ICONTECH INTERNATIONAL JOURNAL OF SURVEYS, ENGINEERING, TECHNOLOGY
ISSN 2717-7270
Journal homepage: http://icontechjournal.com/index.php/iij
Volume 7 (2023) Issue 3

Generative Text-to-Image Models in Architectural Design: A Study on Relationship of Language, Architectural Quality and Creativity

Emel Cantürk Akyıldız

Assist. Prof. Dr., Kocaeli University, Faculty of Architecture and Design, Kovaeli, Türkiye

Abstarct

Text-guided generation of images with deep learning technology has made significant advances and has seen an increasing interest since 2021. With these mostly web-based models, users can synthesise photorealistic and high-quality digital images from natural language descriptions with no or little understanding of the underlying technology. Although these AI technologies are in the early phases, there is already an explosion in AI-generated architectural activity. While generative AI technologies propose a new design method for designers and architects, it will undoubtedly redefine the skills, knowledge and competencies that designers should equipped with. This research focuses on understanding the "artificial intelligence – architect" interaction as a design method, specifically the "language as a design driver", and interrogates the role of the designer in AI-driven design. In the context of the research, the textual inputs ("prompts") and the outputs of the architectural design studies of 36 subjects generated in Midjourney – a text-to-image latent diffusion model – were analysed in terms of the possible relationships between the language of the prompts, (1) prompt length, (2) descriptive language, (3) specific architecture-related indicators, and the quality of the outputs in two terms of architectural quality and architectural creativity.

Keywords: Artificial Intelligence, architectural design, text-to-image generation models, design method, architectural quality, architectural creativity

1. Introduction

The relationship between design thinking and the new tools of design, representation and production for architectural design, which emerged due to evolving technologies, has been an important subject of interest and discussion for theorists and designers from different disciplines. Schön's (1987) theory of "reflective practice" is one of the most influential and widely accepted theories in architectural design (Webster, 2008). Schön (1987) sees design as a reflective process that involves the constant iteration of ideas and the generation of new knowledge through dialogue between the designer and the design context. As a part of the design context, design tools shape the design dialogue. Clark and Chalmers (2003) approach design tools and technologies in the context of the "extended mind" theory, arguing that the mind is not limited to the brain but can extend beyond the body and into the physical world. They propose that some objects in the physical world, including technology, can be used to become part of the mind itself and to increase the cognitive abilities of the mind. In this context, the tools of design and representation in architectural design thinking can be seen as a way for the designer to develop the ability to perceive and reason.

Focusing on the relationship between technology and architecture and how technology has transformed architectural design throughout history, Carpo (2017) emphasises the importance of understanding the historical context of technology and its impact on architectural design, as well as the potential for new technologies to transform architectural practice in the future. For

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023;
Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

example, digital design tools and digital fabrication technologies have enabled architects to create complex forms and structures with greater precision and efficiency than ever before.

Today, we are witnessing another technological revolution, the "deep learning revolution", which enables designers to generate endless design alternatives with text-to-image generation with deep learning through a mostly web-based, easily accessible ecosystem of generative models such as Imagen, DALL-E, DeepDream, Midjourney, Stable Diffusion, and so on. This implies a milestone in human-machine interaction. For the first time, users can generate high-quality images without any or little knowledge or understanding of the underlying technology. These technologies also do not only imply new mediums and tools for design but, more importantly, new methods and understanding of design. They "call into question our design methodologies, understanding of our culture and interactions with the world around us, among other things" (Bolojan, 2022: 24). As generative models become more widespread, questions about originality, authorship and creativity arise, and how humans interface with AI becomes an important research area (Reynolds & McDonell, 2021; Dang et al., 2022; Deckers et al., 2023; Oppenlaender et al., 2023). If anybody with a computer or even a smartphone can generate digital images with high aesthetic qualities, can we talk about the creativity of the designer anymore? What is the role of the designer in an AI-generated design world? How can we employ these technologies as a new design method? In light of these questions, the research focuses on understanding language as a design driver in AI-generated architectural design and how the designers can interact with the computer to enhance their creativity. The aim of the research is to understand the relationship between textual inputs (prompts) and outputs (synthesised digital images) via the following research questions:

- What determines the quality and creativity of the AI-generated outputs?
- How can we assess the AI-generated outputs?
- What is the relationship between language and architectural quality/creativity of the output images?

2. Artificial Intelligence, Deep Learning Systems and Architecture

Artificial Intelligence (AI), first proposed in 1956 by Dartmouth Summer Project, is a broad and evolving concept. The field's pioneers, McCarthy, Minsky, Rochester and Shannon (2006 [1955]: 12), defined the study of AI in their 1995 proposal for the Dartmouth Project as "to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it." Currently, AI is defined as a multidisciplinary field focused on developing systems and machines capable of performing tasks that typically require human intelligence, such as problem-solving, learning, reasoning, perception, language understanding, and decision-making. It basically refers to the simulation of human intelligence in computers and other machines.

Even though it has undergone two periods of decreased funding and interest throughout the 1960s and 1990s, the so-called "AI Winters" (Russell and Norvig, 2003; Crevier, 1993), artificial intelligence research has come a long way since 1956. AI research has witnessed the emergence of diverse methodologies aimed at replicating human cognitive abilities. Two prominent concepts within this realm are "expert systems" and "learning systems." After the first AI Winter, the 1970s responded to a revival in the field with the emergence of expert systems. Expert systems, known as knowledge-based systems, became very popular in this

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

period. These models enable machines to reason based on a set of rules and collections of data (Chaillou, 2021: 21). These systems use a knowledge base that contains a vast amount of expert knowledge and rules. The rules are typically represented in the form of "if-then" statements, where the system can make inferences and decisions based on the input provided. On the other hand, learning systems are built on algorithms and protocols that allow machines to learn from repeated observations after iterative exposure to vast amounts of data and improve their performance with time. These systems leverage machine learning techniques to identify patterns, relationships, and trends in large datasets.

