. For filmmaking and cinematic experiences, maybe it makes sense. But for architecture, text-to-video doesn't make sense to me.

All these text-to-video software are like slot machines. The ones I've tried, like Luma and Runway, are very good in quality but not in control, and that's what's missing. The videos you see online, everything, probably involved at least 20 to 30 videos that turned out badly before finally getting a good result. This is a concern of the community.

If you have enough time, you will eventually get a nice generation, but if you have more control and predictability over what you're going to get, you're not losing that time. The whole point of AI is to make it more efficient and faster. But if you're spending so much time on the slot machine to get a good image or video, then it kind of loses the whole point of efficiency.

Mattia : Next question, do you see AI as a continuation of technological evolution architecture, or does it represent a completely new area or new paradigm?

Fredy : Yeah, of course, it's a continuation of architecture. I try to always give the idea that it's not an AI architect or AI specialist. It's an AI tool, and you use it as a tool. You leverage it from one direction or the other. You are the one that chooses. You're the one that's bringing critical thinking. You're the curator of things. But it is a change, or it will more and more become a change in the way we design. I think that already we're starting to see examples, publications where you can definitely tell that it was inspired by AI. You start to see a sort of MidJourney-esque language. It's not just AI, but you definitely start to identify a sort of language, and that's the paradigm shift that I'm talking about. You will have more ideas and more interesting ideas with a smaller team. Then you start to refine the ideas. There's this super article, I don't know if you've read it already, called "Intelligence Super Abundance." I recommend it, super interesting.

Mattia : What are the potential and limitation of AI in generating design ideas? Can I ever be a true partner in the creative process?

Fredy : I think it will do both. For now, the people who start using it first, the practices that start using AI, will have an advantage, for sure. As with any new technology, the early adopters are the ones who can have the advantage of being able to produce more, to create more.

Mattia : What are the potential and limitation of AI in generating design ideas? Can I ever be a true partner in the creative process?

Fredy : Old school architects will think no, but I think yes. It's definitely true. More and more people are going to start using it in the creative process. As you've seen in the design approach image that I shared on social media, you can use it in all parts of the design process, not just brainstorming, not just AI rendering, but everything in between. Regarding limitations, I see consistency, repeatability, control, but also limitations in the ideas that come out. You get ideas that are not sustainable, not constructible, not feasible economically. This is where the role of the architect, with his critical mind, comes in. The architect understands the ideas, and the design becomes more interesting. There's also the limitation of copyright. That's why I say that we at MVRDV only use images that have been trained on our native images. Then there's the limitation of text-to-3D. The models now are very polygon-intensive, not too precise, and need a lot of work to refine them. Hopefully, we'll start getting a bit cleaner models from text-to-3D, but this is just the initial stage. And just as images improved so quickly, I feel that this will improve quickly as well, maybe not as quickly as images because they're easier and less heavy to train, but eventually it will. You'll start to see a lot more clean 3D geometry.

Mattia : Regarding the models itself, there are no checkpoints and models that are just based on architecture, will it change something? Will it be very good change? because I see that not a lot of the industry is still capable to do something like this, training up whole models and starting building nodes and workflows, just for architecture, right now that field is just on anime, cartoons and other kind of stuff. Film making and architecture is trying to get inside, what do you think about this?

Fredy : I think that's why, for example, Ismail Seleit has done so well, because he's one of the few people, or the initiators, of adding this just for architecture. We've done it here at the office as well, and this is why we're also training LoRas. For example, if you have a LoRa, you can merge it into a checkpoint. In that way, you have a specialized checkpoint. So all models come from one base model, and you can refine or change the model by just merging it and utilizing a fine-tuner. If we talk about Stable Diffusion 1.5, as Flux, every other model you see in Civic AI has a base model that was trained by the big companies that have the resources to actually train a very complicated or very

A highly parameterized model, then what you need to do is simply fine-tune it by teaching it a little bit extra to the base model. Using LoRas with the base model helps to avoid heavy components in the process. Training the base model is a bit better, but doing LoRas is much more efficient and easier. For sure, we're going to start seeing more architectural models that are very specific to architecture, trained with more images. It's just a sort of no man's land right now, where people are experimenting, but it will become more and more professionalized and tailored to architecture.

Mattia :how do you see AI being integrated across different project phases, such as player in linear design, PD, concept design and schematic?

Fredy : Talking about diffusion models, I see them integrated everywhere: brainstorming, super Pinterest, collage design or segmentation, design, material, materialization, massing, ambience change, rendering. From preliminary design to schematic design. If we talk about a more global AI, then I would see larger language models such as ChatGPT or oLama. The LLM that Meta released (LLaMA 3) is super interesting because it's open source, just like Stable Diffusion. You even see Mark Zuckerberg talking about it, saying he knows that the future is open source, because it's other people who are improving your model, and he's not having to pay for that. So, at the end, it's like the community improves its own, and the community gravitates toward a model that's free, but improves it at the same time.

Mattia Next question, do you think AI will primarily serve to enhance efficiency in architectural practice, or do you see it as a tool that could foster entirely new forms of creativity and design?

