



EXPLORING THE VISIONARY ART CREATIVE POTENTIAL OF DALL·E 2 WITH PROMPT ENGINEERING: BRIDGING THE GAP BETWEEN TEXT TO IMAGES

XIAOXIIN WANG

DR. AN'S AI GROUP

APR 15ST, 2023





AGENDA

01
INTRODUCTION

02
QUICK START METHOD

03
AUXILIARY
TOOLS

04
FUTURE
DIRECTIONS

05
REOURCES AND
REFERENCES

Preface

Anime boy, black hair, black eyes, head tilted to the right, surprised with a big mouth, white background, wallpaper, Hayao Miyazaki style



All the pictures you are about to see are *not real* except for the principle architectures, they are created through AI.

- ⚠ photos that are not real photos
- ⚠ paintings that are not real paintings, and
- ⚠ people, places and things that do not exist.

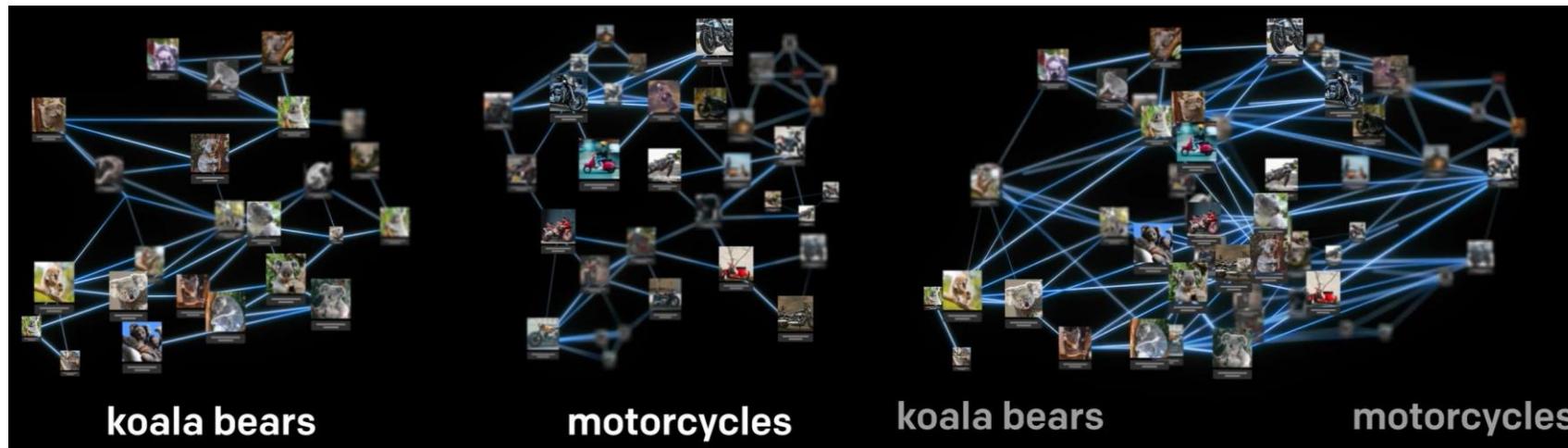
Perhaps some of them may be familiar to you, but it could be that the DALL·E 2 user used a certain artist or a certain scene's prompt.

01 AI INTRODUCTION



1.1 DALL·E

DALL·E is an artificial intelligence (AI) model developed by OPENAI, a research organization co-founded by Elon Musk and other prominent tech figures, and it is known for its ability to generate incredibly realistic images from textual descriptions.



1.2 DALL·E research

DALL·E research

1. Help people express themselves visually
2. Knowing whether the system understands text through AI images
3. Helps humans understand how AI systems see and understand our world

1.3 DALL·E 2

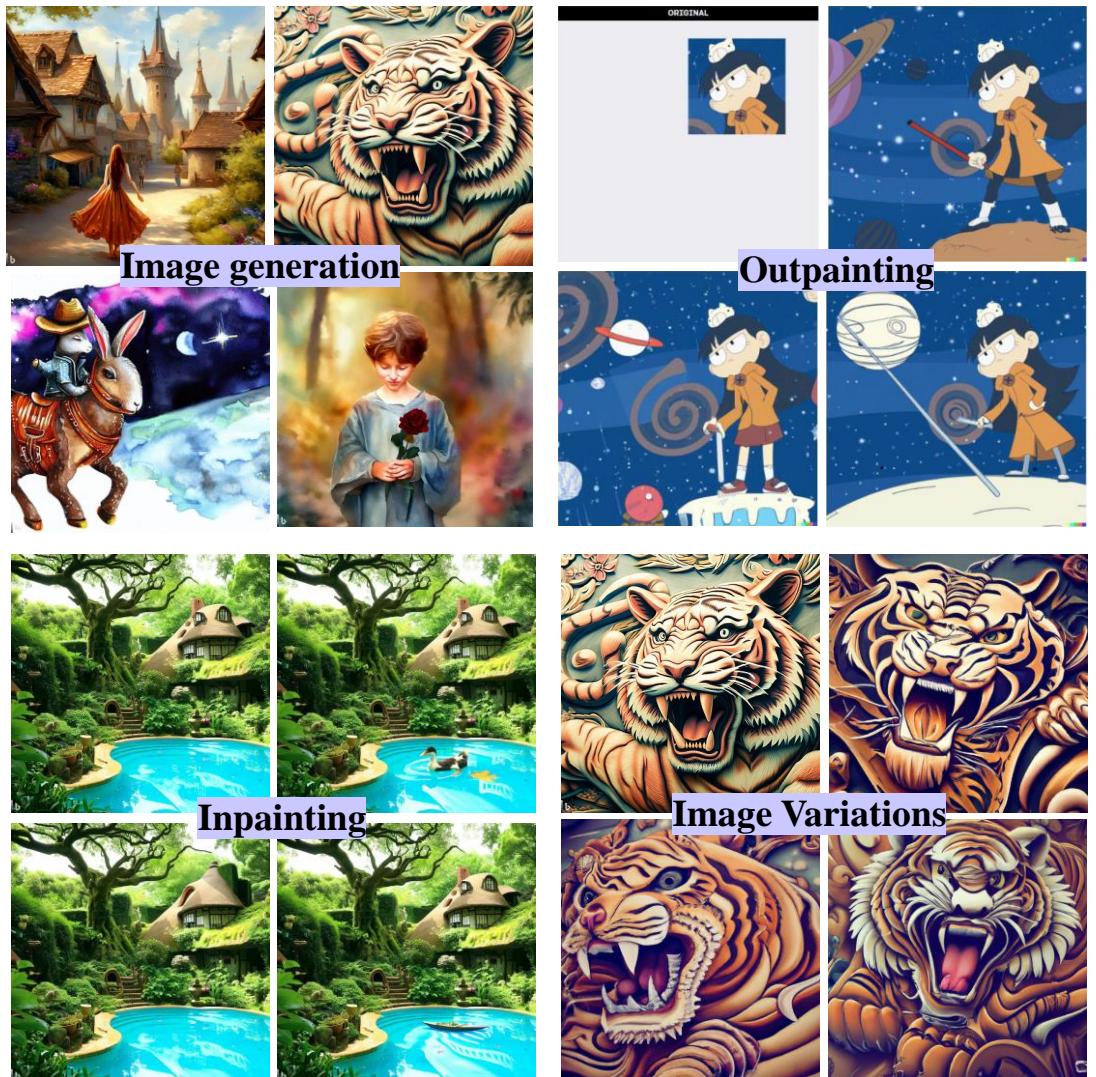
DALL·E 2 is an AI system that can create realistic images and art from a description in natural language.



DALL·E 1



DALL·E 2



1.4 DALL·E VS DALL·E 2

DALL·E 2 is preferred over DALL·E 1 when evaluators compared each model.