Throughout the 1990s and 2000s, AI research shifted toward machine learning-based methodologies. The limitations of expert systems prompted new explorations such as neural networks, Bayesian networks, evolutionary algorithms, etc. (Chaillou, 2022: 25). The rise of learning systems marked a fundamental shift in AI's approach from codifying human expertise into rule-based systems to harnessing the power of data-driven learning. This is one of the most profound shifts in computer science and architecture that we are witnessing today (Bolojan, 2022: 24). The main difference lies in their approach and purpose: expert systems (early versions of AI) use predefined knowledge to make decisions within a specific domain, while learning systems (recent versions of AI) utilize data-driven approaches to improve their performance over time, making them more adaptable to changing conditions and broader domains.

The 2010s ushered in a period defined as the "deep learning revolution" in which advances in artificial intelligence accelerated. Deep learning refers to the constant change taking place within the AI community, meaning that "artificial neural networks" (ANN) have taken over from expert systems or other architectures as the primary focus of AI research (Chaillou, 2022: 25). Learning systems leverage the vast amounts of data available in the digital age to enable machines to learn and adapt autonomously. Deep learning algorithms, such as neural networks, enable systems to identify patterns, relationships, and trends within data and have redefined the landscape of tasks like image recognition, natural language processing, and speech recognition. Neural networks draw inspiration from both the architecture and operational principles of the human brain and how humans gain specific knowledge. Instead of relying on predefined solutions, these networks learn predominantly through examples (Bolojan, 2022: 24). While the human brain learns to identify semantic aspects of images through extensive learning from experiences, neural networks can identify the semantic aspect of images after learning from large amounts of labelled data (Bolojan, 2022). They extract features from images and progressively learn to associate these features with specific labels.

2.1. Image generation with deep learning: Generative-AI as a design method

As AI technologies continue to advance, they have found a compelling array of applications within the architectural domain. The first category is applications of AI in the optimisation of the building's energy consumption and performance, construction processes, and material consumption; and the second is the investigation of architectural design, including creativity and intuition, "which are difficult to translate into code" (del Campo & Leach, 2021: 7). Image generation with AI, a recent achievement within machine learning, falls into the second category, and has the potential to transform the practice of architectural design.

Image generation is a relatively recent achievement within machine learning. In 2014, Ian Goodfellow and his colleagues theorised generative adversarial networks (GANs), a class of machine learning models designed for generative tasks (Goodfellow, et al., 2014). These

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

models can synthesise extremely realistic images and videos. StyleGAN, an extension of the Generative Adversarial Network (GAN) architecture, has emerged as a powerful tool that revolutionizes the generation of images with unparalleled levels of realism and artistic finesse. In 2018, Karras and colleagues showed that StyleGAN could generate extremely realistic high-definition human faces (Karras, et al., 2018).

The early 2010s can be considered a significant turning point when breakthroughs in deep learning and neural networks led to remarkable progress in AI capabilities, leading to increased attention, investment, and innovation in the field. This ongoing period is called AI Boom or AI Spring (Bommasani, 2023), with an increasing interest in generative AI being the essential aspect of this boom, which began with the establishment of an American research laboratory, OpenAI, in 2015.

Although it is a relatively recent technology and in its early phases, image generation with deep learning models found usage in architectural practice already. “Architecture, as a discipline of form-making, often absorbs new ways of form-thinking from other domains as well as it adapts new technologies as tools for form-modelling” (Koh, 2022: 111). In this context, generative AI models are used as a design tool in the field of architecture at an early stage of the design process, which can be defined as the creativity stage. In particular, ZHA (Zaha Hadid Architects) and Coop Himmelblau's work on form modelling through image synthesis with machine learning is considered pioneering work in this field. It can be predicted that the applications of AI as a design tool will soon become widespread in the field of architectural design and have great potential to transform the design industry. Synthetic images generated by AI can stimulate the human mind to push the boundaries of architectural creativity. “Because they are based on existing information, they are familiar enough to be construed as architecture but strange enough to provoke us and challenge us as designers” (del Campo and Carlson, 2022: 179).

The current research Office of Coop Himmelb(l)au, on artificial intelligence “Deep Himmelblau”[1] focuses on using AI as a new tool within the whole design process and methodology. The machine-learning protocols and algorithms learn the semantic characteristics of the projects designed by the office over time, to generate new interpretations of the existing visual data and open up new possibilities. The office states its motivation: “DeepHimmelblau explores the possibility – in connection with human beings – of teaching machines to be creative, to interpret, perceive, propose new designs, augment design processes and augment design creativity” (dPrix et al., 2022: 16).

Similar to Clark’s ‘extended mind’ theory (1998), Crespo and McCormick (2002: 56) define AI as augmentation of the mind or “‘mental prosthetics’ that can be used as a means to expand our imagination of the mundane”, and del Campo and Leach (2022: 11), as “a muse an extension to the human imagination” to expand the designer’s cognitive abilities as Clark suggested.