Fredy : I think it's not one or the other. I think it's both. Enhancing efficiency could be as simple as taking an image and improving it, making it much more realistic looking. Instead of spending so much time rendering it, you can render a low-quality image and then upscale it, or just render the base structure of the image and then animate it. So that's an enhanced efficiency. But also, if you go in the middle of the process, you can have, for example, brainstorming, and that's where I see the real magic coming. It's enhancing efficiency because it generates a lot of proposals quickly, but it also increases creativity. So, it's more than just a yes or a no. It's more like a gradient. There are smaller teams with more specialized architects, that's how I see it. Big companies will be able to do more with fewer people, rather than having a lot of extra people. So, roles will definitely change to more technical roles. For example, I don't like the term prompt engineer, but more like AI engineer. AI engineer instead of prompt engineer, because it gives the idea of not a specific thing, but something more related to the overall, and this helps to control the outcome much better. I feel that we're going to have more competition. Smaller offices will be able to do what big offices were able to do only because they had much more resources. Smaller offices will have the same capacity. This will definitely change the landscape of who's doing what, with only the best surviving. The final image will lose importance, but the idea, the step diagram, the storytelling, and how you can convince the client that it was the best idea, will become very valuable. That's why, here at MVRDV, you not only learn that you got to an idea because you like it, but more that there was a process behind it.

Mattia What are the AI tools that you're currently utilizing?

Fredy : The current tools that we're using here at the office are ChatGPT as an LLM tool, Stable Diffusion, and recently Flux for Diffusion Models. We also have a MidJourney account, but it's more for beginner users just getting started with these tools and for brainstorming. The core, more interesting, customizable part is using tools like ComfyUI, which is a multimodal, node-based connection system. In this tool, we can create our own workflows and really develop it into a more complicated tool. I'm also seeing a trend where AI tools are starting to become more and more expensive as more people realize how useful they are. For example, the SORA model is now 200 euros a month to use. That's why I say the future is in starting to use tools like ComfyUI, where you can download an AI minimax, or have your own model, video models that you can use for the same amount of money, but open-source, free, and customizable even if you have to pay a license for them. But it's not about the cost of the license.

Completely different from the cost of using a monthly subscription. I feel that the more open-source, customizable tools are the most interesting ones, like ComfyUI. Again, everything that you have to go through a paywall with, a monthly fee, and wait for them to improve certain things, I think it's good for the general public and for general users, but to really push forward, or have a research and development department like MVRDV NEXT, we need more tools that we can shape to what we want, rather than waiting for the next update that MidJourney offers.

Mattia We need more people who are capable of working with this kind of stuff. This is because, in that way, we can utilize something that's not dependent on some companies, on the release of updates and software. For example, we saw with OpenAI that they currently have huge problems with capacity, so you don't have the freedom to use it as you would like. Instead, if you have a team that is specialized in this, it will be much stronger, helping to control the outcome better.

Fredy : Yeah, I agree. There's also another issue, and that's privacy. When you use local models, you know that all your information is kept inside MVRDV, and they're not going to use it for some other purpose or change the rules. You know that your LoRas stays within MVRDV, and no one has access to it unless they take it purposefully, illegally. But with tools like MidJourney, you have no idea what they're going to do with the images you create, or what happens if you have a top-secret competition and somehow it leaks that's a problem.

Mattia What do you believe will be the most significant part of the shift in architecture due to AI over the coming decades? Do you foresee the emergence of new roles or restructuring of the architecture process as a result?

Fredy : There's going to be smaller, more specialized teams. This doesn't mean that everyone has to become an AI architect, an expert at using the tools. I think we'll just simply have more AI ambassadors, called in for certain specific tasks. We need, just like in MVRDV NEXT, the scripting expert or the BIM experts, people specialized in that, who can really make your project much more efficient.

You can still do it the old way; you can still do everything in AutoCAD. This doesn't mean that all architects have to go in that direction. Architecture is a broad field. You can also focus on studying history or detailing.

For example, detailing is something that I don't see happening in the near future, maybe if we have an AGI (Artificial General Intelligence) that understands the concept of what you need, instead of just understanding the image. There's an interesting article I read about LLMs where it was mentioned that the neurons activated when it reads the word 'cat' and when it sees an image of a cat are the same neurons activated by the San Francisco Bridge. It suggests that it's not just about understanding images, but that it understands the concept.

Until we have AI tools that can grasp concepts or principles, for example detailing, you need to understand the principles of water-tightness, R-values, and the climate of a location. You're not going to do the same detail here as you would in a tropical area. So, until that level of understanding, I still think the human architect is what truly excels.

Mattia And the last question looking ahead, do you believe AI will enable more personalized and responsibly architectural designs, or will it standardize and limit creativity expression could lead to new methods of communication between architects, clients and public?

Fredy : Yeah, so, of course, we kind of touched on this already, more personalized and responsive architecture comes from doing your own training, or now it's called LoRa training. In the future, it might be something else, but basically it's fine-tuning a model so that it does what you want it to do.

Many people say, 'Oh no, I'll just wait until I use the commercial tool.' But if you wait until then, you're going to have a standardized, limited creative expression. That doesn't interest me in particular. There will definitely be new ways of communicating, making everything more efficient. But also, if you have a story to tell, you can tell it much easier through a video than through text.

The live aspect isn't quite there yet, but with augmented reality and AI, you can really start to have a different discussion. It's live, and you don't have to wait until you get a

result and then the client says, 'I like this.' You can start working together and see it evolve in real-time, humanizing the architectural design process. This could reduce the back-and-forth, making meetings more efficient. The client would know much more about what they want, having explored it together with you. It could also generate the idea that clients expect everything to change quickly, but there are pros and cons to everything. Personally, I'm a technological optimist if you have a better tool, you'll probably have better relationships with the client and better results.

Mattia The bonus question, the last one: sustainability. We can see that AI can help a lot in improving how we read data, especially large amounts of data, and that's a more technical part. But it will also be highly productive in terms of the environment. As I understand it, the world is moving towards a direction of producing more, producing more, producing more and AI is only boosting this further. So, what's going to happen? Will it be beneficial for sustainability, or will it make things worse?