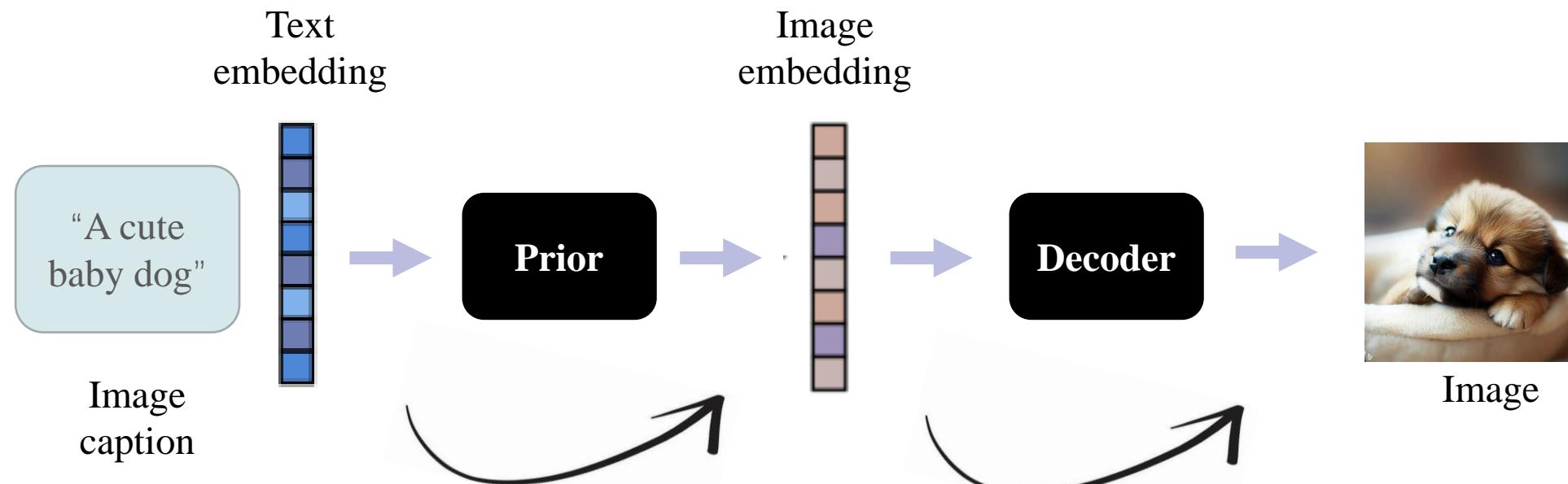
71.7%

Preferred for caption matching

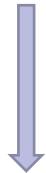
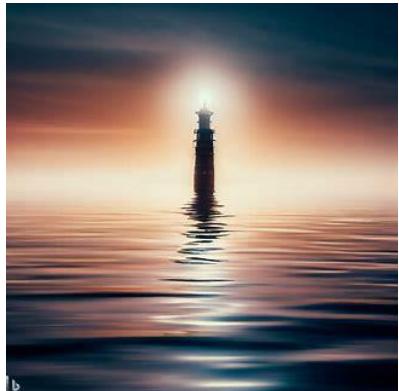
88.8%

Preferred for caption matching

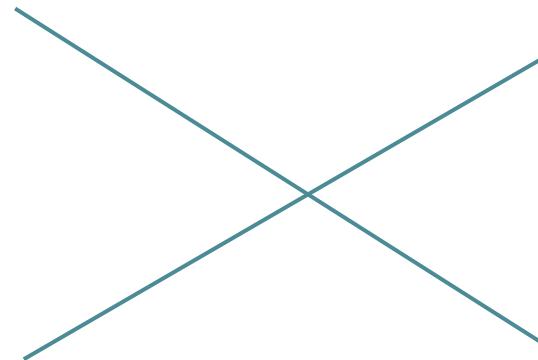
1.5 How DALL·E 2 Actually Works



1.5 How DALL·E 2 Actually Works



“Sea lighthouse with
reflections visible”



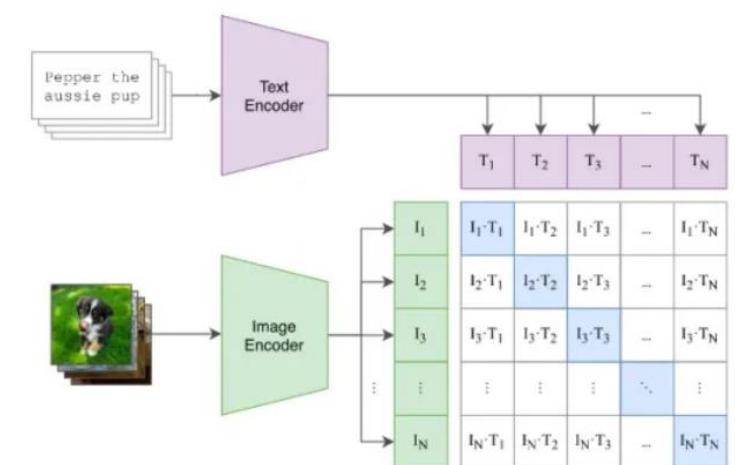
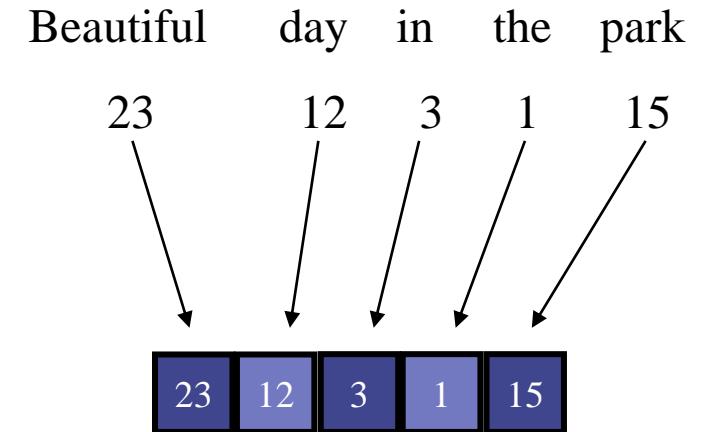
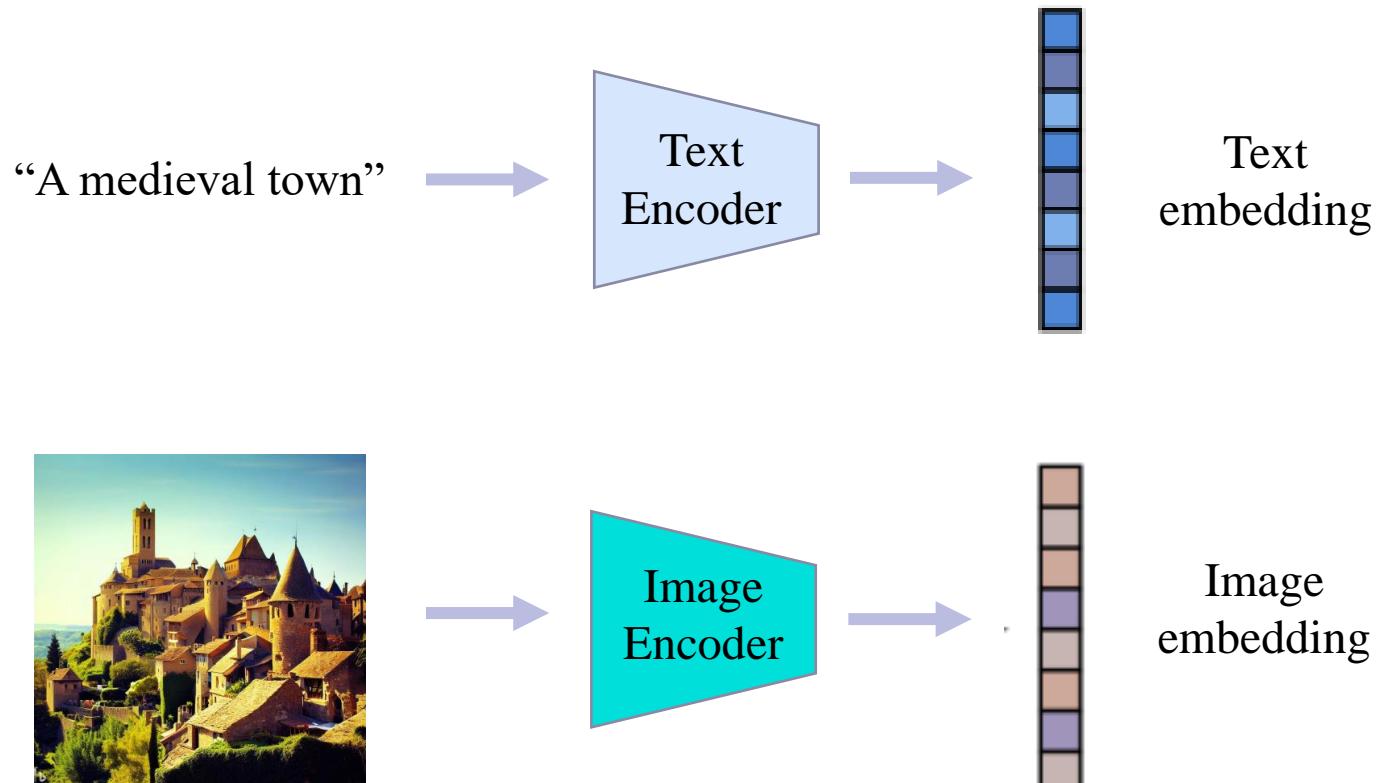
A medieval town

