Research Article

Generating Images with AI: A Case Study from Facebook Public Group “Free Prompt Gen Image”

Nawarit Rittiyotee[1]

Article History

Received: November 27, 2024

Revised: May 28, 2025

Accepted: June 11, 2025

Abstract

This study investigates generating image with AI in early-stage period to artificial intelligence (AI) by examining user interactions and prompt structures on the Facebook Public Group "Free Prompt Gen Image". The research questions were what are the common patterns in prompt construction among users and what best practices can be developed for crafting prompts to optimize the outcomes of AI image generation. Drawing on a dataset of over 1,000 prompts, the research applicates Content Analysis approach and explores how specific elements—such as subject specificity, environmental context, and stylistic details—impact the quality and accuracy of AI-generated visuals. The findings found 7 distinct seven distinct groups: Animals, Food, Graphics, Humans, Imagination, Objects, and Views. Each category demonstrated unique structural patterns and user preferences in prompt composition. The research highlight the critical role of prompt design in optimizing outputs, as well as the challenges posed by technical complexity, language barriers, and ethical considerations.

The study also compares the use of web-based versus computer-based AI platforms, identifying a preference for accessible web-based tools despite their limitations in flexibility and customization. Ethical issues, including copyright concerns and data ownership, are also addressed. By emphasizing the need for improved user education and the development of intuitive interfaces, the research contributes to enhancing the usability and creativity of generative AI technologies.

[1] School of Communication Arts, Bangkok University

E-mail: nawarit.r@bu.ac.th

Future directions include refining prompt-crafting methodologies, advancing multimodal input systems, and fostering collaborative repositories of effective prompts.

Keywords: *AI, Image Prompts, Content Analysis, Social Media, Digital Imagery, Creative Technology*

1. Introduction

Image-generative AI has revolutionized digital content creation by enabling the synthesis of realistic and creative visuals through advanced models like Generative Adversarial Networks (GANs) and diffusion models. These tools democratize high-quality image generation and find applications across advertising, entertainment, medical imaging, and design.

This study analyzes user interactions with image-generative AI on the Facebook Public Group Free Prompt Gen Image, focusing on patterns in prompt composition, their impact on output quality, and platform preferences. It also examines challenges like technical barriers, ethical concerns, and output limitations.

By exploring these dynamics, the research highlights the importance of prompt design in optimizing AI-generated imagery and its implications for creativity, accessibility, and ethical use.

1.1 Research rationales

Prompts as key to effective AI-generated imagery

Prompts are the foundation of image-generative AI systems, serving as the main input that guides the creation of visual content. The effectiveness of these prompts directly impacts the quality, accuracy, and artistic fidelity of the generated outputs. Despite this central role, crafting effective prompts is not intuitive and requires users to learn, practice, and refine their input techniques. Challenges such as ambiguity, lack of specificity, and varying model interpretations necessitate a deeper understanding of prompt structures and their influence on AI-generated imagery. Research into prompt design is critical for developing frameworks that empower users to achieve their desired results. By analyzing common patterns, recurring elements, and effective practices in prompt crafting, this study aims to bridge the knowledge gap and provide actionable insights. This focused not only improve user outcomes but also contribute to making generative AI tools more accessible and intuitive for diverse audiences.

Exploration of early-stage image-generative AI platforms

The field of image-generative AI is in a nascent stage, with many tools and services, such as MidJourney, Stable Diffusion, and DALL-E, launched within the past year. These platforms are undergoing rapid development and experimentation, offering unprecedented capabilities but also revealing limitations such as inconsistent output quality, user accessibility challenges, and ethical concerns. Given this experimental period, it is essential to explore how these tools are being used, their strengths and weaknesses, and the evolving preferences of users. Additionally, as users experiment with different platforms and models, patterns in usage, challenges in adaptation, and areas for improvement emerge, providing a valuable basis for shaping future advancements.

However, until now there has no clear guideline on how to craft an effective prompt for an image generative AI. Users are still struggling on a trial and error. Studying prompt structures from other users can be a way to achieve prospected image generating. Therefore, this research is aimed at exploring and analyzing prompts to extract prompt structures for those interested in generating image with AIs.

In summary, this research examines user interactions with image-generative AI systems, specifically analyzing prompt patterns to find image prompt structures through data collected from the Facebook Public Group "Free Prompt Gen Image." The study focuses on how prompt composition affects output quality, explores platform preferences among users, and identifies key challenges including technical barriers, ethical concerns, and output limitations. Therefore, the research can be utilized to develop comprehensive prompt frameworks and provide educational resources that improve user outcomes with image-generative AI tools.