Fredy : I think that if you look at it just from the standpoint of generating images, it could be worse, because it definitely requires more energy. I think we've already gotten worse. Ten years ago, the client wasn't expecting a rendered image, they were expecting a sketch or a hand drawing, which was much more sustainable. Rendering farms and companies that operate independently of AI already consume energy, so, sustainability-wise, in terms of creating images, we're going down the wrong path, that I agree with. However, I also see that AI opens new opportunities and new economies. Now, data is the new gold or the new oil. AI helps you organize data and discover new opportunities. For example, when looking at cities, instead of using a linear approach where materials are taken, used, and then discarded, you could use AI to identify resources in a different way. Instead of mining natural resources from the land, you could mine resources from cities, from data. That's why you need certifications and methods for organizing this data.

AI, in principle, is pattern recognition it helps you recognize and repeat patterns. So if you ask the right questions, you can get the right and better results. It really depends on what you ask and how critically you approach it.

But, yeah, if you see it only from the energy perspective, we are always searching for ways to create more energy. That's the human goal more energy, the better. AI could allow for more efficient models, or models developed by companies like Google or OpenAI, which might lead to more efficient materials and possibly breakthroughs in energy production. That's what I think; as I said, I tend to be more of a technological optimist.

UnStudio Answering

The impact of AI on the world of architecture is explored. For Olaf, AI can certainly improve efficiency by being able to generate fast renders and optimise projects, but some doubts are expressed regarding the data on which AI is based and its effect on creativity. Although in the future there could be an exponential increase in productivity thanks to AI, enabling a specialised architect to do the work of ten, the human role of the architect remains fundamental.

Olaf Olden

Sustainability Technologist
at UnStudio



Petru : [How do you perceive the current landscape of AI in architecture? Is it truly transformative, or is it being overhyped in some respects ?](#)

Olaf : I think it's truly transformative. It's way easier to generate ideas, and you don't have to sketch everything, which you can make it tangible making quick AI renders, or simply just changing materials on the go in a matter of minutes, let's say, or seconds even. I do think it's a bit overhyped. Architecture is not only making nice images. And I think a big part that architect plays is being the main director of the process as well, with all the consultants, with the client. I think the architect is like in the middle of that circle, and making sure that all the different disciplines come to place in terms of, like, visualization, yes, there plays a big, big part. But I think it's a bit like, overshadowed. Especially in the era where we live, everything needs to be there in 5 or 10 seconds, we need AI render here, AI render there. So I think it's a good tool, but it's not much more than a tool. And we still need the architects much more than the AI renders.

Petru : [Yeah, we don't have something that keeps everything together. We can generate images, but, I mean, we can not make details or solve even more complex architectural topics.](#)

Olaf : No, no, no, indeed. So in that sense, it's still really like an empty vessel. It understands how to put textures on a building from the outside, but it doesn't communicate back with what's happening inside. But that will change. I believe, yeah.

Petru : [What do you see as the most critical research areas or questions regarding Generative AI in architecture today ?](#)

Olaf : One thing is the whole essence of using AI is based on historical data. So I mean, if you generate an AI image based on Tadao Ando or Zaha Hadid, Kengo Kuma, whatever, how original is it and does it? Does it block creativity? Because, yes, we can make AI renders, but they are inspired by historical data. And I am wondering, is that also, being a limiting factor for true creativity to still be there, right? Because we stay in the bubble that we already know, and I think that's a very important part, because I'm really wondering simply because the essence of AI is based on historical data, how do we go towards the future and still have 100% creativity there? There are questions regarding generative AI and architecture today. Yeah. But I also think because what's happening in the past. You always need to be very careful with renders, because as soon as you put a nice render to a client, that is the image they print in their head, and that's the building that they want, that's something that was always there. And since we can produce so much more images as such a rapid increase of, like, time efficiency. How is that part gonna go forward? Right? Is it then, so that maybe the client is also gonna use his like, comfy UI and generate some images and tell you, hey, I want this.

Petru : Also, as an architect, you have to be sure that the image that you show to the client it is feasible, and not futuristic images that are clearly made with AI.

Olaf : Yes, indeed. It does need to be feasible. And I think the nice thing of sketching, if you see a sketch in itself, it's a work in progress, it's like a way of thinking, and maybe you need to put your AI renders with like a Lora of like sketch on top of the AI render that you get out of it. But since that's always such a critical part of the past, I'm wondering how that is going to continue towards the future, where it's becoming very more tangible and visually driven. In images, let's say.

Petru : Next question, do you see AI as a continuation of technological evolution in architecture, or does it represent a completely new paradigm ?

Olaf : For sure, I would say it's a continuation. I think architects always like to try to capsule their information and their knowledge within images. But I think back in the day, it's a bit more like diagrammatic stuff, or really like sketches. Now we have the power to within that same time frame, we can come up with very realistic renders. So I do think it's more like upping our game in order to have visually talking with the client. Like, what do you like? Do you like a wooden facade or not? I think those parts so for sure, it's more continuation. But I think the more and more AI gets, get stronger. As soon as inside and outside are going to talk with each other, I think then it really becomes like a paradigm, because that basically means that at that point you can make a building with the prompt. I think then it's really becoming a paradigm, because then you have like architects that is just like prompt to architecture, and now it's really prompt to render, but then it becomes prompt to architecture.

Petru : What are the potential and limitations of AI in generating design ideas? Can AI ever be a true partner in the creative process ?