A cute baby dog

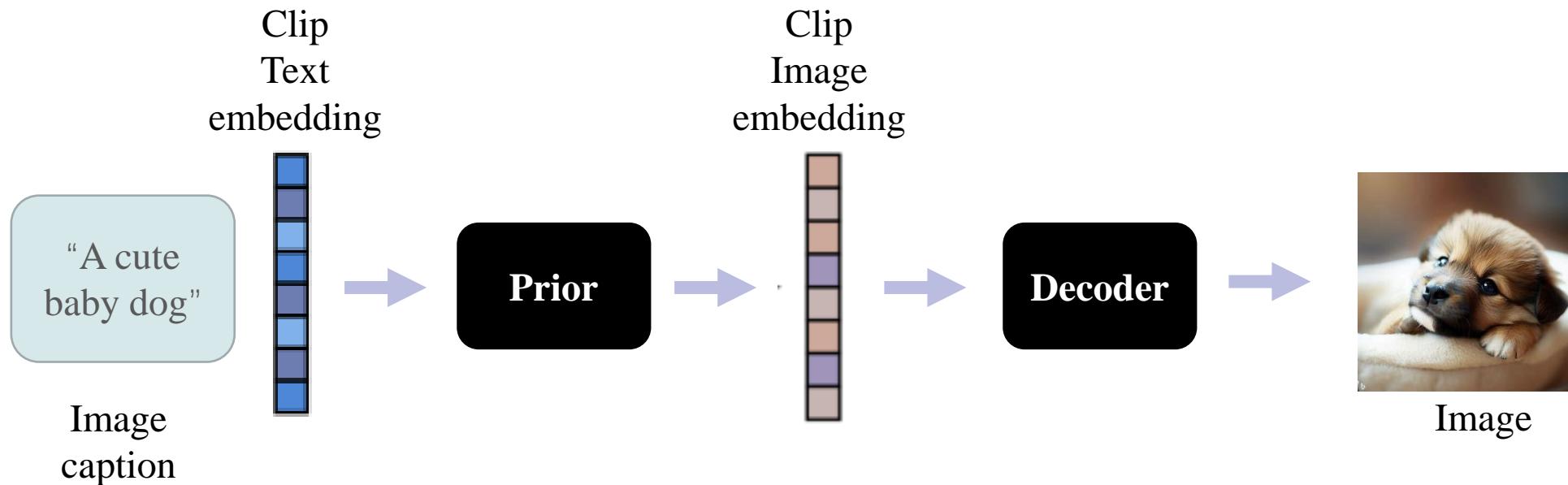


Sea lighthouse with
reflections visible

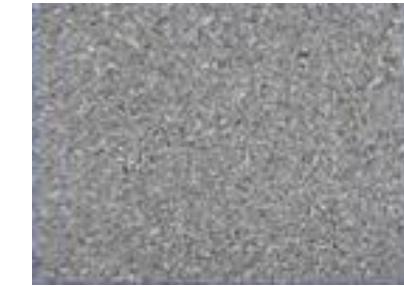
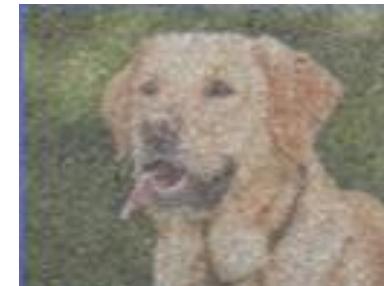
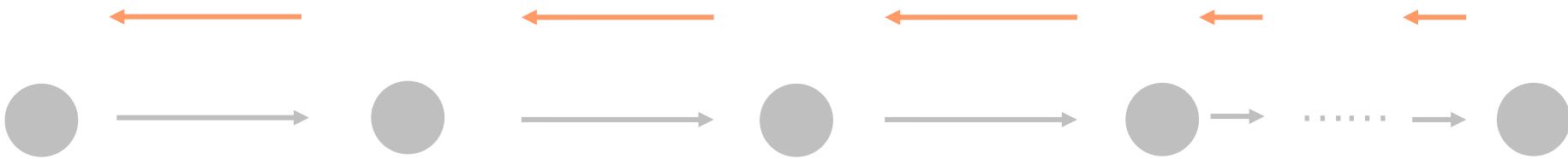
1.5 How DALL·E 2 Actually Works