2.2. Text-to-Image Generation: Language as a Design Driver

One of the latest advancements in generative AI technologies, at the interface of AI and linguistics, is the introduction of text-to-image generative models. Text-guided generation of images with deep learning technology has made significant advances and has seen an increasing interest since 2021 when OpenAI - an AI research laboratory founded in 2015 – released their

[1] For information about Deep Himmelblau Project: https://coop-himmelblau.at/method/deep-himmelblau/

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023;
Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

language models GPT-3[2] (Brown, et al., 2020) and DALL-E[3] (Ramesh et al., 2021). These models can translate textual information into potentially linked visual representations. In other words, they can synthesise photorealistic and high-quality digital images, using natural language descriptions – referred to as “prompts” – as inputs and by drawing data from the Internet to match images with captions. After DALL-E, a range of deep learning models for image generation have been introduced, such as diffusion models, which are far more successful than previous architectures in image synthesis (Oppenlaender et al., 2023; Dhariwal and Nichol, 2021). Mostly browser-based, generative text-to-image models such as DALL-E, Midjourney, Stable Diffusion, XKool (one designed specifically to synthesize architectural images), and more are already widely used by architects. There is a considerable increase in seminars, programs, workshops, and degrees to engage architects and designers of various disciplines with new AI technologies.

del Campo and Manninger (2022:45) discuss the motivation to explore attentional generative adversarial networks (AttnGAN) as a design technique in architecture. Traditionally, architectural design begins with visual inspiration, but AttnGAN – and other architectures of text-to-image generation – explores using “language as a starting point for design”. It translates written programmatic needs into visual representations that can be translated into 3D architectural models. The award-winning project of the 2020 competition entry for the 24 Highschool in Shenzhen highlighted AttnGAN's potential to function as a successful design technique for a complex architectural program. “This technique allows shape to be interrogated through language: an alternative design method that creates its own unique sensibility” (del Campo and Manninger, 2022: 45). According to them, following the initial phase of experimental projects, the subsequent phase involves introducing enhanced datasets tailored exclusively for architectural design. This will enhance the effectiveness of the presented design methods by incorporating a comprehensive AI-driven approach to architectural design. Integrating text-to-image generation models into architectural design implies a paradigm shift in how architects conceptualize, communicate, and iterate through design processes.

3. Material and Methods

The author conducted an online four-day workshop[4] that covers exploring the creative potential of generative AI models to drive architectural design explorations. Throughout the workshop, attendees were introduced to a range of state-of-the-art generative diffusion models, including Midjourney, Stable Diffusion and DALL-E, and design workflows that incorporate these generative models. The workshop was structured as a combination of informative sessions, tutorials, design sessions and design reviews. The attendees were asked to create and expand architectural concepts using mainly Midjourney and incorporate DALL-E, Chat GPT and Stable Diffusion into their design workflows. Before the workshop, attendees were asked to fill in a form, including demographic items, educational backgrounds, experience -and competency- with generative AI, experience in the profession, and a consent form explaining the study they were participating in.

3.1. Workshop Participants

[2] First GPT was introduced in 2018 by OpenAI and GPT-4, capable of accepting text or image inputs, was revealed on March 2023.
[3] DALL-E uses a version of GPT-3 modified to generate images.
4 Synthetic Architecture: AI-Generated Design Explorations workshop conducted in the scope of Digital Futures 2023 Workshops, between July 22-25. (https://digitalfutures.international/synthetic-architecture/)

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023;
Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

36 attendees participated in the workshop, 21 (%58.3) of them were professionals, and 15 (%41.7) were architecture students. Professionals included architects, 2 architectural visualisers, 1 industrial designer, 1PhD researcher and 3 academics. The ages of the attendees ranged between 20 and 60. %52.4 of the professionals had less than 5 years of experience in the profession, while %28.6 had 5-10 and %19 had 10-15 years. More than half of the attendees, %55.6, had no experience with text-to-image generation models, while %44.4 had some experience with Midjourney (12 attendees), Stable Diffusion (2 attendees), and DALL-E (2 attendees). However, the majority of the %44.4 assessed their competency low in using these models.

3.2. The Workshop

Through the workshop, attendees were introduced to Midjourney, DALL-E, Stable Diffusion and Control Net, and given four design challenges to perform. On the first day of the workshop, they were introduced to the basics of the AI generative models and "prompt crafting" -or called "prompt engineering", "prompting", and "prompt design" (Oppenlaender, 2022)- in Midjourney. In design **challenge #1,** they were asked to create an architectural concept and generate images by combining two controversial or unlikely concepts together. They were asked to avoid traditional combinations, edit their prompt until they find an output worth to study on it, and do as many iterations as possible. In design **challenge #2** they were expected to be more descriptive of their concept by adding descriptions of geometric characteristics, style, context, lighting, mood and atmosphere.

On day two, participants were introduced to advanced prompting[5], such as image prompts and multi prompts, and controlling the settings and parameters in Midjourney. In design **challenge #3,** attendees were asked to expand their architectural concepts by adding detailed descriptions of the materials, surfaces, visual style and composition/frame. And also using different parameters and settings to explore through a trial-and-error process. Design challenge #4 was about influencing the prompts with additional media such as sketches and images.

On the third day, participants were introduced to advanced prompting, such as image prompts and multi prompts, and controlling the settings and parameters in Midjourney. In design **challenge #4,** attendees were asked to expand their architectural concepts by adding detailed descriptions of the materials, surfaces, visual style and composition/frame. And also using different parameters and settings to explore through a trial-and-error process.