2. Research Objectives

1. To analyze the structural patterns of image-generating prompts shared in the Facebook group "Free Prompt Gen Image" to identify elements that influence the quality and accuracy of AI-generated images.

2. To develop actionable insights and best practices for crafting prompts to optimize AI image generation outcomes

To achieve the objectives, two research questions were:

1. What are the common patterns in prompt construction among users?

2. What best practices can be developed for crafting prompts to optimize the outcomes of AI image generation?

3. Literature Review

3.1 Introduction to image-generative AI

Image-generative AI focuses on creating visual content using advanced deep learning models like Generative Adversarial Networks (GANs) and Diffusion Models, enabling photorealistic image generation from textual prompts, sketches, and more (Yang et al., 2023). These innovations have transformed creative industries, including art, entertainment, medical imaging, and autonomous driving (Pise et al., 2024).

The progression from simple, low-resolution outputs to highly detailed visuals highlights the integration of natural language processing (NLP) and computer vision. Tools like GPT-3, Stable Diffusion, and MidJourney have expanded the field's accessibility and creative potential (Srivastava, 2024).

3.2. Technological foundations of image-generative AI

Image-generative AI relies on advanced deep learning models, primarily Generative Adversarial Networks (GANs) and Diffusion Models. These architectures generate high-quality images by learning patterns from large datasets.

GANs, introduced by Goodfellow et al. (2014), use two neural networks—a generator and a discriminator—in a competitive framework, resulting in increasingly realistic images. GANs excel in creative applications like artwork and style transfers but face challenges like mode collapse, which reduces output diversity (Sakirin & Kusuma, 2023).

Diffusion Models, introduced by Ho et al. (2020), progressively add and remove noise to generate realistic images. They address GAN limitations, such as mode collapse, and excel in tasks requiring high resolution, including inpainting, image translation, and super-resolution (Srivastava, 2024).

3.3. Advancements in image-generative AI

Advancements in AI have revolutionized image generation, enabling models like DALL-E, MidJourney, and Stable Diffusion to create high-resolution, realistic images from textual prompts.

This integration of natural language processing (NLP) and computer vision allows these models to interpret complex prompts and produce nuanced outputs, empowering users without technical expertise to generate professional-quality images (Srivastava, 2024).

The democratization of AI tools has enhanced creative workflows for artists, designers, and hobbyists. Text-to-image AI has increased productivity by 25% while introducing innovative workflows blending human and AI creativity, described as "generative synesthesia" (Zhou & Lee, 2024). These tools are widely adopted across industries, from architectural design visualization to advertising content creation (Pise et al., 2024).

3.4. Applications of image-generative AI

Image-generative AI has applications beyond art and recreation, spanning industries like data augmentation, creative content, and entertainment:

3.4.1 Data Augmentation: AI generates synthetic data to supplement real-world datasets, enhancing model training in fields like medical imaging and autonomous driving. This improves diagnostic accuracy and robustness (Pise et al., 2024).

3.4.2 Creative Content Generation: Artists and designers use tools like MidJourney and Stable Diffusion to create unique digital art, product designs, and game visuals, lowering entry barriers for exploring new styles (McCormack et al., 2024).

3.4.3 Style Transfer and Image Translation: Techniques like style transfer merge artistic elements from one image to another, while image-to-image translation transforms sketches into realistic visuals or shifts between styles (Srivastava, 2024).

3.4.4 Entertainment and Advertising: AI-generated imagery is used for special effects, character designs, and backgrounds in films and games, while advertising benefits from customizable, cost-effective visuals for rapid concept testing (Pise et al., 2024).

3.5 Impact of prompts in image-generative AI

Prompts are crucial in image-generative AI, guiding models like DALL-E and Stable Diffusion to generate specific outputs based on detailed textual descriptions. The quality of these prompts directly impacts the aesthetic and conceptual quality of the images.

Research by McCormack et al. (2024) highlights that prompts often focus on surface aesthetics, such as color schemes, rather than deeper conceptual themes. Effective prompt design

can enhance creative potential, especially in professional settings like architecture, where precise prompts ensure visuals align with design intentions (Zhang & Liu, 2024). However, some platforms limit creativity by discouraging detailed prompts, restricting exploration of novel ideas (Torricelli et al., 2023).

3.6 The importance of prompts in image generation

The quality of AI-generated images depends on the precision and clarity of the prompts, which guide the AI in capturing details like lighting, color, textures, and spatial relationships. Vague prompts lead to generic images, while detailed prompts produce more accurate visuals. McCormack et al. (2024) note that users often focus on surface aesthetics, reinforcing conventional imagery and limiting creative exploration.