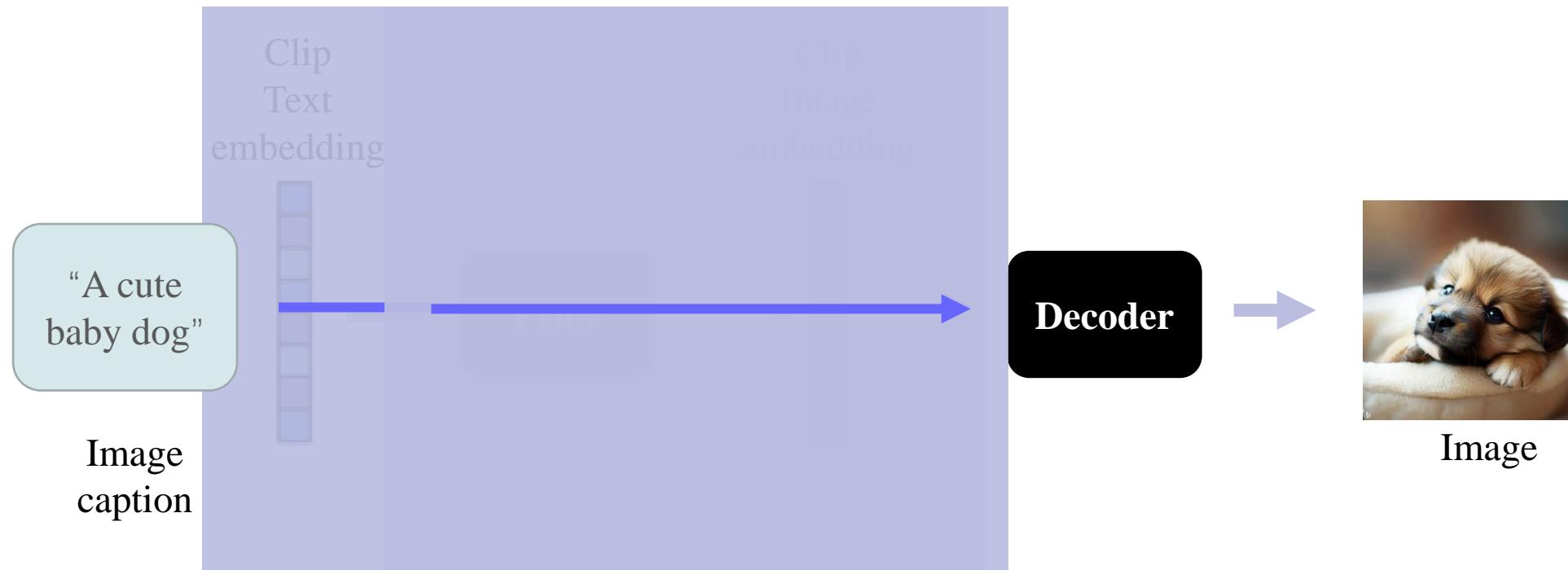
1.5 How DALL·E 2 Actually Works



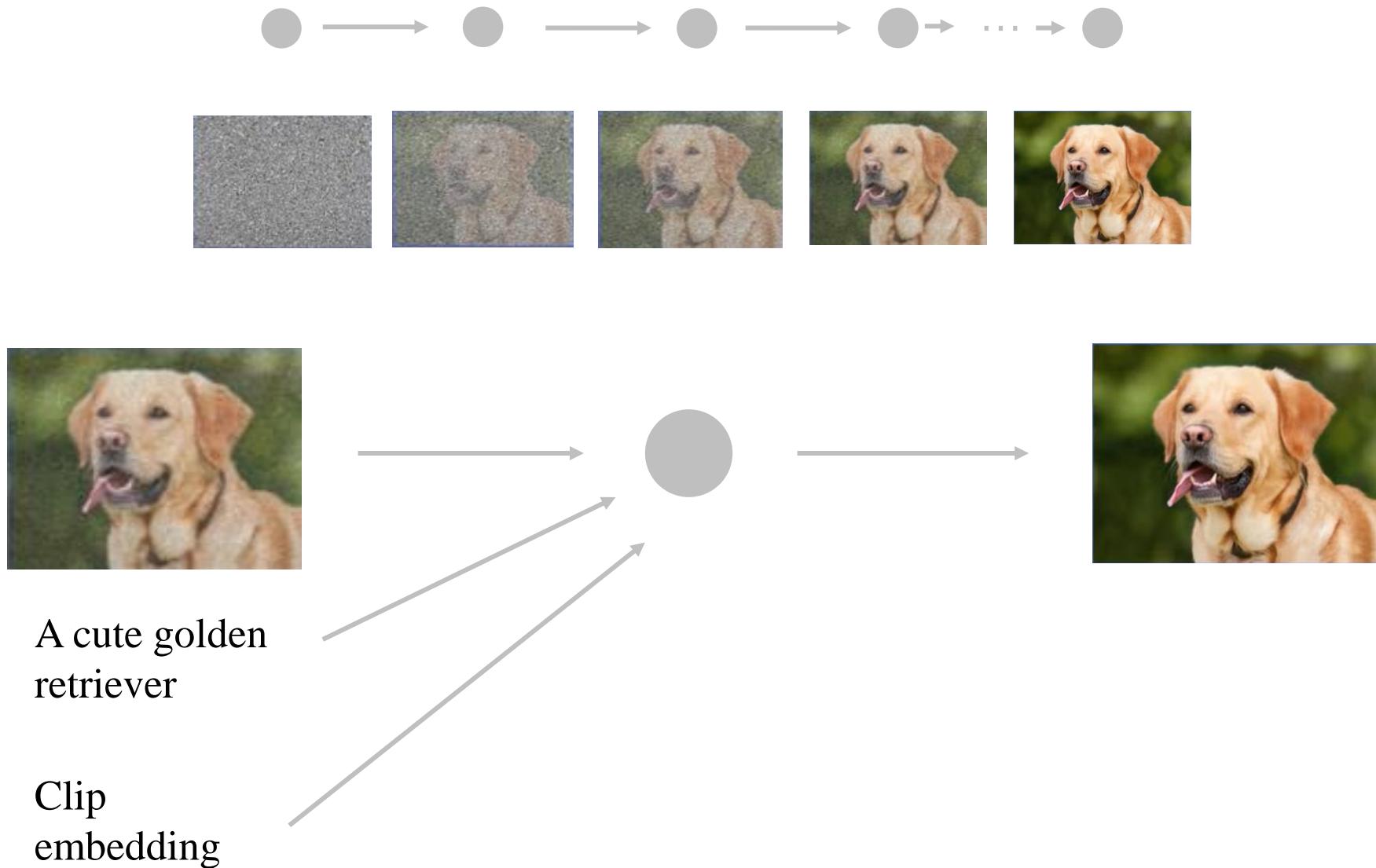
1.5 How DALL·E 2 Actually Works



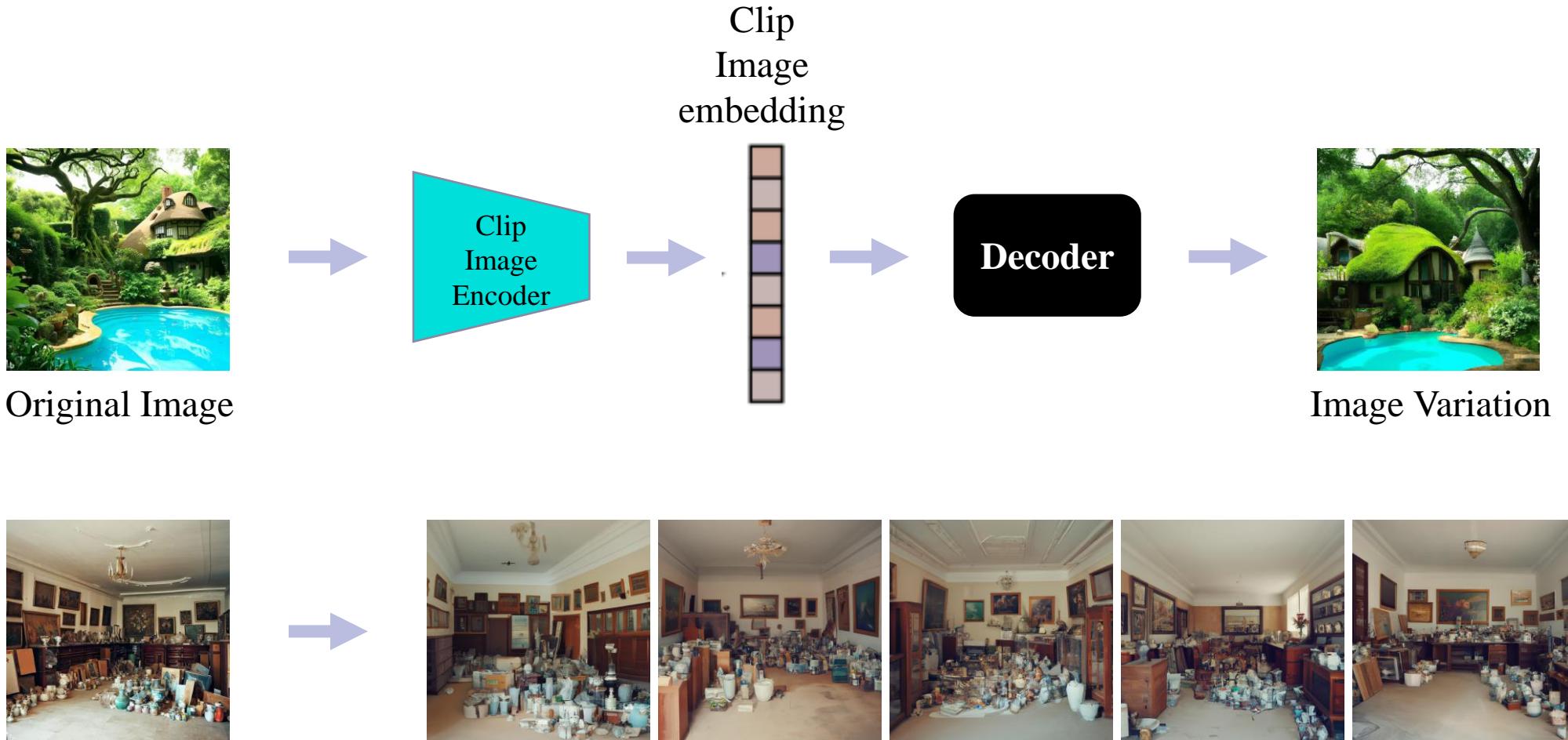
1.5 How DALL·E 2 Actually Works



1.5 How DALL·E 2 Actually Works



1.5 How DALL·E 2 Actually Works



1.6 Prompt engineering

- Prompt engineering aims to incorporate human intuition, creativity, and insights into the machine learning models.
- Prompt engineering is plays a crucial role in text-to-image creation because the generated image will depend heavily on the quality and specificity of the prompt.
- A well-crafted prompt can improve the performance of the image generative model and enhance the user experience.