The last day was dedicated to using Stable Diffusion and Control Net, and expanding the architectural concepts by incorporating these models into the design workflows. In the final design challenge, **challenge #5,** attendees were expected to expand and refine at least 3 architectural concepts by using Midjourney, Stable Diffusion, Control Net and additional media.

3.3. Curation of the materials and assessment of the output images

For the dataset, 48 images were selected from the workshop outputs. 15 academician-architects from two universities were asked to score the images according to the following criteria: (a) the architectural quality of the space/building depicted in the image (b) the architectural creativity/novelty of the space/building depicted in the image on linear scales from 1 to 5 point. Also, extra explanations for the criteria were given. The voters were sent the images as Google Forms, and the images were shuffled to prevent the voters from having the impression that the

[5] https://docs.midjourney.com/docs

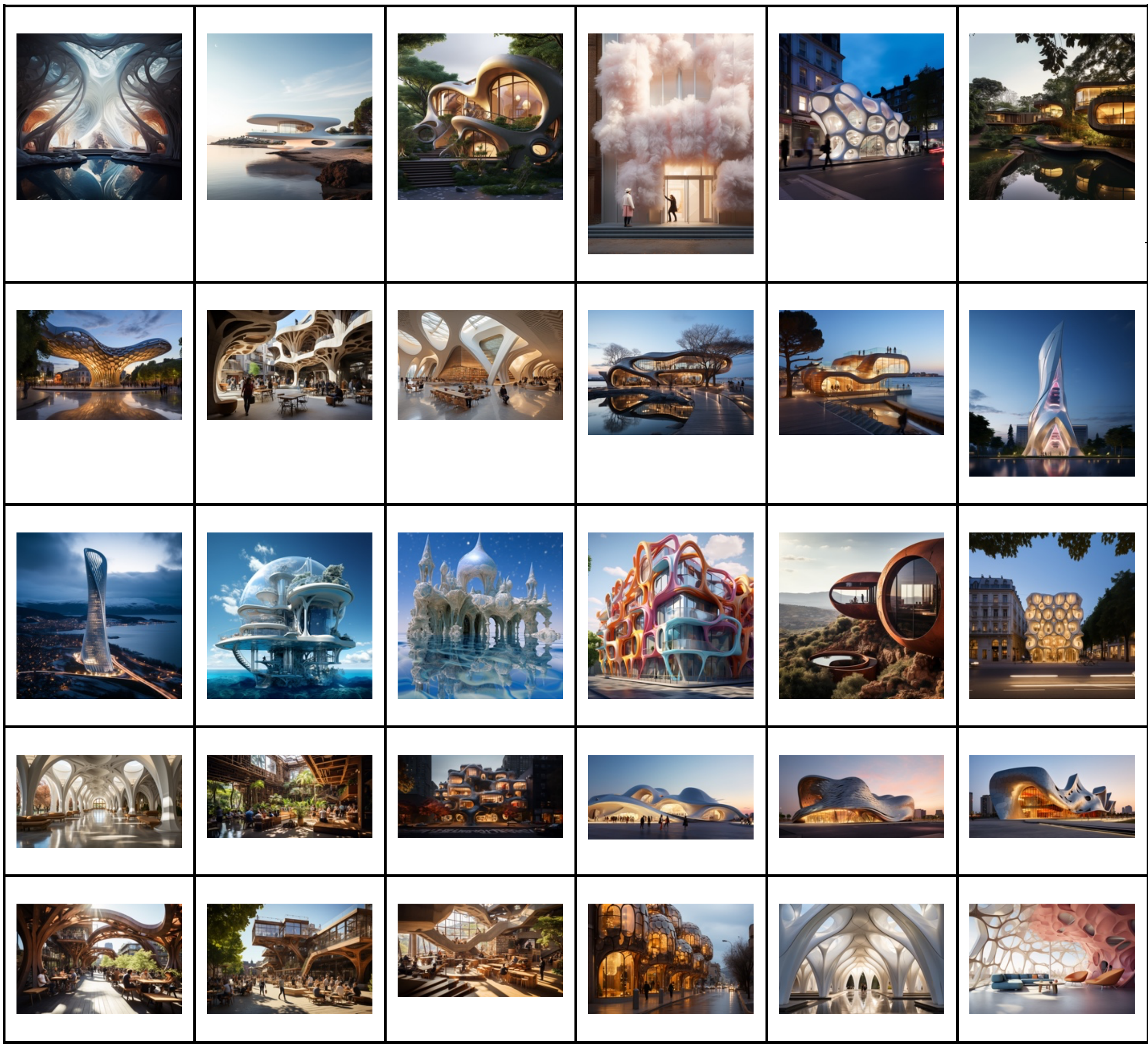
Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

ICONTECH INTERNATIONAL JOURNAL OF SURVEYS, ENGINEERING, TECHNOLOGY
ISSN 2717-7270
Journal homepage: http://icontechjournal.com/index.php/iij
Volume 7 (2023) Issue 3

images were sorted from the highest to lowest quality or vice versa. For consistency, check a set of very similar images placed in the form (Fig. 1).

3.4. Analysis

Scores for the 48 images in two categories from 15 voters were collected, and two average scores for the categories of architectural quality (AQ) and architectural creativity (AC) were calculated for each image (96 average scores in total). Since average values do not provide reliable data on their own, the standard deviation (SD) values were calculated for each average score, and the scores differentiated from the average score more than SD values were eliminated, and new average scores were obtained. For each assessment category (AQ and AC), the scores of the images were sorted from the highest to lowest and grouped into three groups of images with high, medium, and low scores. Then, each prompt was analysed in terms of four following aspects: (1) prompt length (number of words in the prompt excluding prepositions and parameters), (2) descriptive language of the prompt (number of the specific words vs. general statements) (3) architecture-related indicators in the prompt (a) style (b) volume and shape (c) material and surface (4) architect name. Then, in two categories, the relationship of these aspects to the success of the output is analysed by statistics.



Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

ICONTECH INTERNATIONAL JOURNAL OF SURVEYS, ENGINEERING, TECHNOLOGY
ISSN 2717-7270
Journal homepage: http://icontechjournal.com/index.php/iij
Volume 7 (2023) Issue 3

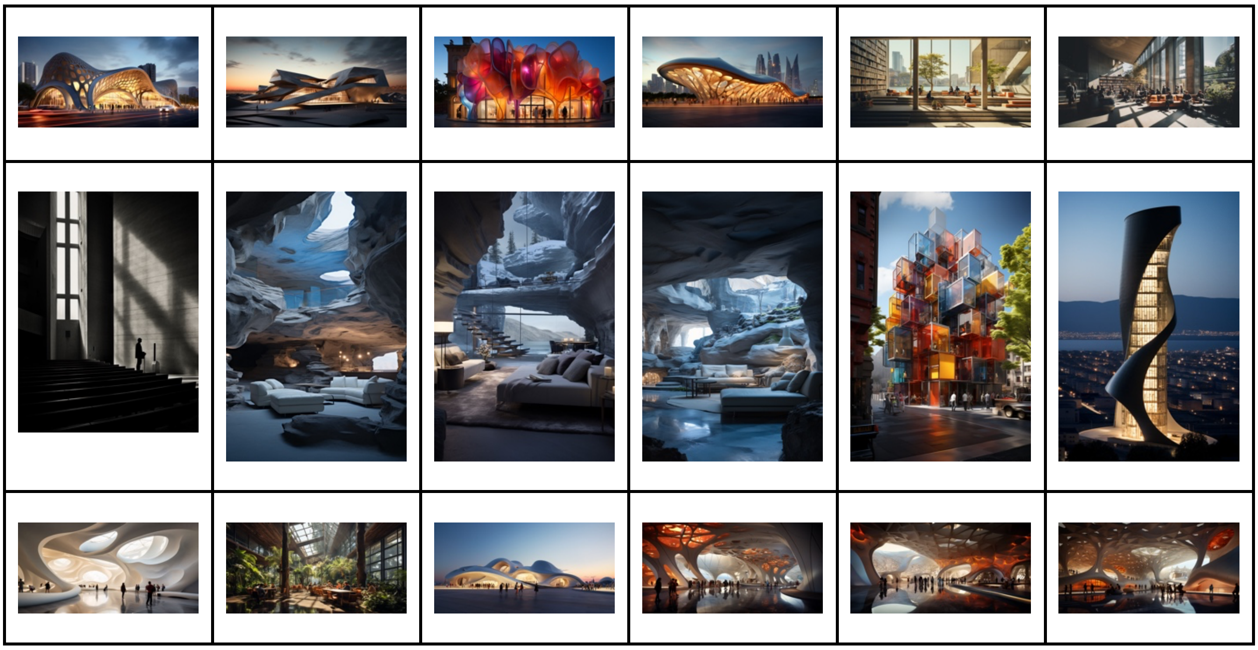


Figure 1. Dataset of the research, images selected from the outputs of the workshop

4. Findings and Results

4.1. Architectural quality

The voters were asked to score the architectural quality (AQ), with the highest score of 5 for the space/building depicted in the image with recognisable forms and patterns and the lowest score of 1 for unrecognisable, too complex and "alien" forms. The highest average score was 4.6, and the lowest was 2.3 (Fig. 2). According to the scores, images categorised as (1) images of high AQ: 4 and scores above, (2) images of medium AQ: scores between 3.5-3.9, and (3) images of low AQ: 3.4 and below scores. According to this categorisation, 13 images were identified as having high architectural quality (AQ), with the highest score of 4.6. Prompt lengths of the high AQ images ranged between 12 and 50 words, with an average of 26.76 and a median of 18. Prompt lengths (PL) of 24 images identified as medium AQ ranged between 7 and 56, with an average of 26.04 and a median of 18 words; low AQ images' PLs ranged between 2-18, with an average of 13.5 and a median of 13.5. Although, a significant difference in PL between high-medium scored images and low ones observed, such a difference was not observed between high and medium. The relationship between the prompt length and the architectural quality of the image was not clear (Table. 1).

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023;
Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

ICONTECH INTERNATIONAL JOURNAL OF SURVEYS, ENGINEERING, TECHNOLOGY
ISSN 2717-7270
Journal homepage: http://icontechjournal.com/index.php/iij
Volume 7 (2023) Issue 3

		
AQ score: 2.3 (low)	AQ score: 3.6 (medium)	AQ score: 4.6 (high)

Figure 2. Examples from images with the lowest, medium, and highest AQ scores

Regarding descriptive language (DL), high-quality images' prompts contained an average of 9.46 and a median of 9 descriptive words, while medium-quality images had an average of 7.85 and a median of 7 words. The low-quality ones contained an average of 4 and a median of 3 descriptive words. A meaningful difference was detected between the descriptive words used in high-medium and low-quality images, while the difference between the high and medium quality is slighter (Table 1).