In fields like architecture and design, precise prompts are essential for achieving specific visual goals. Zhang and Liu (2024) stress the importance of well-crafted prompts in controlling AI outputs, underscoring the need for further research to refine prompt-crafting techniques and unlock the full creative potential of image-generative AI.

4. Methodology

4.1 Research methodology

This study employs a qualitative research approach to investigate the structure of image-generating prompts shared on the Facebook Public Group "Free Prompt Gen Image" which has more than 169,000 members sharing image prompts. Qualitative methods are particularly suitable for this research as they allow for an in-depth exploration of the shared prompts, enabling the identification of common structures and patterns (Creswell & Poth, 2018). This approach provides the flexibility to analyze the nuanced language and context of the prompts, which is crucial for understanding their composition and effectiveness in AI image generation.

4.2 Scope of study

This research focuses on the use and efficacy of Image Generative AI systems, specifically analyzing the prompt structure shared by users of the Facebook Public Group "Free Prompt Gen Image." The study aims to explore the impact of prompt crafting on the quality of AI-generated visuals and examines the recurring patterns in user inputs to identify best practices for maximizing

creative and technical outputs. The research encompasses prompt composition, image model analysis, user behavior and challenges, ethical and practical considerations.

4.3 Data collection

The study collected secondary data from the Facebook Public Group "Free Prompt Gen Image." This data consisted of user-shared image-generating prompts posted on the public group between October and November 2024. Only posts with clear prompt structures were included to ensure that the patterns could be analyzed while posts without prompt structures such as prompts with a single word or fragmented words were excluded.

4.4 Data analysis

Content analysis was the primary method used to analyze the collected prompts. This method is particularly effective for systematically examining text data to identify patterns, themes, and structures (Krippendorff, 2018). The content analysis followed these steps:

1. Initial reading: Researchers familiarized themselves with the collected prompts. Example of collected prompt is “Create a 4K image full body of a handsome 25-year-old man, white bikini with a muscular build, lying in a luxurious bathtub. The background should feature an elegant, high-end bathroom setting.”.

2. Coding: Prompts were coded based on their content, structure, and other relevant characteristics. For example, 4K image was coded as Quality Emphasis, a luxurious bathtub was coded as a Setting while a bikini was coded as clothing.

3. Categorization: Based on the coding, prompts were categorized into distinct groups. For example, angel, spirit, ghost and devil were categorized as Imagination.

4. Structure analysis: Researchers analyzed the categorized and structured data to identify recurring patterns and trends in prompt composition.

This methodology allows for a comprehensive examination of image-generating prompts, providing insights into effective prompt structures and common practices among users of AI image generation tools.

5. Findings

This study analyzed 1,009 image prompts shared on the Facebook Public Group between October and November 2024. The analysis focused on identifying common patterns, structural elements, and user preferences across different categories of image generation prompts. The prompts were categorized into seven distinct groups: Animals, Food, Graphics, Humans, Imagination, Objects, and Views. Each category demonstrated unique structural patterns and user preferences in prompt composition.

5.1 Image prompt structures

5.1.1 Animals

From 36 prompts, images of animal were generated. This category is identified by the main subject of the prompts which basically are animals such as dog, cat, lion, snake and other species of animals based on users' interest. Some animals are generated as symbols; for example, lion as a representative of elegant and leadership while snake as a symbol of danger and betrayal. It can be Moo-deng which during a sampling selection is a baby Pygmy hippopotamus in a spotlight of the world.

Prompt structure for animals

To analyze the structure of the image prompts, we can look at several common elements that frequently appear in the prompts and understand how they work together to create a vivid and detailed description. Based on the provided examples, here's a breakdown of frequent structural components and their functions:

1. Subject and setting

Each prompt typically begins by identifying the main subject (e.g., "a little hippopotamus," "a grey and white Persian cat," "a tiger"). This is followed by specific details that shape the setting or environment, giving a context in which the subject exists (e.g., "on a snow-covered mountain," "in a swimming pool scene").

Example: "A golden eagle, with black highlights, is soaring in the sky."

2. Attributes and actions of the subject

Descriptions often include specific characteristics or actions, which help personalize and animate the subject. For example, specifying expressions ("relaxed, almost melancholic mood"), postures ("perched delicately"), or emotions ("happy, cheerful, and laughing") gives the image a sense of personality.

Example: "A baby hippopotamus wearing a helmet... happy, cheerful, and laughing."