02
05

QUICK START METHOD

2. Prompt

AI robot, sitting at a computer desk, operating the computer, with a white background



- DALL·E 2 has not been taught anything, it operates by relying on its study of images & captions to generate its own conclusions.
- There is no *regular guidance manual* so far, as even developers may not be sure what response DALL·E 2 will give when facing phrases.
- It is difficult to evaluate the specific impact of individual prompts on the final product.
- Then, if you are interested in DALL·E 2 and want to ‘discover’ its abilities and reactions. Let's start with the prompt template.

2.1 Prompt Design

A **prompt** is a sentence, 400 characters or less, that describes the image you want. **Prompts can be very long – or very short!**



A medieval town



A medieval town, digital



A medieval town, digital, illustration, detailed



A medieval town, digital, illustration, detailed, 4k, trending on artstation, fantasy



A medieval town, digital, illustration, landscape, detailed, 4k, trending on artstation, fantasy, style of David Finch



A medieval town, digital, illustration, landscape, detailed, 4k, trending on artstation, fantasy, style of David Finch, concept art, matte painting, highly detailed, rule of thirds, dynamic lighting, cinematic, detailed, denoised, centered

2.1 Prompt Design

- For DALL·E 2, prompts can't be more than 400 characters, in any case.
- The more detailed your prompt, the image may be closer to your expectations.
- Try to describe your idea using simple words and guide AI step by step to create the image you want.



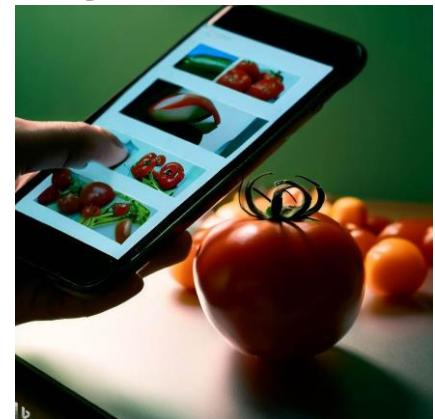
a tomato on Amazon



a tomato picture in an e-commerce platform



a tomato picture in Amazon, Amazon is an e-commerce platform



A photo of a tomato is displayed on an online sales platform, which is a place to sell various items. Electronic devices such as mobile phones and computers can browse product images and make purchases anytime and anywhere

2.1 Prompt Design

- But sometimes, describing too much may not necessarily be useful. A specific phrases may replace many descriptions.



A beautiful girl, Clockwise from top left are style of Picasso, Van Gogh, Hayao Miyazaki and Andy Warhol



Sea lighthouse with reflections visible, sunset, *lighthouse located in the middle of the picture*

Sea lighthouse with reflections visible, sunset, *central composition*

- Familiarity with art fields such as painting, sculpture, drama, photography, film, architecture, literature, etc. may be of great help to your own prompt.

2.1 Prompt Design



A little girl wearing a blue skirt and a straw hat is chasing butterflies on the grass, *Anime, light, fresh, romantic, peaceful, close to nature*



A little girl wearing a blue skirt and a straw hat is chasing butterflies on the grass, *Hayao Miyazaki style*

- DALLE is like a blind box, even with the same prompts, different results will be obtained.
- Imagination + Ability to relate knowledge + Accurate expression ability=Ideal generation results.
- DALL·E also has a creative eye, and has studied over 400 million images. So if you don't have specific ideas, there is nothing wrong with being vague, and seeing what happens!

2.2 Prompt Templates

Prompt:

Main body + Detail settings + Modification part

A word or phrase

Characteristic

Color



A little boy holding a rose

A short sentence

Structure

Time



A little boy holding a rose,
the background is a forest,
colorful, Renaissance
period

A short story

Background

Light



A little boy holding a rose,
the background is a forest,
colorful, Renaissance
period, crepuscular ray

...

Emotion

Type



A little boy holding a rose,
the background is a forest,
colorful, Renaissance
period, crepuscular ray

...

Style

Resolution ratio

...



A little boy holding a rose,
the background is a forest,
colorful, Renaissance
period, crepuscular ray



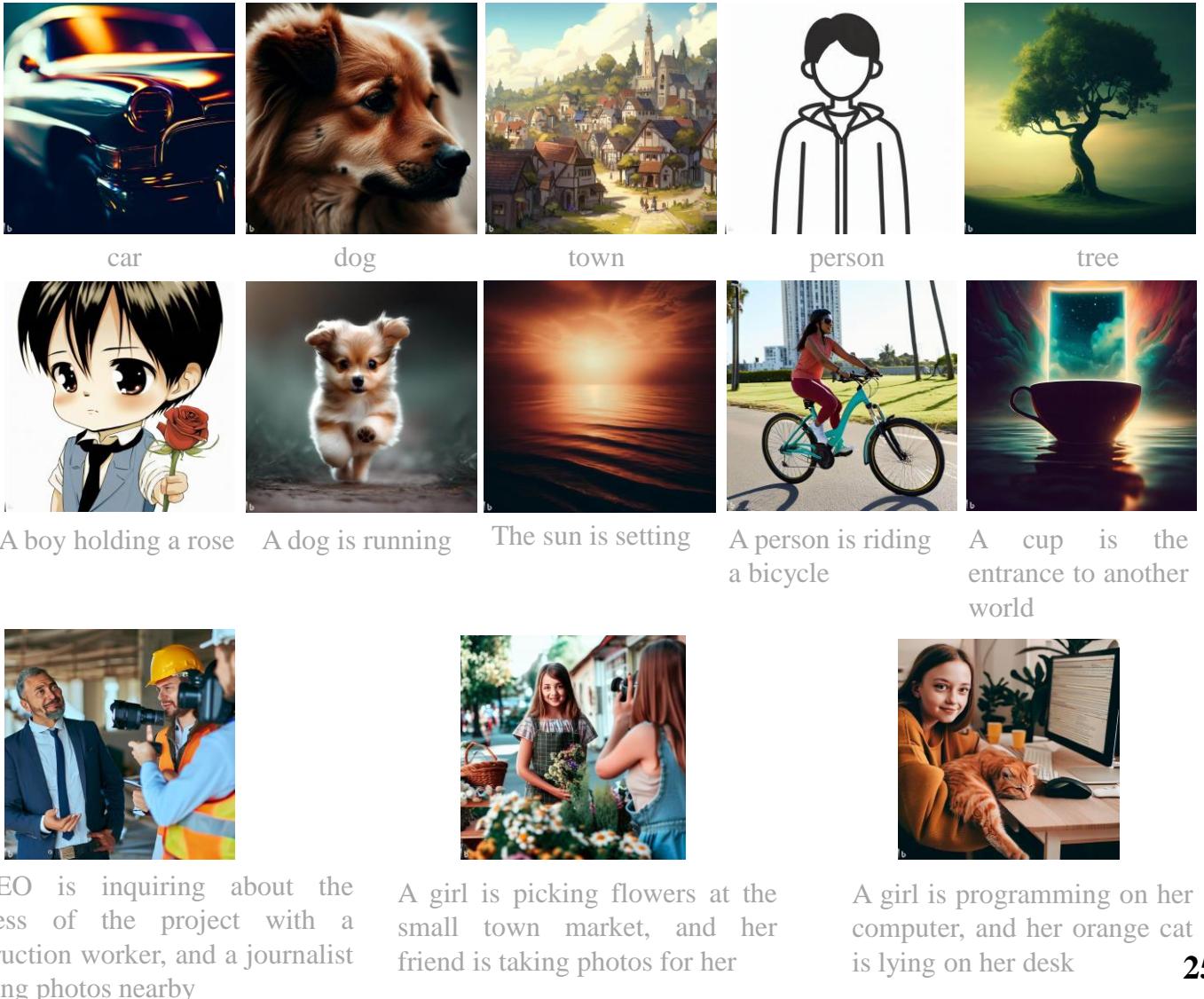
A little boy holding a rose,
the background is a forest,
colorful



A little boy holding a rose,
the background is a forest,
colorful, Renaissance
period, crepuscular ray,
watercolor painting

2.2.1 Main body

- A word or phrase: It's a specific image, such as dog, car, town, person...
- A short sentence : A sentences describing the subject's actions, expressions, states...
- A short story : A short story describes a scene as it may involve interaction between multiple subjects...