Table 1. Relationship between AQ scores of the images and PL and DL

Scores for Architectural Quality	Number of images	Prompt length (PL) (No. of words)			Descriptive language (DL) (No of specific words)		
		Range	Average	Median	Range	Average	Median
High (4-4.6)	13	12 - 50	26.76	18	5 - 19	9.46	9
Medium (3.5-3.9)	24	7 - 56	26.04	18	0 - 18	7.87	7
Low (2.3-3.4)	11	2 - 18	13.5	13.5	0 - 17	4	3

The language in prompts was analysed in detail in terms of indicators of (1) architectural style (2) volume and shape (3) material and surface (4) the architect's name to test if any of these have a significant effect on the success of the image. The following observations were made: (a) Images with high AQ scores: 10/13 prompts had style indicators, 4/13 had volume and shape indicators, 9/13 had material and surface indicators, 9/13 had at least one architect name; (b) Images with medium AQ scores: 19/24 had style indicators, 12/24 had volume and shape indicators, 18/24 had material and surface indicators, 14/24 had at least one architect name; (c) Images with low AQ scores: 7/11 had style indicators, 2/11 had volume and shape indicators, 6/11 had material and surface indicators, 3/11 had at least one architect name. When the frequency difference between high AQ and low AQ images is compared, the frequency of indicators implies no effect on the quality, while the frequency difference between high-medium AQ images and low AQ images is increased (Table 2). As a result, although the input of indicators does not always lead to outputs with high architectural quality, the use of indicators is necessary to obtain an acceptable quality.

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

ICONTECH INTERNATIONAL JOURNAL OF SURVEYS, ENGINEERING, TECHNOLOGY
ISSN 2717-7270
Journal homepage: http://icontechjournal.com/index.php/iij
Volume 7 (2023) Issue 3

Table 2. Relationship between architecture-related indicators and AQ scores

Indicators in the prompt	High AQ		Medium AQ		Low AQ	
	13 prompts		24 prompts		11 prompts	
	No. of prompts	Percent.	No. of prompts	Percent.	No. of prompts	Percent.
Architectural style indicator	10	%76.9	19	%79.1	7	%63.6
Volume and shape indicator	4	%30.7	12	%50	2	%18.2
Material and surface indicator	9	%69.2	18	%75	6	%54.5
Architect's name	9	%69.2	14	%58.3	3	%27.3

4.2. Architectural creativity

Architectural creativity is a broad and loose concept with no consensus on its definition in architectural literature. For this research, the definition of architectural creativity was explained to the voters as the novelty of the forms, materials, and other architectural qualities, yet the 15 scores differentiated in high ranges, implicated the fuzziness of the concept. The highest average score for architectural creativity (AC) was 4, and the lowest was 2.3 (Fig 3). Regarding the average scores, images were categorised as (1) images of high AC: scores above 3.6, (2) images of medium AC: scores between 3-3.5, (3) images of low architectural quality: 3.9 and scores below. 12 images labelled as having high AC had an average of 25.08 and a median of 31 words prompt length, while the 24 images of medium AC had an average of 22.45 and a median of 15 words. Images of low AC had an average of 13.91 and a median of 14 words. The results displayed a meaningful difference between the PL and the image's AC score, which implies that longer prompts generate better images in terms of architectural quality, although there can be some exceptions (Table 3).


AC score: 2.3 (low)


AC score: 3 (medium)


AC score: 4 (high)

Figure 3. Examples from images with the lowest, medium, and highest AC scores

A similar result is observed in terms of descriptive language (DL). The images labelled as high in terms of AC, had an average of 13.08 and median of 11 specific words, medium had an average of 6.87 and median of 6, while low had an average of 4 and median of 4. Moreover, some images of medium and low AC scores had 0 specific words in the prompt, while images with high AC scores had at least 6 specific words. The numbers displayed a relationship

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

between the AC and DL exits, meaning that descriptive language with more specific words generates more creative images when compared to non-descriptive language with general statements. After all, although longer prompts generally generate better outputs, they may not be enough to generate quality outputs unless they include more descriptive language (Table 3).

Scores for Architectural Creativity (AC)	Number of images	Prompt length (PL) (No. of words)			Descriptive language (DL) (No of specific words)		
		Range	Average	Median	Range	Average	Median
High (3.6-4)	12	15 - 50	35.08	31	6 - 19	13.08	11
Medium (3-3.5)	24	2 - 56	22.45	15	0 - 17	6.87	6
Low (2.3-2.9)	12	5 - 20	13.91	14	0 - 7	4	4

Table 3. Relationship between AC scores of the images and PL + DL

The descriptive language (DL) of prompts was analysed in detail in terms of (1) architectural style, (2) volume and shape, (3) material and surface, and (4) the architect's name. The following results were found: (a) Images with high AC scores: 11/12 prompts had at least one style indicator, 7/12 had at least one volume and shape description, 12/12 had material and surface description, and 10/12 had at least one architect name; (b) Images with medium AC scores: 16/24 prompts had a style indicator, 11/24 had at least one volume and shape description, 12/24 had material and surface descriptions, 12/24 had an architect name; (c) Images with low AC scores: 9/12 prompts had a style indicator, only 1/12 had a shape and volume description, 8/12 had material and surface description, and 4/12 had an architect name The percentages of all four indicators used in the prompt gradually decrease from high AC images to low AC images (except the increase of one indicator from medium to low AC). (Table 4). Although it can be interpreted as the usage of indicators enhances the quality of image in terms of architectural creativity, a more detailed analysis of indicators is needed to understand their relation to architectural creativity by asking the question, "Does specific styles, volume, material and surface indicators affect the architectural creativity?"