Olaf : This was the point that I was raising right ? Just to phrase it one more time, AI stuff is based on historical data. And I am wondering, how can that go, like, hand in hand with pure creativity and historical data, maybe you can steer the creative idea that you have a bit better with bouncing back and forth. I think there's a very fine line between where you lose most of your creativity and the wherever you can still use it. So yeah, it is a critical area.

Petru : How do you see AI being integrated across different project phases such as Preliminary Design (PD), Concept Design (CD), and Schematic Design (SD) ?

Olaf : I think the earlier the better. That is the main thing at this point now. But for sure, also, if you talk, let's say from the outside of the building, AI can do a lot, especially now, already from the inside of the building there are so many rules and regulations, even between countries, that AI needs to be very smart at one point. I think there's always a fine balance between structural methods and what kind of span am I gonna do ? What is the thickness of my beam ? What is the thickness of my floor ? It will be nice that if you have like an outline of a building, and you say : I want to have like a wooden structure, can you give me like the most optimal layout, rational layout, and what if you compare that to concrete slabs? What if you show me the embodied carbon, or how much open and close should I do my windows in order to meet Net Zero building and to still have sufficient daylight? If you can ask those things at one point to AI, it's going to greatly increase.

Petru : Now we cannot do that.

Olaf : No, no, it requires too much specific information on regulation and level and also like cultural level. Now it's too complicated at this point, so one day. So at this point, it's really in the preliminary design. But that will change.

Petru : It's like a curve that goes really high at the beginning, then it's going all the way down through all the different architectural phases.

Olaf : Yes, indeed.

Petru : Do you think AI will primarily serve to enhance efficiency in architectural practices, or do you see it as a tool that could foster new forms of creativity and design ?

Olaf : I think mostly efficiency. I do believe that we shouldn't rely that much on the creativity of AI because it's based on historical data. So I would like to keep the creativity spirit free and try to use that one as much as possible from yourself or in a team environment, but for efficiency, yes, like either making fast renders going with an optimum structural grid. So it's also about efficiency, how to produce stuff faster, but while keeping the quality in there. So don't lose grip on your quality. Maybe even enhance quality to find those optimizations, which you normally do now with the parametric design, with the parametric model, I think AI can help you there. So in that sense, mostly on efficiency.

Petru : What AI tools are you currently utilizing, and how do you assess their capabilities and limitations in practical terms?

Olaf : So for me, I played a bit around with Comfy UI, also the previous one, the Automatic 1111, very limited. Comfy UI, is the one that I'm most comfortable with and used a lot, what I would say, the limitations there are, the limitations are basically the checkpoints, the models and the Loras, because everything is very nuanced. I think the checkpoints are a bit more flexible, but also the ways how I like to use this, very specific architecture and see if I can make an AI render of that. I don't think it's working that good yet. If you simply have a box, I think it works perfectly. It can give you the most amazing renders. But as soon as your project becomes very complicated, it has a lot of geometry, maybe a lot of, like, not regular shapes that you normally see in the built environment, it doesn't give you the good results.

Petru : For instance if you already have a facade with balconies, closed parts, open parts, and it's hard to change something at that, right? Maybe you just want to change a bit of design with AI quickly, but it's gonna be hard to do that for the limitations.

Olaf : I think we, as architects, we can work quite well with it. Because I think going from a model to image, I think that works pretty well, but like, halfway during the process to make those changes that you are also saying it's quite limiting, so maybe it's also, I think it's all about prompting. But I think we're not good in prompting yet. And I think we are really a bit like early adopters. I think we understand that game a bit better. But for people that don't have a clue, they also don't get anything nice out of it, because they don't know how to prompt. So that's a bit like a limiting factor. MidJourney I've never touched, to be honest, I was not a fan of using, you need to have discord or something. To be honest, I like going from my model to images, not so much from text to image.

Petru : You have more control in ComfyUI. The nice part of ComfyUI, it's also, it's free. You can access everything for free. And it's amazing.

Olaf : Also you could just send an image to someone, someone put that image in there, and you have the whole workflow.

Petru : What do you believe will be the most significant paradigm shifts in architecture due to AI over the coming decade ? Do you foresee the emergence of new roles or a restructuring of the architectural process as a result ?

Olaf : I think in 10 years from now, the amount of work that 10 architects do at this point in 10 years, one person can do that, and one, and one person with good skills in using AI to do their work. People really need to reconsider what they want to focus on. And you either should focus really on a niche, for example, what I'm doing, the sustainability will always be there. But for the normal architects, most of them will disappear, I'm afraid, because with AI, we can simply work that much faster. I think what I also said a bit in the beginning, the architect is more like the one that dictates the overall design process. However, it's simply like the middleman that makes sure how we cover our structural part, we cover the map, we cover the wishes that the client has. So it's really more about the process. And then you can use all the AI stuff to produce stuff. I think in 10 years time, one guy can do the work of 10 architects at this moment.

Petru : Looking ahead, do you believe AI will enable more personalized and responsive architectural designs, or will it standardize and limit creative expression? Could it lead to new methods of communication between architects, clients, and the public ?

Olaf : I'm still wondering, since AI is based on historical data, how does it go hand in hand with pure creativity? Maybe there is a way, and maybe they can really strengthen each other. But I think it really lies also then, within the prompting. On the other hand, what I do understand is that, especially for the public, let's say, because if you design an airport or a station within the city, and then one of the stakeholders is also the people in the city, right? Because they will use the building. I think there you can go to a way more interactive discussion with the user, what would you imagine? How is your ideal station? And even for people that don't know a lot about architecture, at a certain point where they have tools to simply say : Okay, I want this and this and this, I want to have this kind of feeling, and before you know it, they can also start to produce renders or come up with certain design arguments. So I do think that AI will give more hands towards the clients and public to also interact into the discussion that you have in order to establish somewhere, because now it's simply : 'What do you want ?'. They don't have the tools to say something, right? From the architects perspective, I find it hard to see how it can foster creativity. On the other hand, from the public and the clients, the other stakeholders, they have way more tools to engage in that discussion, which is a good thing, I think.

