Open in app ↗                                                                    Sign up          Sign In

◖◗❙

tds    Published in Towards Data Science

Leonie Monigatti    Follow

Sep 20, 2022 · 9 min read · ✨ · ▶ Listen

☐⁺ Save    🐦    f    in    🔗

# A Beginner's Guide to Prompt Design for Text-to-Image Generative Models

Learn these prompt engineering tricks before you waste your free trial credits



"a cat wearing sunglasses"

"a black and white photograph of a cat wearing sunglasses by annie lebovitz, highly-detailed"

If you have already played around with a text-to-image generative model, you know how difficult it is to produce an image you like.

👏 114    💬 1

With the release of <u>Stable Diffusion</u>, <u>Midjourney</u>, and <u>DALL·E2</u>, people have been saying that prompt engineering could become a new profession. Because DALL·E2, the Midjourney Discord server, and StabilityAI's DreamStudio have a credit-based pricing model [3,5,7], users are incentivized to use as few prompts as possible to get an image they like.

## Users are incentivized to use as few prompts as possible.

This article will give you a quick guide to prompt engineering before you waste all your free trial credits. This is a general guide, and there are differences between DALL·E2, Stable Diffusion, and Midjourney. Therefore, not all tips might apply to the specific generative model you are using.

We will use the base prompt **"a cat wearing a pair of sunglasses"** similarly to [11]. The images will be produced with DreamStudio (GUI for Stable Diffusion) with the **default settings and a fixed seed of 42** to generate similar-looking images for comparison.

For more inspiration on prompt engineering, you can have a look at <u>https://lexica.art/</u>, which is a collection of prompts and their resulting images produced with Stable Diffusion.

## Fundamentals of Prompt Design for Text-to-Image and Text-Guided Image-to-Image Generation

Currently, most generative models are either text-to-image or text-guided image-to-image generative models. In both cases, at least one input is a prompt, which is a description of the image you want to generate.

### Prompt Length

The prompt should be relatively short. While Midjourney allows up to 6000 characters, prompts should **stay under 60 words** [6]. Similarly, prompts for DALL·E2 must **stay under 400 characters** [9].

### Character Set

From a statistical point of view, your best bet is to **phrase your prompt in English.** E.g., Stable Diffusion was trained on a subset of the LAION-5B database, which contains 2.3 billion English image-text pairs and 2.2 billion image-text pairs from 100+ other languages [1, 4].



Prompt: "a cat wearing sunglasses" (Image made by the author with DreamStudio).

That means you are not limited to the Western European alphabet. You can use non-Roman character sets like Arabic or Chinese, and you can even use emojis.

Prompt: "サングラスをかけた猫" (Japanese for "a cat wearing sunglasses") (Image made by the author with DreamStudio)

Prompt: "🐱😎" (Image made by the author with DreamStudio)

However, as you can see, both the image generated with a Japanese prompt as well as the image generated with an emoji only prompt **fail to produce a pair of sunglasses on the cat.**

While it might not work as well as English prompts, you can use it for enhancement (see section Repetition).

Also, e.g., Midjourney is **not case-sensitive** [6]. That means whether you capitalize your text does not impact the generated image; therefore, you can write your prompt in lowercase.

### Template and Tokenization

A prompt usually follows the following template (adjusted from [8]). We will get to each part in the following sections.

```
[Art form] of [subject] by [artist(s)], [detail 1], ..., [detail n]
```

Tokenization in the context of prompt engineering describes the separation of a text into smaller units (tokens). For prompt engineering, you can use commas ( , ), pipes ( | ), or double colons ( :: ) as hard separators [6, 10]. However, the direct impact of tokenization is not always clear [6].

## 1. Subject

The most important part of a prompt is the subject. [2, 8] What do you want to see? While this might be the most straightforward, it is also the most difficult regarding the amount of detail you want to provide.



Prompt: "a cat wearing sunglasses" (Image made by the author with DreamStudio)

### Plurals

Vague plural words like "cats" leave a lot of room for interpretation [6]. Did you mean two cats or 13 cats? Therefore, when you want multiple subjects, use plural nouns with specific numbers [6].

Prompt: "cats wearing sunglasses" (Image made by the author with DreamStudio)

However, it was reported that while, e.g., DALL·E2 has no problem creating multiple subjects in a scene, it falls short in separating certain characteristics of each from each other [11].