3. Detail emphasis (color, texture, lighting)

High-frequency elements include color (e.g., “golden fur,” “vibrant plumage with a deep blue head”), texture (e.g., “soft brown couch cushion,” “smooth grayish-brown skin”), and lighting (e.g., “soft and natural,” “sunlight illuminates the scene”). These features establish realism, mood, and focus.

Example: "Soft, natural lighting, highlighting the dog's warm, golden fur."

4. Environmental context and background

Many prompts provide a blurred or softened background, often described with color or style (e.g., “a blurred indoor setting with neutral tones,” “fiery red and orange sky”), which focuses attention on the subject and adds visual depth. This technique is particularly common in nature or wildlife scenes.

Example: "The background is a blurred indoor setting with neutral tones."

5. Tone or style descriptions

Prompts also frequently specify the artistic tone or style, such as “photorealistic,” “macro,” “high definition,” or “ultra quality,” which clarifies the intended artistic effect.

Example: "Macro photograph of a ladybug on a leaf, extreme close-up."

6. Mood or atmosphere

Words describing mood (e.g., “relaxed,” “lighthearted and fun,” “enchanting and serene”) help evoke the emotional tone the artist should capture. These phrases are often added toward the end of the prompt to summarize the overall feeling or impact.

Example: "Creating a serene, nature-filled atmosphere."

7. Technical details

Some prompts specify technical aspects like resolution, style variations, or desired aspect ratios, especially for complex scenes (e.g., “4K image,” “--ar 16:9,” “512K”).

Example: "ultra-realistic 4K image of a big whale shark swimming in a deep blue ocean."

5.1.2 Foods

Food is also a category in an interest of users. This category was elicited from 63 prompts. The challenge of generating this type of image is how to create realistic pictures of what users intend to. Food images that users like to generate are Western, Japanese and Thai. From

observations, some image models are more adept at generating realistic foods while some spare more room for imagination which sometimes leads to irrelevant ingredients or even create wrong image of food that the users meant.

Prompt structure for foods

Analyzing the structure of these food image prompts reveals several recurring elements that contribute to clarity and visual appeal. Here's a breakdown of the common components used to effectively write image prompts:

1. Subject with descriptive naming and ingredients

Prompts start by identifying the food item, often with detailed descriptors that define its appearance and components. Specific ingredients are listed, sometimes down to garnishes or sauces.

Example: "Create an image of dry rice noodle mixed with brown sauce, topped with pork, pork ball, flat round pork mince, and bean sprouts."

2. Preparation style and serving context

Many prompts specify how the dish is prepared (e.g., "fried," "steamed," "grilled") and presented (e.g., "in a bowl," "on a plate"), which gives more context to the scene.

Example: "Grilled Thai chicken presented in a traditional Thai style with dipping sauce, cucumber, and sticky rice on a banana leaf."

3. Detail-oriented descriptions for texture and appearance

Texture is emphasized to enhance the realism, such as describing the dish as "crispy," "gooey," or "tender." Colors and visual qualities of the food are also highlighted to create a vivid picture (e.g., "golden-brown skin," "gooey interior and exterior").

Example: "Grilled bananas on skewers with caramel sauce, focusing on the gooey texture."

4. Backdrop and environment

The background, whether minimalistic or atmospheric, contributes to the image's tone. Common settings include rustic kitchen setups, bustling street food scenes, or natural light backgrounds that emphasize authenticity.

Example: "Set on a wooden table by the lakeside, with a natural backdrop."

5. Lighting and photography style

Prompts often specify the lighting type (e.g., “natural light,” “studio lighting”), as well as stylistic elements like “macro,” “photorealistic,” or “professional food photography,” adding focus to the technical execution.

Example: “Studio lighting, food photography style.”

6. Enhancing elements with added visual effects

Some prompts include specific effects to enhance the appeal, such as “smoke rising” for warmth, “glistening” for freshness, or splashes of liquid for dynamic visuals. These details add liveliness and realism to the scene.

Example: “Smoke rising from a hot bowl of stew with a hint of bubbles.”

7. Cultural and traditional references

Certain prompts incorporate cultural or regional specifics, like describing Thai or Japanese traditional styles, which adds a cultural element and directs the scene’s tone.

Example: “A bowl of Khao Soi noodle soup, street food style photo, shallow depth of field.”

5.1.3 Graphics

Many users love to generate image with various styles of art. Among 1,009 prompts, 57 prompts were generated by the users. This category includes sticker designs, movie posters, drawings, oil paintings, logos, typographies and so on. This type is image is specifically identified due to specific purposes of users which are not meant to be realistic but to deliver artistic styles of image objects or overall pictures.