UnStudio Answering

According to Enes, AI can improve efficiency and creativity, at the moment in the early stages of a project, where it is more conceptual and preliminary, but today there are limits because architecture requires a lot of precision.

Today we are not yet at a level where we can only work with ai, but the world of architecture needs to embrace this innovation.

Enes Sever

Architect at UnStudio



Petru : How do you perceive the current landscape of AI in architecture? Is it truly transformative, or is it being overhyped in some respects ?

Enes : I think some people are over hyping it in a way, how they're using it, because they still don't know how the AI can be transformative. But there's also some other community, they think that is quite transformative, and I feel I belong in the part of that. I really believe that it can create some kind of transformation in architecture fields. It can actually give us some new guidelines, something that we cannot actually think, or something, that we cannot actually be that efficient, since, as we know that in architecture, it's quite, let's say, critical to be on time, and since we don't have so much resources in different offices, I think it can really elevate that design process and make it much more efficient on points and with clear direction. It's definitely a bit more transformative, rather than overhyped.

There are people who are a bit more trying to gain, let's say, they want to put themselves in some kind of labeling, AI expertise, AI architect, AI and designer, I think that is still not there. I don't think there should be that kind of labels yet, because we don't know yet how to use it, we have these tools, we can do renders in architecture, we can do nice inspirations, but those inspiration we can also do before, you know, you can also do find something in Pinterest. I think it will come transformative as soon as we know how exactly we want to use the tools.

Petru : What do you see as the most critical research areas or questions regarding Generative AI in architec-ture today ?

Enes : I think one of the critical thing that, especially myself I'm being critical to things that I generate, is, like a catalog of our in architecture, let's say, 50 options of a facade design, and then I have to choose one. You know, that's something that is a bit for me, because those models are trained from different images so they're not trained specifically for architects. The logic behind it is not working for how the steps are working. How can I actually? Is it realistic, or is it not? It can be nice inspiration, so even the implementing design can be something a bit more clear. But I definitely think that we still need more precise models that will serve for architects.

Petru : I think we still have to understand and work on that, for instance now we can use FormFinder as an integration during our studies, and is the nearest thing that we have as an architectural model.

Enes : It's working, but it's still a bit more rendering or a bit more specific things. And we need now to think, how can we actually approach to train the data sets so we can actually have results, or see the backgrounds a bit more, and take one step back, because now we know what this tools can do. Now we know that they can achieve some really nice results with a very short time with just prompting. But now we need to see one step back and see how this models are created, and maybe how can we actually be a bit more specific and a bit more realistic and something more rather than catalog, a bit more transformative in a way of also modeling experience.

Petru : Do you see AI as a continuation of technological evolution in architecture, or does it represent a completely new paradigm ?

Enes : In my opinion, I feel that this is a bit more evolution rather than new paradigm, because we still need our knowledge. When 3d modeling came up, or AutoCAD came up everyone thought that's okay, it's done, architecture is going to be very easy, but actually it gives us more challenges, gives us more opportunities to do something else. And I think everyone have a similar approach, that AI is more like evolution. Now everything can be very quick and efficient. So I believe that we have this kind of power now to elevate that fields and do a bit more different things, and because we have been so much in this transition with the modeling, let's say parametric architecture, and every time there's something new coming up, there's also a new chapter for us. This creates also for a new generation to think a bit more differently and a bit more smart and bit more efficient, I think that's quite important.

Petru : Now, we are in the beginning because we are trying to understand how the prompts are working, and it's quite hard to understand how to do a proper prompt for example.

Enes : That's very important part. Also, models are not there, since I'm using all these models from a long time and until now, only Dall-E can actually be a bit more coherent with the prompts and that it was introduced almost two years ago. So even after two years, we still don't have a model that can actually combine all this text from the prompts and give you good results. We still need to achieve a bit more, to wait for the models, to be a bit more advanced. And as you said, we don't still know how to prompt, we still don't know what is the subject, what do we want to achieve.

Petru : What are the potential and limitations of AI in generating design ideas? Can AI ever be a true partner in the creative process ?

Enes : I think the potential is, as far as I see, enormous. I'm not saying now, but in future, definitely can elevate a lot of things. And as a designers, we really struggle a lot, we struggle to convince the client, we struggle to convince our directors, our managers, especially when you come in a bit more down level or hierarchy. And it can give us hard time to model things, render and then send to your client, and that takes quite a lot of time. I think we can definitely shrink with this process, if we can have the good workflow that actually can give us more time to think about design rather than do the hard work of modeling things. Mostly of the limitations are when AI generates something, is it something that I want, or it's something that I think this looks nice. ? So I think we should definitely be a bit more critical in these things, because AI can generate really nice images. This was happening also with Pinterest, that you can see really nice references in projects, and then in your patch of projects it was not looking good, because it's something completely different. As I said before, AI in architecture still needs a bit more training data sets and a bit more focus and shapes and space. It's definitely not a tool for me, it's more a collaborative person that I can actually collaborate, I can talk, I can change my prompts, I can change my reference image, and that's actually what elevates your creativity in even further, because creativity is not just thinking out of the box, but also creativity is to link all these dots that you have in your mind, say that you found a nice reference in one of the buildings, but you want to mix with architecture in Netherlands, but you want also this to be in another way, so that's the link, it already gives you some kind of freedom to visualize it without need of model it, it can be a good creative tool for architects.