2.2.2 Detail settings

➤ Characteristic

Describe the characteristics of the subject, such as size, color, shape, texture, number, etc.



The girl is holding lemons

The girl is holding bright yellow fresh lemons in her hands

2.2.2 Detail settings

➤ Structure

Structure is a consideration and arrangement of the overall content and form of the image.



A group of ants,
circular composition



A tall building,
vertical composition



A lighthouse at sea,
rule of Thirds



Glacier Snow Mountain,
mirror composition

- *Vertical(horizontal) Composition*
- *Centred Composition and Symmetry*
- *Diagonal Lines*
- *Rule of Thirds*
- *Circular Composition*
- *Mirror composition*
- *Radioactive composition*
- *S shape composition*
- *Diagonal composition*
- ...



A bustling town market,
located in the center of the
screen, ahead of view



A rotating staircase
extending from the bottom
right corner to the top left
corner of the screen



An eagle, low angle



Fantastic Flower Fairy, facing
left, one-third of the screen

2.2.2 Detail settings

➤ Background

The background is a description of the environmental scenery behind the main body in the image, which is an organic component of the subject and can make the image more layered.



A rabbit on a skateboard, the background is a beautiful beach



A rabbit on a skateboard, the background is a medieval town



A rabbit on a skateboard, the background is a dazzling universe

2.2.2 Detail settings

➤ Emotion

The emotions of the main body or/and the image.

A painting &...



happy

exciting

sad

crumbling

A boy sitting in a chair, **, oil painting



Positive mood, low energy

peaceful, calm, serene, soothing, relaxed, placid, comforting, cosy, tranquil, quiet, pastel, delicate, graceful, subtle, balmy, mild, ethereal, elegant, tender, soft, light



Positive mood, high energy

bright, vibrant, dynamic, spirited, vivid, lively, energetic, colorful, joyful, romantic, expressive, bright, rich, saturated, ecstatic, brash, exciting, passionate, hot, flaming



Negative mood, low energy

tired, muted, funereal, somber, dull, melancholic, mournful, gloomy, dismal, sad, pale, washed-out, desaturated, grey, subdued, dreary, depressing, weary



Negative mood, high energy

overwhelmed, dark, ominous, threatening, forbidding, gloomy, terror, stormy, doom, shadowy, unnerving, harrowing, dreadful, frightful, shocking, hideous, ghastly

2.2.3 Modification part

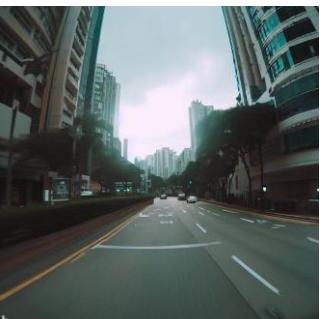
➤ Medium

Based on imagination, directly point out the type of art.

Hong Kong Street View, **



Kodak Gold 200 Black and white Tri-X 400 TX



Polaroid

Double exposure

Dashcam

Type of Painting

- *Chinese painting*
- *Oil painting*
- *Watercolor painting*
- *Gouache painting*
- *Sketch painting*
- *Wash painting*
- *Meticulous painting*
- *Crayon painting*
- *Lacquer painting*
- *Mural painting*
- *Promotional painting*
- *Engraving painting*
- *Teaching Painting*
- ...

Type of Film

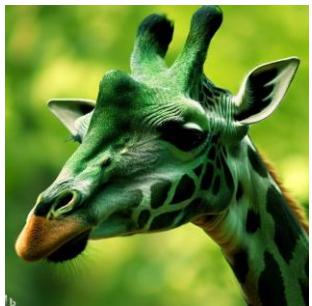
- *Kodak Gold 200*
- *Black and white Tri-X 400 TX*
- *Polaroid*
- *Vintage*
- *Tintype*
- *Autochrome*
- *Lomography*
- *CCTV & Dashcam*
- *Daguerrotype*
- *Double exposure*
- *Cyanotype*
- *Colour splash*
- *Lomo*
- ...

2.2.3 Modification part

➤ Style

Designate a painting or artist style that allows you to directly generate distinctive artworks.

A green giraffe, **



Classicism style

Expressionism style



Fauvism style

Pop style

Dada style

Style of Painting

- *Classicism*
- *Postmodern*
- *Expressionism*
- *Surrealism*
- *Constructivism*
- *Fauvism*
- *Impressionism*
- *Abstractionism*
- *Minimalism*
- *Pop*
- *Dada*
- *Futurism*
- *Cubism*
- ...

Type of Painters

- *Paul Gauguin*
- *Pablo Picasso*
- *Vincent Willem van Gogh*
- *Oscar-Claude Monet*
- *Leonardo da Vinci*
- *Michelangelo Buonarroti*
- *Georges Seurat*
- *Henri Matisse*
- *El Lissitzky*
- *Marcel Duchamp*
- *Hayao Miyazaki*
- *Andy Warhol*
- *Florentijn Hofman*
- ...

2.2.3 Modification part

➤ Time

Infusing your artwork with temporal elements characteristic of a particular era to transport it back in time and evoke a sense of historical context.

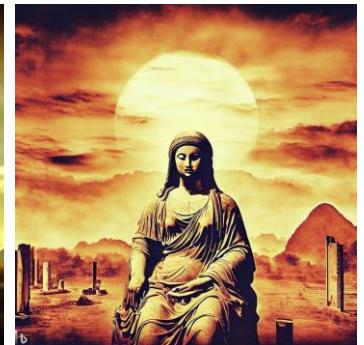
Human Historical Periods

- *Prehistoric era*
- *Ancient era*
- *Medieval era*
- *Early Modern era*
- *Modern era*
- ...

human Society Periods

- *The Stone Age*
- *The Pottery Age*
- *The Iron Age*
- *The Steam Age*
- *The electrical age*
- *The information age*
- ...

Mona Lisa, **



Prehistoric era

Ancient era



Medieval era

Modern era

The information age

2.2.3 Modification part

➤ Light

Utilize the finishing touch of light in creation.

A cute pig in a big blue hat on the streets of Venice's water city, **



with sunset light

with harsh overhead sunlight



with Aurora Polaris



with purple and yellow
lighting



with Silhouette

Outdoor + Light

- *Natural light*
- *Sunlight*
- *Dusk*
- *Sunset*
- *Sunrise*
- *Twilight*
- *Harsh overhead sunlight*
- *Flat lighting*
- *Moonlight*
- *Shadow*
- *Overcast light*
- *Aurora Polaris*
- ...

Indoor + Light

- *Warm lighting*
- *Cold light*
- *'Colourful lighting'*
- *High exposure*
- *Studio lighting*
- *Professional lighting*
- *Silhouette*
- *Chiaroscuro lighting*
- *Cinematic lighting*
- *Backlight*
- *Global illumination*
- *Ray tracing global illumination*
- ...

2.2.3 Modification part

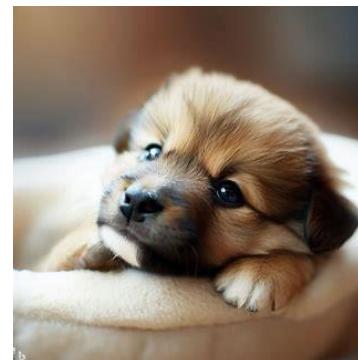
➤ Lens

There are many terms in photography that can also be borrowed and applied.

Camera Parameters

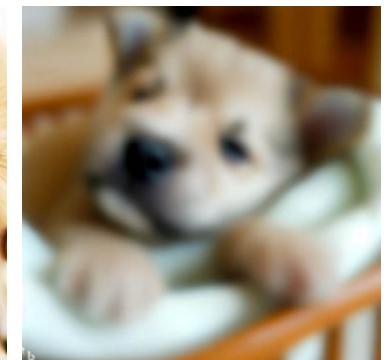
- *Long exposure*
- *Short exposure*
- *Double-exposure*
- *Wide Angle*
- *Extreme close-up angle*
- *Wide shot angle*
- *Medium shot angle*
- *Long shot angle*
- *Over the shoulder point of view*
- *Macro lens sample*
- *Fisheye Lens*
- *Blurry*
- *Tilt Blur*
- *F/22、F/2.8*
- ...