Table 4. Relationship between architecture-related indicators and AC scores

Indicators in the prompt	High AC 12 prompts		Medium AC 24 prompts		Low AC 12 prompts	
	No. of prompts	Percent.	No. of prompts	Percent.	No. of prompts	Percent.
Architectural style indicator	11	%91.6	16	%66.6	9	%75
Volume and shape indicator	7	%58.3	11	%45.8	1	%8.3
Material and surface indicator	12	%100	12	%50	8	%66.6
Architect's name	10	%83.3	12	%50	4	%33.3

To answer this question, the frequency of (a) architectural style, (b) volume and shape, (c) material and surface, (d) architect name used in the prompts were analysed and compared with the scores of the images. According to the results, specific styles do not seem to directly relate to the quality of architectural creativity, nor the specific volume and shape descriptions. For example, the most frequently used style indicator, the word "parametric," and the volume

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

indicator, the word "organic", were used 5 and 7 times in prompts of high AC, while 7 and 8 times in prompts of medium AC. About the architect names, only SANAA and Oki Sato seem to have a positive effect on the score, but the available data is not sufficient to make that conclusion. However, the analysis illustrated that choice of materials and surfaces has a more direct relation to architectural creativity. Prompts included specific materials such as "rock, stucco, corten steel, rough stone, inflatable membrane", generated images with high AC scores, while prompts included common architectural materials such as "wood, steel, glass, ceramic, plaster", generated images with lower AC scores. In terms of "non-architectural" materials (such as "wool, crystal, ice and snow"), the analysis showed no significant result, as these words were related to both high and low scores (Table 5). However, more than the use of specific indicators, finding a way to combine them in a creative and meaningful way seems to be the key issue in the architectural creativity of AI-generated architecture.

Table 5. Frequency of architecture-related indicators compared to the AC scores

Indicators in the prompt		Frequency					
		High AC		Medium AC		Low AC	
		No.	%	No.	%	No.	%
Architectural Style	Parametric	5	41.6	7	29.1	1	8.3
	Futuristic	2	16.6	1	4.15	1	8.3
	High-tech	1	8.3			3	25
	Contemporary	1	8.3	1	4.15	2	16.6
	Bionic	1	8.3				
	Scandinavian	1	8.3				
	Japanese	1	8.3				
	Persian			2	8.3	2	16.6
	Brutalist			2	8.3	1	8.3
	Blobitecture			2	8.3		
	Modern			1	4.15		
Volume and Shape	Organic	7	58.3	8	33.3		
	Cave-like/excavated	2	16.6				
	Biomorphic			1	4.15		
	Modular			1	4.15		
	Fluid			1	4.15		
	Floating/hanging	1	8.3	1	4.15	1	8.3
	Cubic					1	8.3
Material and Surface	Ice	3	25				
	Snow	3	25				
	Rocky	3	25				
	Stacked	3	25				
	Rough stone	1	8.3				
	Stucco	2	16.6				
	Corten steel	1	8.3				
	Glass	1	8.3	2	8.3	2	16.6
	Concrete	1	8.3	2	8.3	1	8.3
	Reflected crystal	1	8.3			1	8.3
	Perforated	2	16.6	5	20.8		
	Voronoi cells pattern	1	8.3				
	Inflatable membrane	1	8.3				
	Wood			5	20.8	4	33.3

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

Indicators in the prompt		Frequency					
		High AC		Medium AC		Low AC	
		No.	%	No.	%	No.	%
	Translucent			1	5.15		
	Steel			2	8.3	1	8.3
	Ripped			1	4.15		
	Wool			2	8.3		
	Mesh					2	16.6
	Ceramic					1	8.3
	Plaster					1	8.3
Architects	SANAA	3	25				
	Oki Sato	3	25				
	Daniel Libeskind	1	8.3			1	8.3
	Zaha Hadid	3	25	1	4.15		
	Frank Gehry	4	33.3	6	25		
	Antonio Gaudi	1	8.3			1	8.3
	Oscar Niemeyer			1	4.15		
	Neri Oxman			2	8.3		
	Alireza Taghaboni			1	4.15		
	Simon Velez			2	8.3		
	Le Corbusier					1	8.3
	Enric Miralles					1	8.3

Table 5. Frequency of architecture-related indicators compared to the AC scores

5. (Un)conclusions

The findings of the research revealed how the prompt structure, in terms of prompt length, descriptive language, and specific architecture-related indicators, affected the generated outputs, underlying the significance of language as a design driver. Furthermore, as generative models synthesise the "existing" data of the visual representations, making conscious and deliberate selections (the ability of decision-making) becomes of critical importance, which also implies a radical transformation of the design process into a "selection" process. Another critical aspect is understanding how to call visual data by textual/linguistic representations. However, finding ways to combine the existing representations in a creative and meaningful way seems to be the key issue in the architectural creativity of AI-generated architecture. Another implication is the research is that, as architects, we do not yet have the necessary means to assess the outputs of generative models in terms of architectural creativity. To be able to assess the outputs of AI-generated architecture precisely, we need to establish new criteria and standards and redefine the concept of creativity, recognizing AI technologies as new methodologies and approaches to design. Although image generation with deep learning technologies has great potential to enhance our cognitive abilities as architects and to push the boundaries of architectural creativity further, further steps are needed to be taken for the adoption of these technologies in overall design processes.

References

Bolojan, D. (2022). Creative AI: Augmenting Design Potency, in Machine Hallucinations: Architecture and Artificial Intelligence, Mathias del Campo and Neil Leach (ed.) *AD*, 3 (92), 22-27.