Petru : How do you see AI being integrated across different project phases such as Preliminary Design (PD), Concept Design (CD), and Schematic Design (SD) ?

Enes : From my experience, AI is definitely a bit more conceptual. When you have a design already in your mind, is easy to adapt and design to in the future, in the further phases. Also what we have done in this office as well, we were focusing on a concept phase, and we saw that the potential is there. When we try to apply this in a bit more design development phases, it doesn't work. It doesn't work because, you know, you don't want to also mess with different things, you already have the design, you already have thoughts before, why do you have an extra addition to your creativity, it's already done, so you need to try to fix that design, instead of giving you new ideas. So for now, it has very big impact in the concept and a preliminary design phase, also for bringing ideas, for example, we had this idea : we have this industry building in one of the coolest projects we're working and we want to transform this to a skate park, and we were searching in google 'skate park industry building', how does it look like ? There's no one designed, and then we came with this idea : let's try to visualize it, how can it actually look? How can AI can generate that? And it gives you idea, you know, like it doesn't have to be perfectly, but just to vision it, it already helps you to go for the design.

Petru : Do you think AI will primarily serve to enhance efficiency in architectural practices, or do you see it as a tool that could foster new forms of creativity and design ?

Enes : I agree in both ways. I think in efficiency wise, it can definitely elevate our process, I'm not talking just for inspiration, but I'm talking also about the graphical drawings can be elevation or plan that we are trying to use, you know, we spend a lot of time on this kind of things, or renders as well, one of the methods that we are using here in this office is also like enhancing our renders so in that case, will be definitely a big push for us. I'm spending now instead of 10 hours, I spent now one hour with AI, what do I do with nine hours after ? So that we need to be smart enough to allocate the time accordingly, instead of like doing over time now, we can actually, as an artist, have a proper work and life balance. We as much as we are passionate about it, we have also our lives. We play a big role in the future, and we as designers are now in charge to actually take care of this, let's say that if I can finish these things for eight hours, I don't need to do overtime to do other things, and it's enough to me for producing this kind of things. So in that case, I believe that it will be more efficient, we're gonna take a deep breath and gonna avoid this long hours and long design process but I also believe that it can also form a new creativity, but that will come, of course, with people using the tools more often, people seeing how can this be implemented in their different design workflows, in their different approach of the creativity. So it will come as soon as people are bit a more emerged with these tools.

Petru : What AI tools are you currently utilizing, and how do you assess their capabilities and limitations in practical terms?

Enes : Currently in our office, we are trying to be a bit in front with AI, and we are using our own tools. We are also now using ComfyUI, as you know, you have a lot of potential there, maybe not now but I think that people are still not aware how powerful that can be, because also the models are not very for architecture, it's more difficult. I was in a conference for generative AI, people were showing ideas of design, of graphics, web design, all those kind of things. Those fields are very liberal, they can actually produce so many things and so many ideas. But when it comes to architecture, you have limitations there, you need to be precise, so I see the limitations there, AI still doesn't understand much more. That is something a bit limited, but I know that the capability is there, especially for ComfyUI, I think all the designs are now going in that direction. There's this tool, and how can we use it ? You know, they're still not very sure how to actually imply all these kind of tools, and how can we actually use it. But a potential is there, we can do a lot of things. We can actually create this super workflow, get one prompt, generate 10 options, and from this 10 options, I can generate 10 videos, and from the 10 videos, I can generate models, I can offer you this, maybe not perfectly right now, maybe the models are not very polished for that, but I believe in the future it's going to be there, when the computers are also going to be a bit more stronger, when the graphics are going to get stronger I believe that in two minutes you can generate 10 option, 10 videos and 10 video models with this tool. And that's why people are also excited. In MidJourney, I also see a lot of potential, is definitely one step ahead for inspiration, considering to other tools, it definitely can actually transfer your thing that you have in your mind, to the visual and with the prompts as well. I really like how much I can control MidJourney through inspiration, but the limitation is that you cannot work from shape, it is just text to image.

Petru :What do you believe will be the most significant paradigm shifts in architecture due to AI over the coming decade ? Do you foresee the emergence of new roles or a restructuring of the architectural process as a result ?

Enes : I think we already are seeing these shifts, that people are a bit more now trying to learn this tool, that's something going because we know that with ChatGPT and all other tools, it's already going big shifts, people are already going crazy, especially the computer science coders, already had that big shift party, they already switched everything to there, you see now that maybe 80% of the people who code use ChatGPT for their own because they see that, how efficient it can be, and in architecture, we are still in not that shift, and mostly because it's more complex. We, as UnStudio, we have our own design language, and we don't want to break that language, and that shift will take some time to become, it should come also from the bigger hierarchy level of people, they need to also understand that it can be also beneficial, if office wills or any practice wills, to train their own models with their own images, and then they see that it can have good results, it can already shift big. Now I think that we need, as architects, to think what kind of role we can actually manage to create, so we can actually have a bit better to control. And I think this role should come from architects itself, someone who has visual knowledge, who has esthetic knowledge, who knows about architecture, should be the ones who would be in charge of this, and not people who are in computer science or technology fields, it should definitely come from designer artists or creatives, because they know what we need exactly.

Petru : Looking ahead, do you believe AI will enable more personalized and responsive archi-tectural designs, or will it standardize and limit creative expression? Could it lead to new methods of communication between archi-tects, clients, and the public ?