Prompt structure for graphics

In analyzing the structure of the graphic design prompts, here are the main elements frequently used to write clear and effective image prompts:

1. Primary subject with detailed visual description

Each prompt typically begins by identifying the main subject (e.g., a character, animal, or object) with a clear visual description. Specific details of the subject’s appearance, including colors, textures, and distinguishing features, are provided to enhance visual clarity.

Example: "A highly detailed close-up portrait of a woman in traditional Chinese imperial clothing and a large vibrant Chinese dragon."

2. Stylistic elements and art style

The prompts frequently specify the desired art style, which could range from “realistic” and “minimalist” to more stylized forms such as “anime-style,” “watercolor,” “cubist,” or “oil painting.” This guides the artistic approach and tone of the image.

Example: "Minimalist illustration of a male astronaut in a spacesuit, standing next to a rocket."

3. Composition and background elements

The composition often includes information on the subject’s setting or background, such as indoor scenes, natural landscapes, or plain backgrounds to enhance focus. This adds depth to the visual and sets the scene for the main subject.

Example: "Standing atop a rugged mountain peak, with layered mountains in the background painted in soft tones of brown and beige."

4. Emphasis on texture and detail

Many prompts include detailed descriptions of textures and materials, which add realism and character. For instance, prompts might specify textures like “silk,” “metallic,” or “stone,” enhancing the quality of the image and the viewer’s sense of materiality.

Example: "The lion’s mane is made of Bitcoin elements with digital coin textures."

5. Lighting and mood

Descriptions of lighting, such as “warm golden light,” “subtle rays of sunlight,” or “eerie white light,” set the mood and tone of the artwork. This aspect frequently influences the emotional impact of the image.

Example: "The scene is bathed in warm golden light, evoking a soft, dreamy feeling."

6. Color scheme

A defined color palette (e.g., “blue, black, and cream” or “green and gold”) is commonly provided to guide the tone and cohesion of the image. Using specific color schemes also helps unify the visual style across the image elements.

Example: "A high-resolution 4K Halloween-themed image featuring a large carved pumpkin... illuminated from within."

7. Dynamic poses or actions

The poses or actions of the characters are often specified, which adds a dynamic aspect to the prompt. For instance, “sitting,” “reclining,” or “running” gives a sense of movement and liveliness to the design.

Example: "A chibi girl in a red lace dress, sitting on a luxurious stool with a relaxed pose."

8. Cultural or traditional elements

Cultural references like traditional attire, symbols, or patterns add authenticity and can be important for the theme. This aspect personalizes the scene, especially when capturing styles specific to a culture or historical period.

Example: "Traditional Asian art style, with two koi fish on a dark background, detailed patterns."

5.1.4 Humans

It is challenging for Imaginative AI to generate a realistic picture of human. Previously there was a clear border between human photography and generated image of human. However, the border seems to be faded by the development of Imaginative AIs. During observations, it can be clearly seen that generating human images are the most popular comparing to other types of images, 470 from 1,009 prompts. Most images are usually female and male models in different poses. Users often identified nationality of the models such as Chinese, Thai, Japanese or sometimes loosely specify it as Asian or Western. Famous models of celebrities are also generated such as Lisa in a costume designed by Victoria Secrets. Apart from these, users liked to generate humans in festivals such as Loi-kratong and Halloween.

Prompt structure for humans

The prompts showed several recurring structures that can help guide how to write prompts effectively based on commonly used elements:

1. Setting and background descriptors

Prompts often begin by establishing the location, time of day, or background scenery. This serves as a grounding element and often defines the mood of the image.

Example: "In a room with a white wall and a wooden shelf in the background."

2. Character and clothing details

The character's appearance is described in detail, including clothing, posture, expressions, and physical attributes. Emphasis on attire is particularly common for prompts describing specific attire, cultural elements, or aesthetic goals.

Example: "A woman wearing tank top and denim shorts, long brown hair with bangs, looking directly at the camera."

3. Emotional or expression-based cues

Emotional tone or facial expressions of characters are frequently specified to establish the intended mood or personality.

Example: "Smiling happily, exuding warmth."

4. Lighting and effects

Specific lighting styles such as "warm tones," "natural light," or "volumetric lighting" are common, adding depth and mood to the images. Effects like "soft bokeh" or "HDR" are also used to specify visual clarity.

Example: "Soft, natural lighting, giving a calm, uncluttered atmosphere."

5. Detailed style and quality emphasis

Many prompts include qualifiers for quality like "realistic," "8K resolution," "masterpiece," and "photorealistic," especially when aiming for high-resolution or highly detailed images.