While the above image generated with Stable Diffusion's DreamStudio produced two separate cats, it shows its struggles in the following image. You can see that the cat on the left is not wearing sunglasses. Instead, the pair of sunglasses seems to be floating behind the cat.

Prompt: "three cats wearing sunglasses" (Image made by the author with DreamStudio).

Also, it was reported that DALL·E2 can handle prompts with up to three subjects well, but prompts with more than three subjects are difficult to create even if you say "12", "twelve", "a dozen", or say it multiple times in multiple ways [6].

Again Stable Diffusion is showing a difference to DALL·E2 regarding this issue. However, it also shows that generating exactly 12 cats is difficult.

Prompt: "twelve cats wearing sunglasses" (Image made by the author with DreamStudio)

### Weights

If you want to give a specific subject a heavier weight, there are various ways to do so.

1. **Order:** Tokens near the front of a prompt are weighted more heavily than the tokens in the back of a prompt. [10]

2. **Repetition:** Repeating the subject by phrasing it differently can impact its weighting [8, 12]. I have also seen prompts repeating the subject in different languages or using emojis.

3. **Parameters:** E.g., in Midjourney, you can suffix any part of a prompt with `::weight` to give it a weight (e.g. `::0.5` ) [6].

### Exclusions

Prompts containing negative words like "not", "but", "except", and "without" are difficult for the text-to-image generative models to understand [6]. While Midjourney has a special command for cases like this ( `--no` ) [7], you can bypass this

issue by **avoiding negative phrasing** and instead positively phrasing your prompt [6].

## 2. Art Form

The form of art is a crucial part of the prompt. Commonly used art forms in prompts are [2]:

- **photography:** studio photography, polaroid, camera phone, etc.
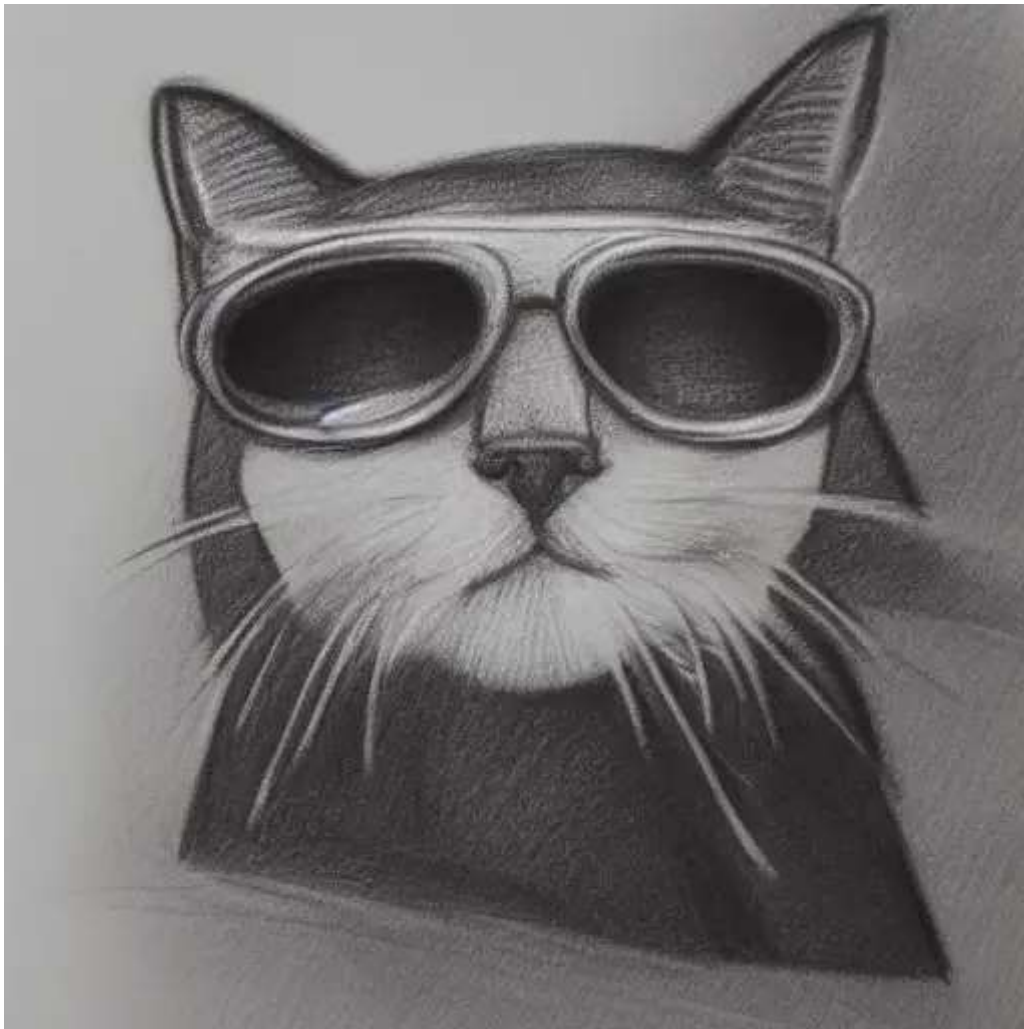