A cute baby dog lying in the cradle, **



Long exposure

Double-exposure



Wide Angle

Extreme close-up angle

Blurry

2.2.3 Modification part

➤ Scene

What scenario should this image be used in?

- *Sports events*
- *Street photography*
- *Advertising*
- *Image photography*
- *Fashion events*
- *Teaching illustrations*



A picture of cheering and shouting, used in sports event promotion



A fashion photography used in the bulletin board of a bus stop



A building structure diagram used as a teaching illustration

➤ Special Efficiency

Creating a large visual experience.

- *Smog*
- *Whole gale*
- *Flame*
- *Halo*
- *Torrential rain*
- *...*



A robot walks on the moon, surrounded by the moon, surrounded by big smoke



A robot walks on the moon, surrounded by big blazing fire



A robot walks on the moon, surrounded by typhoon

2.2.3 Modification part

➤ Dimension

Creating a more impression of space in a photo can enhance its visual impact and engage the viewer on a deeper level, leading to a more impactful and memorable visual experience.

- *3D*
- *Multiverse*
- *Beyond-Dimensional*
- *Extradimensional*

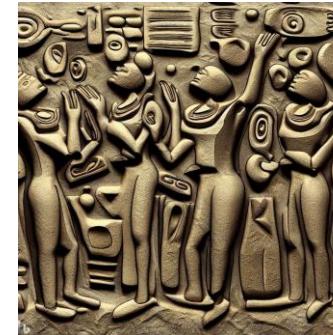
➤ Resolution ratio

Improving photo resolution can enhance image details and effects.

- *HD / Ultra-HD / Full-HD*
- *4k / 8k*
- *Super-Resolution*



A primitive human mural



A primitive human mural,
3D



A primitive human mural,
extradimensional



Maritime Lighthouse



Maritime Lighthouse, 4K



Maritime Lighthouse,
Super-Resolution

2.2.3 Modification part

➤ Detailed & Digital art & Trending on Artstation

Finally, perhaps you would like to add the word ‘detailed’ , ‘digital art’ and ‘trending on artstation’ to comprehensively upgrade your generation.



A amagical town inspired by victorian england and amsterdam, trending on artstation, concept art, matte painting, art by greg rutkowsk, craig mullins, octane render, 8 k, unreal engine, highly detailed, intricate, digital art

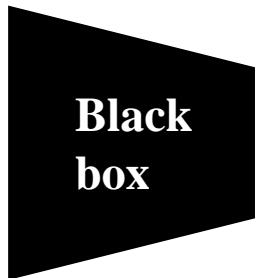


The iron giant standing in the snow looking towards a town, full body image, steam punk, sci-fi, extremely detailed digital painting, in the style of Fenghua Zhong and Ruan Jia and Jermy lipking and peter mohrbacher, mystic colors, deep aesthetic, 8k, highly ornate intricate details, cinematic lighting, rich colors, ray tracing, hyperrealistic, photorealistic, cinematic landscape, highly detailed,digital artwork

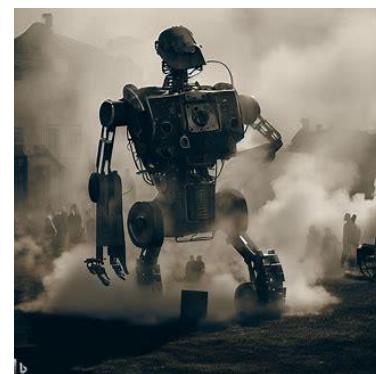


A poster from the animated film Your Name describes a boy and a girl meeting in a dislocation space-time. The boy and girl stand at both ends of the space-time gate, reach out and touch each other's hands. The background is a mountain. Blue and purple lights, Miyazaki style, symmetrical composition, mirror composition, modern, highly gorgeous complex details, cinematic lights, high photo realism, 4k, digital art

2.3 What will AI say



Imaginary



Generated

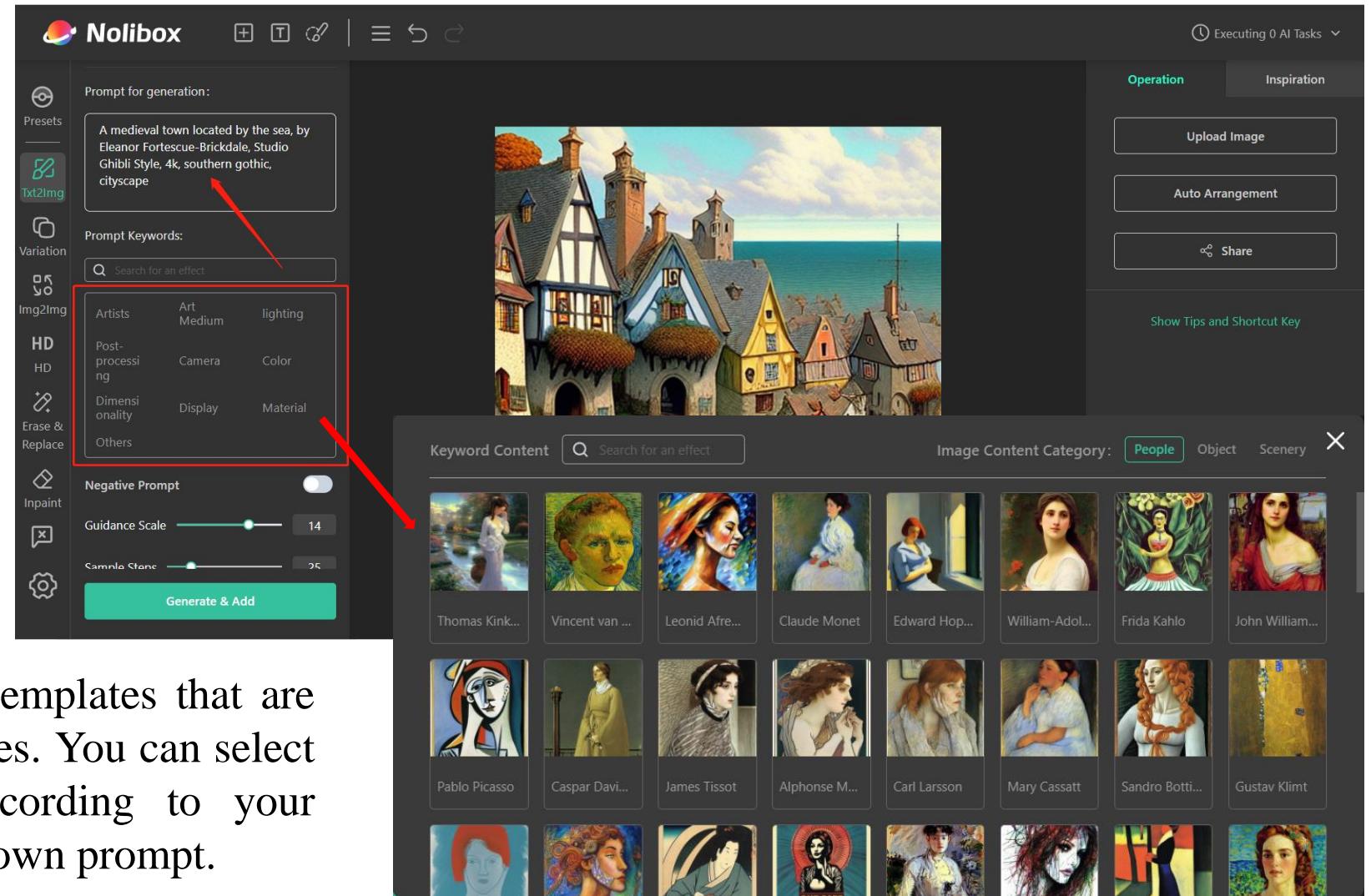
AUXILIARY TOOLS

03

03



3.1 Nolibox



Nolibox has some prompt templates that are divided into different modules. You can select specific prompt words according to your needs and add them to your own prompt.