Example: "Super extremely ultra detailed, masterpiece, high-res."

5.1.5 Imagination

This category is identified by 271 prompts aimed at generating human, animal, object and scene which do not exist in a real world. Animal from mythology such as Naga (a huge snake from Buddhist tales), dragon, ghosts or spirits, witches and half-human-half- animals were commonly identified. Also, caricatures of politicians, celebrities, and historic events were generated.

Prompt structure for imaginations

Based on the analyzed prompts, several recurring structural patterns emerge, revealing how detailed prompts are crafted to achieve specific image outcomes. Here’s a breakdown of these patterns, examples, and a synthesis for creating effective prompts:

Key structural patterns in prompts

1. Primary subject or character

Most prompts begin by clearly identifying the main subject or character, often with distinct details on appearance, attire, and personality. This establishes the core focus of the image.

Example: "A fierce dragon with glowing red eyes, perched atop a rocky mountain."

2. Detailed description of setting and background

Adding environmental context helps to place the subject in a cohesive scene, whether mystical, futuristic, or natural. Specific setting elements are frequently included, like “forested mountains” or “futuristic cityscape.”

Example: "A massive underwater city with glass domes and bioluminescent flora."

3. Lighting and atmosphere

Many prompts specify lighting conditions, which significantly impact mood. Common descriptors include “soft, warm light,” “harsh moonlight,” and “neon-lit atmosphere.”

Example: "Soft, morning sunlight streaming through forest trees."

4. Pose and action

Descriptions of dynamic poses or specific actions add life to the image, making it more engaging. Including the pose or motion of a character adds depth.

Example: "A samurai mid-swing, katana slicing through the air."

5. Art style and quality

Specificity in style helps clarify visual goals, with terms like “photorealistic,” “watercolor,” “8K resolution,” and “oil painting style” frequently used.

Example: "Watercolor style, pastel colors, soft lighting, 4K resolution."

5.1.6 Objects

This category is identified by analyzing 112 prompts aimed at generating objects with specific details such as a city model made of glass, a royal barge, a flower made of gems, figures, accessories and so on.

Prompt structure for objects

The prompts in this category follow a structured approach to detailing objects, scenes, and settings with specific recurring elements. Here’s a breakdown of the patterns, frequent descriptors, and a synthesis to write prompts that are both coherent and rich in imagery.

1. Primary focus (object) and detail description

The main object is often described with intricate details regarding its appearance, form, and materials. These specifics give a clear image of the primary subject.

Example: “A premium jewelry display featuring a set of emerald and diamond earrings, presented on a black velvet stand with gold trimmings.”

2. Setting and background context

Adding environmental elements like the background, lighting, or specific surroundings gives context to the object, helping to situate it in a visually coherent scene.

Example: “An intricate bow made of gold ribbon, adorned with diamonds and onyx, placed atop a giant X-Mas themed background.”

3. Lighting and atmosphere

Lighting is essential in defining the mood or feel of the scene. Descriptors such as “soft light,” “dramatic lighting,” or “sunlight streaming through” add dimension.

Example: “A minimalist style bedroom with clean lines and a neutral color palette, with natural light filling the space.”

4. Material and texture emphasis

Many prompts highlight the texture or material quality, which can enhance the visual realism or artistry.

Example: “A cozy, atmospheric café interior that features a blend of rustic and industrial design elements.”

5. Artistic style and quality terms

Specific styles (e.g., hyper-realistic, 3D render) and quality indicators (e.g., high resolution, cinematic lighting) are often included to specify visual expectations.

Example: “Hyper-realistic, ultra-detailed, cinematic lighting, octane render.”

5.1.7 Views

This type of image depicts a beautiful scene or place in a specific atmosphere. Users aimed at creating images of landscape, view or place usually seen in photos. Some images specify famous places or sightseeing locations in the real world.

Prompt structure for views

Analyzing these prompts reveals several patterns, particularly in the structure and content that aim to create vivid, detailed images. Here's a breakdown of the key patterns and an analysis of how they can guide prompt writing:

1. Subject description with focused detail

Pattern: Each prompt begins by specifying the main subject or scene, such as "a glowing, full yellow moon" or "a wooden path stretching across lush green hills." These subjects are often enhanced by including specific descriptors.

Example: Instead of "A forest scene," use "A dense forest scene with towering pine trees, their trunks illuminated by soft morning light."

2. Environmental and lighting details

Pattern: Prompts frequently use adjectives to describe lighting, mood, and the overall ambiance of the scene, such as "golden light of the setting sun" or "ethereal morning mist." Lighting is crucial for setting the mood and enhancing realism.