Prompt: "polaroid photo of a cat wearing sunglasses" (Image made by the author with DreamStudio)

- **paintings:** oil paintings, portraits, watercolor paintings, etc.

Prompt: "watercolor painting of a cat wearing sunglasses" (Image made by the author with DreamStudio)

- **illustrations:** pencil drawing, charcoal sketch, etching, cartoon, concept art, posters, etc.

Prompt: "charcoal sketch of a cat wearing sunglasses" (Image made by the author with DreamStudio)

- **digital art:** 3D renders, vector illustrations, low poly art, pixel art, scan, etc.

Prompt: "vector illustration of a cat wearing sunglasses" (Image made by the author with DreamStudio)

- **film stills:** movies, CCTV, etc.

Prompt: "CCTV still of a cat wearing sunglasses" (Image made by the author with DreamStudio)

As you can see, you can even define the specific medium for each art form. E.g., for photography, you can become very specific by defining details like [9]:

- film type (black & white, polaroid, 35mm, etc.),

- framing (close up, wide shot, etc.),

- camera settings (fast shutter speed, macro, fish-eye, motion blur, etc.),

- lighting (golden hour, studio lighting, natural lighting, etc.)

There are various other art forms like stickers and tattoos [11]. For more inspiration, you can have a look at [11].

If the art form is not specified in the prompt, the generative models will usually choose one it has seen the most during training. For many subjects, that art form will be photography [6].

## 3. Style or Artists

Another part of the template that can heavily impact the outcome of the generated image is the style or the artist [6, 8]. Simply use "by [artists]" [11] or "in the style of [style or artist]".



Prompt: "oil painting of a cat wearing sunglasses by van gogh" (Image made by the author with DreamStudio)

Two tips for generating interesting images are:

- Mixing two or more artists [2]

Prompt: "oil painting of a cat wearing sunglasses by van gogh and by andy warhol" (Image made by the author with DreamStudio)

- Using fictional artists [12]

Prompt: "oil painting of a cat wearing sunglasses by max mustermann" (Image made by the author with DreamStudio)

## 4. Combining Features

On the note of combining artists to generate interesting images, you can also combine two well-defined concepts [6]. You can try out the following templates [11]:

```
- "[subject] made of"
- "[subject] that looks like"
- "[subject] as"
```

Prompt: "a cat as a rockstar" (Image made by the author with DreamStudio)

## 5. Adjectives and Quality Boosters

Adding details like adjectives and quality boosters can significantly impact the overall aesthetic of your image [8].

Commonly used adjectives usually describe:

- the framing (close up, landscape, portrait, wide shot, etc.)

- the color scheme (dark, pastel, etc.)

- the lighting (cinematic lighting, natural light, etc.)

- other: epic, beautiful, awesome

But there are also some "magic terms" the community has already found that seem to generate better-looking images [2, 8]:

- "highly-detailed"

Prompt: "a cat wearing sunglasses, highly-detailed" (Image made by the author with DreamStudio)

- "trending on artstation"

Prompt: "a cat wearing sunglasses, trending on artstation" (Image made by the author with DreamStudio)

- "rendered in Unreal Engine"

Prompt: "a cat wearing sunglasses, rendered in unreal engine" (Image made by the author with
[DreamStudio](#))

- "4k" or "8k"

## Conclusion

In this article, you learned how to design a prompt to produce images with text-to-image generative models in fewer tries.

We discussed how you could improve an acceptable-looking image from a prompt that only contained the subject like "a cat wearing sunglasses".

Prompt: "a cat wearing sunglasses" (Image made by the author with DreamStudio).