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

Bommasani, R. (2023). AI Spring? Four Takeaways from Major Releases in Foundation Models, Stanford Institute for Human-Centered Artificial Intelligence. Retrieved from https://hai.stanford.edu/news/ai-spring-four-takeaways-major-releases-foundation-models, 30.07.2023.

Brown, T.B, Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, R., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henghan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., & Dario, A. (2020). Language Models are Few-Shot Learners. *ArXiv*, 22 Jul 2020, https://doi.org/10.48550/arXiv.2005.14165

Carpo, M. (2017). The Second Digital Turn: Design Beyond Intelligence. MIT Press.

Chaillou, S. (2022). Artificial Intelligence and Architecture: From Research to Practice. Basel: Birkhauser.

Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis*, 58(1), 7-19. https://doi.org/10.1093/analys/58.1.7

Crespo, S., & McCormick, F. (2022). Augmenting Digital Nature: Generative Art as a Constructive Feedback Loop, in Machine Hallucinations: Architecture and Artificial Intelligence, Mathias del Campo and Neil Leach (ed.) *AD*, 3 (92), 54-59.

Dang, H., Mecke, L., Lehmann, F., Goller, S., & Buschek, D. (2022). How to Prompt? Opportunities and Challenges of Zero- and Few-Shot Learning for Human-AI Interaction in Creative Applications of Generative Models. ArXiv. 3 Sep 2022. https://doi.org/10.48550/arXiv.2209.01390

Deckers, N., Fröbe, M., Kiesel, J., Pandolfo, G., Schröder, C., Stein, B., & Potthast, M. (2023). The Infinite Index: Information Retrieval on Generative Text-To-Image Models. *ArXiv*, 21 Jan 2023. https://doi.org/10.48550/arXiv.2212.07476

Dhariwal, P., & Nichol, A. (2021). Diffusion Models Beat GANs on Image Synthesis. *arXiv*, 1 Jun 2021. https://doi.org/10.48550/arXiv.2105.05233

del Campo, M., & Leach, N. (2022). Can Machines Hallucinate Architecture?: AI as a Design Method, in Machine Hallucinations: Architecture and Artificial Intelligence, Mathias del Campo and Neil Leach (ed.) *AD*, 3 (92), 6-13.

del Campo, M., & Manninger, S. (2022). Strange but Familiar Enough: The Design Ecology of Neural Architecture, in Machine Hallucinations: Architecture and Artificial Intelligence, Mathias Del Campo and Neil Leach (ed.) *AD*, 3 (92), 38-46.

del Campo, M., & Carlson, A, I. (2022). Strange but Familiar Enough: Reinterpreting Style in the Context of AI, in Artificial Intelligence and Architecture: From Research to Practice, Stanislas Chaillou (ed.) Basel: Birkhauser, 72-79.

dPrix, W., Schmidbaur, K., Bolojan, D., & Baesta, E. (2022). The Legacy Sketch Machine: From Artificial to Architectural Intelligence, in Machine Hallucinations: Architecture and Artificial Intelligence, Mathias del Campo and Neil Leach (ed.) *AD*, 3 (92), 14-21.

Goodfellow, I. J, Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014) Generative Adversarial Networks. *arXiv*, 10 June 2014. https://doi.org/10.48550/arXiv.1406.2661

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738

Karras, T., Laine, S., & Aila, T. (2019). A Style-Based Generator Architecture for Generative Adversarial Networks. *arXiv*, 29 Mar 2019. https://doi.org/10.48550/arXiv.1812.04948

Koh, I. (2022). Architectural Plasticity: The Neural Sampling of Forms, in Artificial Intelligence and Architecture: From Research to Practice, Stanislas Chaillou (ed.) Basel: Birkhäuser, 110-117.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, E. S. (2006 [1995]). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 1955. *AI Magazine*, 24(4), 12-14. https://doi.org/10.1609/aimag.v27i4.1904

Schön, Donald A., (1987). Educating the Reflective Practitioner: Towards a New Design for Teaching and Learning in the Professions, San Francisco, CA: Jossey-Bass Publishers.

Webster, H., (2008). Architectural Education after Schön: Cracks, Blurs, Boundaries and Beyond. *Journal for Education in the Built Environment*, 3(2), 63-74.

Oppenlaender, J., Linder, R., & Silvennoinen, J. (2023). Prompting AI Art: An Investigation into the Creative Skill of Prompt Engineering. *arXiv*, 13 Mar 2023. https://doi.org/10.48550/arXiv.2303.13534

Oppenlaender, J., (2022). The Creativity of Text-to-Image Generation. *arXiv*, 31 Oct 2022. https://doi.org/10.48550/arXiv.2206.02904

Howe, J. (1994). Artificial Intelligence at Edinburgh University: a Perspective. Retrieved from https://www.inf.ed.ac.uk/about/AIhistory.html, 11.04.2023.

Reynolds, L., & McDonell, K. (2021). Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm. *ArXiv*, 15 Feb 2021. https://doi.org/10.48550/arXiv.2102.07350

Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., & Sutskever, I. (2021). Zero-Shot Text-to-Image Generation. *ArXiv*, 26 Feb 2021. https://doi.org/10.48550/arXiv.2102.12092

Russel, S. J., & Norvig, P. (2003). Artificial Intelligence: A Modern Approach, New Jersey: Prentice Hall.

Received 28 August 2023; Received in revised form 17 September 2023; Accepted 20 September 2023; Available online 25 September 2023;
doi: 10.5281/zenodo.10057738