Example: "The early morning sun filters through thick fog, casting a pale golden glow over the quiet lake."

3. Foreground and background layers

Pattern: Many prompts emphasize depth by layering the scene into foreground, middle, and background, describing elements in each to add complexity. This layering creates a more immersive effect.

Example: "In the foreground, a solitary cherry blossom branch frames the scene; in the middle, a quaint bridge stretches over a serene pond, with a lush forest in the background."

4. Mood and emotional language

Pattern: Descriptions of mood or atmosphere, like "evokes nostalgia" or "a peaceful yet mysterious night," are common. These statements guide the AI to infuse the image with a certain emotional quality.

Example: "A peaceful, timeless ambiance envelops a quaint cottage under a soft, lavender twilight."

5. Specific artistic styles and camera techniques

Pattern: Some prompts specify art styles (like "Ukiyo-e") or photographic terms (like "low angle shot," "soft focus") to convey a particular visual style.

Example: "Captured in a soft focus, watercolor style, evoking the delicate hues of a gentle sunrise."

6. Utilization of cultural or geographic markers

Pattern: Location-specific markers, such as “Santorini” or “Maruyama Park, Kyoto,” are used to anchor the image in a real or imagined place, often adding authenticity.

Example: "A serene pagoda stands on the shores of a tranquil lake, surrounded by Kyoto’s iconic autumn maples."

7. Rich visual vocabulary and adjectives

Pattern: The use of vivid and varied adjectives like "luminous," "whimsical," "ethereal," etc., helps create striking mental images. This approach combines sensory and descriptive language for stronger impact.

Example: "Under a luminous silver moon, an ancient forest glows softly, its trees shrouded in ethereal mist."

5.2 Types of image generative AIs

Apart from studying image prompt structure, this research also aims at exploring how users generate image to know what kind of AIs often used by users. From the data shared by users, there are two main types of Image Generative AIs often used by users; Web-based AIs and Computer-based AIs.

5.2.1 Web-based AIs

The Web-based AI is the service provided by AI developers on websites or applications. This kind of service is convenient and easy to understand. Users can browse directly to the websites or install apps on their mobile devices. There are few options to select such as image ratios, image models (only on some websites or apps), and number of generated images. Then users can enter prompts in the textbox with human language. This method needs only a smartphone or an ordinary computer to browse the website or install an app to generate images.

From all sampled prompts, the Web-based AIs often used by users are Bing, Seaart, Microsoft Designer, Kling, Gemini, Mojen, Alisa, Leonardo, Bigbot, Claude, Microsoft Copilot and Varoriya accordingly. These services are all in a trial period as Image Generative AIs are still new to users and also developers. There are more online services apart mentioned but they might not be

known much among users. Also, many new Web-based AIs could be introduced in this year as the image generative AI is still in an early stage.

5.2.2 Computer-based AIs

Opposite to the Web-based AIs, the Computer-based AIs is much more complicate. Users need to install software, source codes, and image models on their own computers. It needs higher knowledges and understandings on computer skills because all softwares, source codes and image models are in a one-click installer. They must find each element from various sources. Moreover, this method needs high performance computer, especially high-spec graphic processor which is costly. Therefore, most users prefer the Web-based AIs rather than this method.

The image models widely accepted by users are Midjourney, Stable Diffusions, Flux and Dall-E. Each model is adept at different areas. For example, users aiming at generating human images tend to use Flux while using Midjourney to create object images rather than other image models. However, it is too soon to conclude that which model is better on what areas because all image models are continuously developed every second.

This analysis reveals sophisticated patterns in user prompt construction across different categories. While each category demonstrates unique characteristics, common elements such as detailed subject descriptions, environmental context, and technical specifications persist across categories. The preference for web-based platforms suggests a trend toward accessibility over technical complexity in image generation applications

6. Discussions

Image Generative AI is a powerful tool for creating images, enabling users to generate content through text input without requiring artistic expertise. However, the quality of the outputs heavily depends on prompt structure, model choice, reference precision, user knowledge, language barriers, and ethical concerns.

6.1 Importance of prompt structure

The detail and sequence of prompts directly impact image quality. Vague prompts give the AI more interpretive freedom, leading to varied results. For example, "a woman on the beach" could produce different physical attributes or settings depending on the AI's interpretation. The following images with vague prompt and well-structured prompt can show a significant difference.



Prompt: a woman on the beach



Prompt: a woman on the beach without people, sunset. She's wearing a wedding dress; her hair is black. She is smiling. The warm light shines to her face. Super-high definition.