The essential tricks were:

- defining a fine-grained form of art (e.g., black and white photograph)

- adding a style or artist (e.g., by Annie Lebovitz)

- adding boosting adjectives (e.g., highly-detailed).

By following these simple tricks, the resulting image already looks much more interesting, as you can see below.

Prompt: "a black and white photograph of a cat wearing sunglasses by annie lebovitz, highly-detailed"
(Image made by the author with DreamStudio)

## Enjoyed This Story?

*Here is a collection of my other Generative AI articles:*



*If you'd like to get my new stories directly to your inbox, subscribe!*

*Become a Medium member to read more stories from other writers and me. You can support me by using my underline*referral link underline *when you sign up. I'll receive a commission at no extra cost to you.*

---

**Join Medium with my referral link — Leonie Monigatti**

Read every story from Leonie Monigatti (and thousands of other writers on Medium). Your membership fee directly...

medium.com

---

*Find me on* <u>Twitter,</u> <u>LinkedIn</u>, *and* <u>Kaggle</u>!

## References

[1] R. Beaumont, "LAION-5B: A NEW ERA OF OPEN LARGE-SCALE MULTI-MODAL DATASETS", laion.ai. <u>https://laion.ai/blog/laion-5b/</u> (accessed September 10, 2022)

[2] DreamStudio, "Prompt Guide". dreamstudio.ai. <u>https://beta.dreamstudio.ai/prompt-guide</u> (accessed September 10, 2022)

[3] DreamStudio, "General Questions". dreamstudio.ai. <u>https://beta.dreamstudio.ai/faq</u> (accessed September 5, 2022)

[4] Huggingface, "Stable Diffusion with 🧨 diffusers", google.com. <u>https://colab.research.google.com/github/huggingface/notebooks/blob/main/diffusers/stable_diffusion.ipynb#scrollTo=gd-vX3cavOCt</u>

[5] J. Jang, "How DALL·E Credits Work". openai.com. <u>https://help.openai.com/en/articles/6399305-how-dall-e-credits-work</u> (accessed September 4, 2022)[9] Stability AI, "Stable Diffusion Dream Studio beta Terms of Service". stability.ai. <u>https://stability.ai/stablediffusion-terms-of-service</u> (accessed September 5, 2022)

[6] Midjourney, "docs", github.com. <u>https://github.com/midjourney/docs/</u> (accessed September 10, 2022)

[7] Midjourney, "Midjourney Documentation". gitbook.io.
https://midjourney.gitbook.io/docs/ (accessed September 4, 2022)

[8] J. Oppenlaender, A Taxonomy of Prompt Modifiers for Text-To-Image Generation
(2022), *arXiv preprint arXiv:2204.13988*.

[9] G. Parsons, The DALL·E 2 Prompt Book (2022), https://dallery.gallery/the-dalle-2-
prompt-book/ (accessed September 10, 2022)

[10] "pxan", "How to get images that don't suck: a Beginner/Intermediate Guide to
Getting Cool Images from Stable Diffusion", reddit.com.
https://www.reddit.com/r/StableDiffusion/comments/x41n87/how_to_get_images_th
at_dont_suck_a/ (accessed September 10, 2022)

[11] "rendo1#6021" and "luc#0002", "DALL·E 2 Prompt Engineering Guide",
google.com.
https://docs.google.com/document/d/11WlzjBT0xRpQhP9tFMtxzd0q6ANIdHPUBkM
V-YB043U/edit#heading=h.8g22xmkqjtv7 (accessed September 10, 2022)

[12] M. Taylor, "Prompt Engineering: From Words to Art", saxifrage.xyz.
https://www.saxifrage.xyz/post/prompt-engineering (accessed September 10, 2022)

Artificial Intelligence      Prompt Engineering      Editors Pick      Hands On Tutorials

Stable Diffusion

## Enjoy the read? Reward the writer. [Beta]

Your tip will go to Leonie Monigatti through a third-party platform of their choice, letting them know you appreciate
their story.

Give a tip

# Sign up for The Variable

By Towards Data Science

Every Thursday, the Variable delivers the very best of Towards Data Science: from hands-on tutorials and cutting-edge research to original features you don't want to miss. Take a look.

By signing up, you will create a Medium account if you don't already have one. Review our Privacy Policy for more information about our privacy practices.

 Get this newsletter

About    Help    Terms    Privacy

Get the Medium app