Prompt Engineering in Diffusion Models: A Research Perspective

Based on the research paper: [Investigating Prompt Engineering in Diffusion Models](#)

Methodology:
Textual prompts are dissected into two core elements: the tangible content (e.g., a mainecoon cat kneeling) and the stylistic execution (e.g., lighting, portrait style vs. landscape style). The CLIP model generates a 768-dimensional embedding for each of its 77 tokens, influencing image output.

Experiments & Observations:
Linguistic elements—adjectives, nouns, and proper nouns—distinctly impact image generation. Adjectives slightly tweak images, while nouns significantly alter content. Artist styles, like "in the style of Leonardo da Vinci," drastically change image composition.

Experiments:
Modifiers like repeated words, lighting phrases, and artist styles were tested against a base prompt to observe their effects.
- Repeated words had minimal impact, akin to adjusting weight but not always as expected.
- Descriptors influenced images less than nouns, artist names, or lighting cues.
- Lighting phrases varied in effect, from subtle to substantial changes in content and mood, often altering the background.

Detailed Findings:
Influence of Artist Names:
- Adding artist names to prompts changes the image greatly, affecting style, medium, and mood.
- Changes include switching from photo-like images to artistic styles like oil paintings or watercolors.
- The color palette adjusts to match the artist's typical use of colors, setting a unique mood.
- The image might reflect the artist's common themes or the era they belong to, affecting subjects and overall layout.

Impact of Prompt Ordering:
- The order of words in a prompt matters; earlier words have more influence on the image.
- Start with the most important details to make sure they're emphasized in the final image.
- Later details in the prompt have less impact, so prioritize your prompt's structure for the best effect.
- This ordering lets you control what the model focuses on, helping to shape the image's main theme or focus.

Guidance for choosing your prompts where each stage builds on the previous ones:

1. Start with a clear noun-based statement that contains the main subject you wish to have in the image.

2. Be systematic about recording what seeds are effective, since it is easier to iterate on images based on a firm foundation.

3. The properties of the Stable Diffusion model allow it to be very good at replicating the style of various artists. Because of this, it is often beneficial to look up artists and key styles that you would like to emulate and add that to your prompt.

4. Finally experiment with descriptors that you wish to add, considering the importance of (for instance) having a simple or a busy/intricate background. This can be done through adding lighting effects phrases and repeating words.

Ultimately the process of "Prompt Engineering" is an artistic endeavor and the techniques required for success must be honed over time, using experimentation, imagination and creativity.

Prompt Engineering Guide:

Step	Component	Description	Impact
1	Seeding & Iteration	Use specific seeds to ensure consistency and reproducibility of results. Document effective seeds for future reference, allowing for systematic experimentation and refinement of images. Iteration based on successful seeds can help in fine-tuning the prompt to achieve the desired outcome with greater precision.	Seeding ensures that the randomness inherent in the image generation process is controlled, allowing for the reproduction of specific images or styles. Iteration with documented seeds enables the refinement of prompts and the exploration of subtle variations, contributing to a more deliberate and controlled prompt engineering process.
2	Main Subject (Noun)	Start with a clear, noun-based statement that specifies the main subject of the image. This sets a concrete foundation for the image content.	Nouns introduce new content and dramatically shape the image, establishing the primary focus.
3	Ordering	The sequence in which elements are introduced in the prompt can influence their prominence in the generated image. Items mentioned earlier tend to have a greater impact, while those listed later may be less influential. This flexible approach allows for nuanced control over the emphasis of elements.	The order of prompt elements can subtly shift the focus and hierarchy within the image, with earlier-mentioned aspects receiving more emphasis, offering control over the composition and focus areas of the generated image.
4	Artist Style	Incorporate the name of an artist or a specific art style to infuse the image with a distinct aesthetic. This can include references to art mediums, color palettes, and stylistic elements typical of the artist's work.	Adding an artist's name or style can significantly alter the image across multiple dimensions, including medium, color palette, and overall aesthetic, closely emulating the referenced style.
5	Descriptors	Experiment with adjectives or adjectival phrases to describe qualities, moods, or additional details about the subject or scene. Be cautious with the choice and quantity, as their impact varies.	Descriptors generally have a smaller impact compared to nouns or artists, subtly refining the image by adding details or modifying existing elements.
6	Lighting Effects	Use phrases that describe lighting conditions, such as "ambient lighting" or "volumetric lighting." The impact of lighting descriptors can range from minimal to significant, affecting the mood, appearance of the subject, and the background.	Lighting can either slightly adjust the image's mood and aesthetics or bring substantial changes to the content and background, depending on the specificity and nature of the lighting description.
7	Repeated Words	Integrate repeated words to emphasize certain elements or qualities. The effect of repetition can be subtle and may not always align with expectations.	Repeated words may act as a form of weighting, potentially emphasizing aspects of the image, but their impact can be minimal or unpredictable. With extensive repetition, significant alterations, such as background removal or detail loss, may occur.