Picture 1 Comparison of AI-generated images by vague prompt and structured prompt
Source: Generated by the author with AI on microsoft designer website

These two pictures were generated by the same AI model DALL-E but with different prompts. The first picture with vague prompt gave too much room for AI to generate. Without a setting description, AI created five people on the background. Additionally, AI generated swimming wear for the lady. Comparing to the right picture, specific setting, lighting, and facial expression could control the result as expected.

6.2 Lack of clear guidelines for prompts

Despite image generative AI's capabilities, outputs may not always align with user specifications, such as resolution or stylistic elements. This inconsistency arises from a lack of comprehensive guidelines, making trial and error a common practice for users. There is a need for shared resources to standardize effective prompt use.

6.3 Differences among image models

Different models are favored for specific tasks, such as Flux for human images and DALL-E for food images. These preferences are based on trial and error, as models continually improve by learning from prompts. Using identical prompts across models helps users identify the most suitable ones for consistent results.

6.4 Precision through image reference

Some platforms offer image references, which enhance precision by providing visual cues. This is particularly helpful for terms that are difficult to define through text alone. However, not all platforms offer this feature, limiting its availability.

6.5 Limitations in user knowledge and skills

Computer-based AI platforms require advanced technical knowledge, including software installations and hardware selections. This technical barrier makes it more difficult for users without specialized skills to effectively use the systems.

6.6 Quota limitations in image generation

Web-based AIs often provide limited free image generation quotas, with some platforms offering only a few images per day. Users must purchase credits or subscribe to access additional outputs, especially for beginners who require multiple attempts to achieve desired results.

6.7 Language barrier

Language proficiency is critical in prompt crafting, as most AI platforms are optimized for English. While some accept other languages, prompts in English tend to yield more accurate results, putting non-English speakers at a disadvantage.

6.8 Ethical considerations

Ethical concerns, such as copyright and ownership, arise from the use of vast image datasets. Generated images may resemble the works of artists or recognizable individuals, leading to disputes over ownership between users and developers.

7. Conclusion

This research highlights the early-stage potential of image-generative AI, emphasizing how effective prompt structures shape the quality of generated outputs. Analysis of over 1,000 prompts from the Facebook group Free Prompt Gen Image identified critical elements like subject specificity, environmental context, and stylistic precision, all contributing to high-quality image generation.

Findings reveal diverse user preferences and challenges across platforms. Web-based services offer accessibility, while computer-based models provide flexibility but require technical expertise. Ethical concerns, such as copyright and data ownership, underscore the need for transparent guidelines.

User skills, particularly in English and artistic knowledge, play a vital role in effective tool use. Challenges like language barriers, quota limits, and technical complexities highlight the need for more intuitive and inclusive interfaces.

Future research should focus on refining prompt methodologies, developing real-time refinement tools, advancing multimodal systems, and creating collaborative repositories to unlock the full potential of generative AI for innovative and ethical applications.

8. References

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014*).* Generative adversarial nets*.* In *NIPS'14: Proceedings of the 28th International Conference on Neural Information Processing Systems (*pp.2672-2680). Springer.

Ho, J., Jain, A., & Abbeel, P. (2020). *Denoising diffusion probabilistic models*. arXiv. https://arxiv.org/pdf/2006.11239

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). Sage Publications.

McCormack, J., Llano, M., Krol, S., & Rajcic, N. (2024). No longer trending on artstation: Prompt analysis of generative AI art. *EvoMUSART, 14633, 279-295. https://doi.org/10.1007/978-3-031-56992-0_18*

Pise, M., Yadgiri, N., Gaikwad, P., Dusawar, Y., & Nandanwar, P. (2024). AI image generator. *International Journal of Advanced Research in Science Communication and Technology (IJARSCT)*, *4*(4), 768-773.

Sakirin, T., & Kusuma, S. (2023). A survey of generative artificial intelligence techniques. *Babylonian Journal of Artificial Intelligence, 2023*, 10–14.

Srivastava, A. (2024). Advancements in text-to-image generation through generative AI. *International Journal of Scientific Research in Engineering and Management*, *08*(05), 1-5.

Yang, X., Zhang, J., & Liu, W. (2023). Prompt recommendations for AI art. *Journal of Artificial Intelligence and Art*, *15*(3), 45-67.

Zhang, Y., & Liu, W. (2024). *Boosting architectural generation via prompts: Report.* Scispace. https://scispace.com/pdf/boosting-architectural-generation-via-prompts-report-2nzs41gf0r.pdf

Zhou, E., & Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS Nexus, 3*(3), 1-8.