Enes : Yes, I think we already said this, especially with the communication. We already see that some clients don't want us to use AI, I've heard also that, for example, in some competitions, they are strictly about the use of AI because they feel like is very standardized.

Petru : But it's more for online competition, students competition. What about competitions that we are doing in the architectural offices ? We don't care if we are using AI for our design, it is right ?

Enes : Yes, that's the thing, but some clients are bit tricky, and sometimes they don't want AI to generate your inspiration, they want you to be in charge of this. We also are in charge to teach that, if you're going to shift in this kind of process, because we cannot stop this, people are already using it, it's already out and already giving good results, sooner or later, everyone will use it. We, as architects, need to also teach the client and show them that actually we can achieve good results with AI, we can control these things, that there's not something just randomly generated. Then I think they would also be a bit more in ease, and they can understand. Otherwise I would also not placing my chair in as a client, I would also be skeptical, is this just anyone can type this prompt and so maybe everyone can have this? But if I customize that, if I personalize that, if I can make something very specific that I want to do, and I design, it will definitely be more open mind also for them.

Petru : For instance, teach to the client that with AI you can achieve good results, I feel that now we are trying to reach the level where the client cannot understand if it is AI or not, on the other hand, maybe we can teach the client that you can control a part of the process and achieve good and maybe completely new results.

Enes : Now the results are so realistic that you can not even know if it's fake or not, so if you actually say it : 'okay, this AI', it's not that I just did the random prompt, I had to spend some time, because it's not easy to generate good images, you need some time also to control it, you need some time to generate it, you need to make iterations and all these kind of things, it becomes something that you're actually controlling it, and not just some random things. So as soon as you can actually state this to the client, then I think they will be a bit more open.

UnStudio Answering

*According to Cristobal, ai today is very overrated being used for much simpler tasks than its full potential, but at the moment we are not able to exploit the potential. Tools such as MidJourney are extremely simple to use and therefore lead to a common result for all, while ComfyUI allows for maximum customisation and learning as well as new challenges and opportunities.
In the future, the most likely role to emerge from ai is the prompt engineer, but perhaps this could be more of a skill than an actual role.*

Cristobal Burgos

Senior Digital Developer



Petru : [How do you perceive the current landscape of AI in architecture? Is it truly transformative, or is it be-ing overhyped in some respects ?](#)

Cristobal : I think, as most things with AI are, is being heavily overhyped, which is a bit of detrimental for the industry itself. I think it's been presented as the One Shot solution for everything, while it's not, in my opinion, and I think architects are just using it, are not using it to the full potential, they're mostly using it for generating images and things that might not be very heavily important for them, but they are just using it because they can do it. And it feels like an easy piece of technology to use, while it might not bring out of value. And my biggest concern with this is that it can really unfocus you, you can get really lost of what should be your north because you can do things, and I think that's problematic sometimes.

Petru : [But for instance, now we can just generate images, it's more complex to do other things with AI during the architectural process.](#)

Cristobal : I would say so, for example, the biggest challenge that we are doing now is trying to generate floor plans, because that requires a lot of information, very complex based data that you need to deliver into the model, and that's a lot more heavy computational and it's a lot more challenging to do, especially because architecture unfortunately sits in a very niche situation, doing developments for architects is very expensive, but also architects usually don't have the money to pay for them. So there's a lot of resource scarcity to develop those things.

Petru : [What do you see as the most critical research areas or questions regarding Generative AI in architecture today ?](#)

Cristobal : For me, it's about data. I always remember that my first company, they used to say : 'we are poor because we don't know any assets', but our biggest asset is our project. I think architects have a passion in the good sense of the word for reinventing the wheel and coming back to the same problem again and again and again with a different approach, which it's nice for some things, but for others it's not, and I think AI can really help you engaging with the data that it was a bit hidden before, and it was hard to understand, hard to process, because for me, AI just reduces the cost of managing those kind of things, because you can automatize several process of data management, data flows, and while it's not flawless, it really helps you bringing down the barrier to develop those things. So that's the way that I see AI working a bit more and at least for larger enterprises like ours, that we are 400 people around seven offices across five different continents.

Petru : Do you see AI as a continuation of technological evolution in architecture, or does it represent a completely new paradigm ?

Cristobal : It's for sure a continuation, like the logical next step, because by the end of the day is nothing new. All of these has been since the 80s, I think the first neural network is from the 50s or 60s. Yeah. I think that it was more of a cultural shift, when open AI came with chatgpt, that's where things really started moving, just because it felt very accessible, and things that before it took you a year to develop, now you could do it in a day. I still remember the first time that I trained an AI model, that was 2016, it took me a week of training and the results were not even that great, so now I can just train a model in a couple of hours that will get the results I need.

Petru : What are the potential and limitations of AI in generating design ideas? Can AI ever be a true partner in the creative process ?

Cristobal : I think yes, but not in the creative part of it. So far, I haven't seen anything that would tell me that AI can create it, basically, AI is just like a very good librarian that you can ask for things and get them back. Also, AI is like an umbrella thing for everything, when you are doing a model of linear regressions to calculate the value of a number, it's also AI and LLM is also AI, but you will not put the same things in the same box, but I would say yes, they can be a sparing thing, especially because it allows you to focus on what matters more, now you don't have to spend 90 hours thinking about the render and how the light looks, you can just like, smooth the process with that, and you can still focus on what I want to do with this render.

Petru : How do you see AI being integrated across different project phases such as Preliminary Design (PD), Concept Design (CD), and Schematic Design (SD) ?

Cristobal : I feel like for concept and schematic is all about the images, the conception, the idea, the quick thing, the quick testing of ideas, basically, and then the more you go into later phases of projects, I think it's when it start getting related back data. So you can manage the data by itself, you can check things, you can build agents that kind of check if you're doing it's correct or not, if you're complying with regulation and stuff like that. It boils down from a very idea based thing and concept testing to a very data driven and heavy and precise numbering.

Petru : But for instance, we are still not there with the data, with the complex data, because we cannot use it in complex phases, we still do that manually. We are not integrating AI complex architectural phases. Also architects in general, I don't know if someone else is doing that.

Cristobal : So far, I haven't seen it use it in a way that I've been out. I feel like there's companies that are doing it, but not in like, something that I will say : 'this is the way to go'.

Petru : I would say that we are not there yet.

Cristobal : Especially because that comes always with the question about liability in the early phases, an error is not as expensive as later, if you miscalculated one square meter of area probably that's gonna have a big impact into delaying, but it's not as critical as once you have everything built already, and it was like a meter bigger than needed. Those are decisions that start happening at the latest stages of projects, that's when you need to kind of accept the fact that you need to take responsibility for that, and I think that's why they haven't pushed for the developments of that still.

Petru : Do you think AI will primarily serve to enhance efficiency in architectural practices, or do you see it as a tool that could foster new forms of creativity and design ?

Cristobal : I think yes, I'd like to see things as it's just another tool in your tool set. As an architect, I do think it allows you to do things different than you were doing before, especially with these rendering things that you can test for different modes in a couple of seconds instead of a couple of hours, and it allows to smaller teams to work as they were bigger, because basically, you can produce more output with not more input. I have failed to see new ways of doing architecture so far, but I also think that's because of this nativist conversation of : 'I'm gonna be the first native AI architect'. I feel like it's a bit disingenuous to say that, because I don't think the tech is there, I'm not sure about the techniques to be there.

Petru : What AI tools are you currently utilizing, and how do you assess their capabilities and limitations in practical terms ? And what are your thoughts on tools like MidJourney and ComfyUI ? Have they changed the way you approach architecture ?

Cristobal : Well, we use a lot of LLMs, not because of the reasoning capability, because they don't have to, but we use them more of a like the gate for a couple of things and to process information. Obviously, we are developing tools like a chat GPT clone with our internal data protected all of that, we are building our own MidJourney, but I also use a lot of models of math just to predict data, predict trends, to do forecast analysis and all that. We should quite a lot for that to make expectations on what should happen. We do a lot of economical business case and models, that's still not in production, because we want to be very rigorous with the names, with the numbers that we are given. But that's one of our biggest application. It's more we do the MidJourney for more the architects, so you can just iterate through ideas faster. MidJourney comes with a very nice set of tools that really makes it easy to do, they also come with a trade off, that it's so easy that everyone is doing the same thing. I think MidJourney is not exciting because everyone's doing exactly the same thing, and unless you can have a very specific set of ideas that you can develop, it's gonna look exactly as everyone else's images. ComfyUI allows you to be more specific, do more new things, your own custom things. That's why, for example, you see startups working on top of ComfyUI to use them as a back end to do their whole business. And can be quite unique, obviously, that comes with a trade off that is a lot more challenging to use, but doing that a lot more is what differentiates you from the rest.

Petru : MidJourney is user friendly, so it is super accessible, for ComfyUI was already hard just to install it and understand the basic things.

Cristobal : I think they're moving towards like an easier model system of doing things, but it will always have a lot of technical challenges, but I think that's a good thing, doing the complex parts, because that really allows you to understand what are you actually doing. You're not just not using tokens to generate more tokens in the way of like a JPEG, you know? So you're actually doing something and putting your process.

Petru : What do you believe will be the most significant paradigm shifts in architecture due to AI over the coming decade ? Do you foresee the emergence of new roles or a restructuring of the architectural process as a result ?

Cristobal : I don't want to say prompt engineers, because I really despise the concept, but the people that have more expertise into prompting, I think they're gonna have a bit of an advantage, especially because prompting is more about how you set the things in your mind to understand all the requirements. Prompting is not something as trivial as just like knowing where to copy and paste, it's also about having a mental structure and how to process the data in a way that the computer can understand it very quickly. The problem is that the prompting is a lot more open, a lot of the people complain : 'the results are not that great', it's because it's not a machine problem, it's your problem. It's just another thing in your toolkit, personally, I don't think that it will change fundamentally on the short term. I don't think like you're gonna have mass layoffs. You see those layoffs in architecture for more economical reasons.

Petru : Looking ahead, do you believe AI will enable more personalized and responsive archi-tectural designs, or will it standardize and limit creative expression? Could it lead to new methods of communication between archi-tects, clients, and the public ?

Cristobal : Yes, especially because now you will be able to engage with more data, and that will force a way of communicating that data in a new way, in a way that we're not doing as it is now. I also think we're just like circling back to the same things that we were doing 10 years ago, 20 years ago, just in a different rubber, which is nice, but I also feel like it comes back to the first question about, is it being overhyped? Yes, I think it's being over I agree with that. So we're doing something cool and new, but at the same time, it's not as all the innovations. I think it opens new doors and it lets people that really likes to innovate out that extra new thing that we could not do before. Sometimes those are the smaller, invisible things that you didn't saw, but you had to do it anyways.

In my mind it always helps you avoiding the blank page, that's like the biggest advantage for me, at least, because now when I answer an email, I don't have to start from scratch for example, I can just use AI to process a bit what I'm thinking, because it helps me organize and move faster with my own thoughts.