3.2 Mid journey

Midjourney
@midjourney

Today we're releasing a /describe command that lets you transform images-into-words. Give it a shot! We think this tool will transform your linguistic-visual process both in terms of creative power and discovery.
翻译推文

下午7:50 · 2023年4月3日 · 35.7万 查看

295 转推 114 引用 2,101 喜欢次数 256 书签

Q T L H B U

GooRee @GooRhee · 4月3日

Quick thread on my experiment with /describe using an image I made using Element3D. It's really nice to able to collaborate with Midjourney platform with own work to see what comes out. So many possibilities to create variations.

GooRee @GooRhee · 4月3日

Gonna remix one of my images I created with Element 3D on AE

Using the /describe function to see what it says on #midjourney v5 is really interesting for prompt generation so will now see what they make.

显示这个主题帖

1 forms, rendered in cinema4d, tightly cropped compositions, explosive abstracts, stark simplicity --ar 134:135

2 a black and white origami star shaped with white paper, in the style of rendered in cinema4d, experimental videos, nathan wirth, elongated shapes, hardedge abstraction, blink-and-you-miss-it detail, silver, streamlined design --ar 134:135

3 a black and white abstract shot of paper - looking, in the style of whimsical animation, crystalcore, rendered in cinema4d, the stars art group (xing xing), rusan lomonov, sharp angles, streamlined design, graceful sculptures --ar 134:135

4 a large paper star, which is design

Q T L H B U

19 1.3万

prompt-chat > Talk - Typography Text Words Print

Qoelet 今天08:14 how would you photobash this in order to get an F instead of that random shape?

m3chanical 今天08:49 Dont know

But what is it out of the image you are trying to keep?

The letter "F", a floral letter surrounded by flowers in the background, in the style of realistic and hyper-detailed renderings, baroque-inspired drama, jarring juxtapositions, dark tonality, trompe l'oeil realism, luminous 3d objects, muted tones, surrealism --v 5 -

Gave me

您正在查看以前的消息

Midjourney Bot 机器人 今天14:00

Subscription required

You must have an active subscription to use this feature.

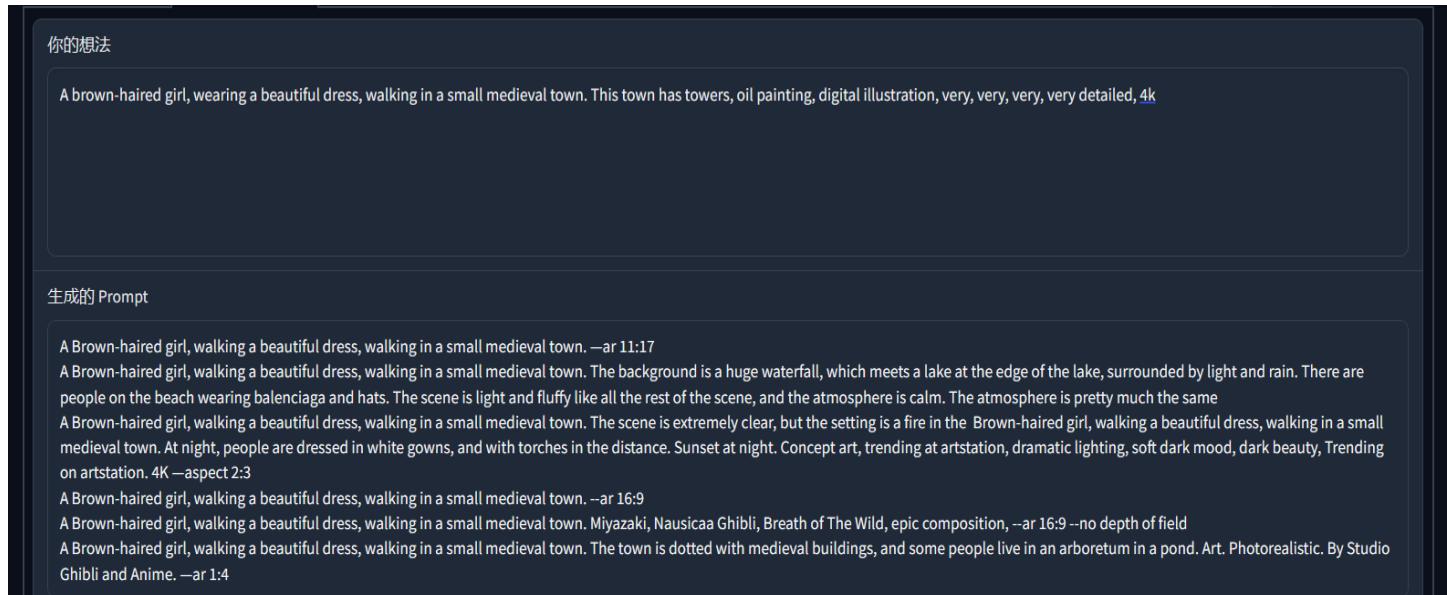
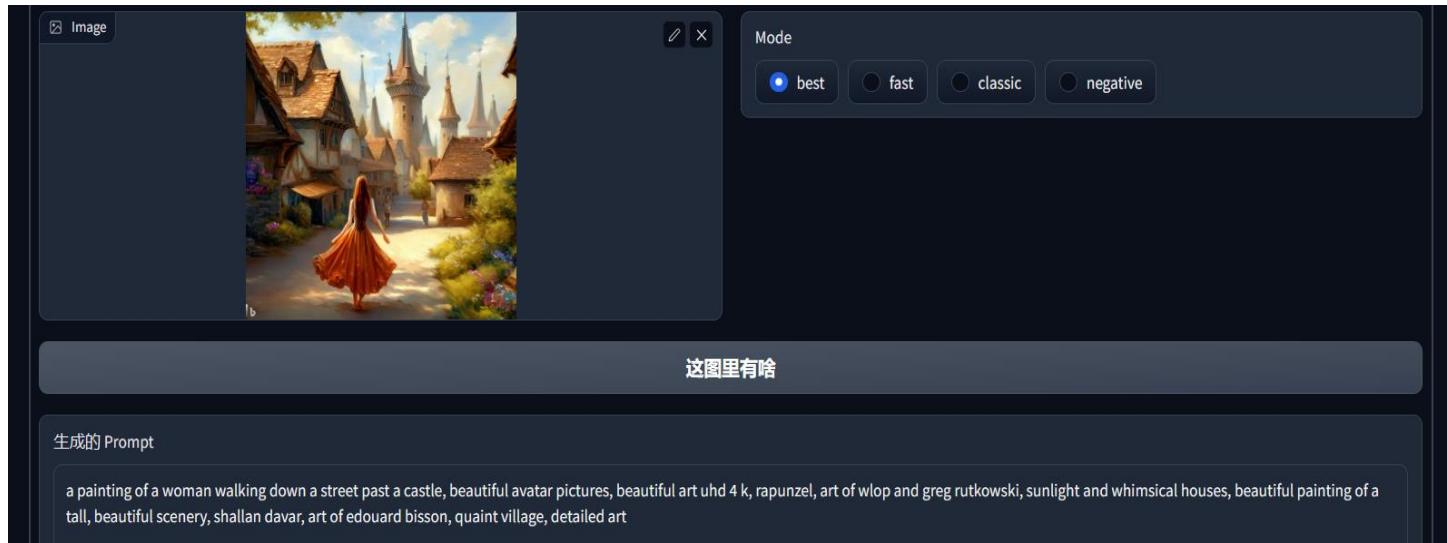
/describe https://cdn.discordapp.com/ephemeral-attachments/1092492867185950852/109578539401175131/sample_3.jpg

Open subscription page

On April 3rd, Mid journey launched the /describe feature, which can help users translate images into text.

3.3 GitHub

It can translate text from images uploaded by users, and can also perform secondary processing on its own text.



3.3 GitHub

A painting of a woman walking down a street past a castle, beautiful avatar pictures, beautiful art uhd 4 k, rapunzel, art of wlop and greg rutkowski, sunlight and whimsical houses, beautiful painting of a tall, beautiful scenery, shallan davar, art of edouard bisson, quaint village, detailed art



A brown-haired girl, wearing a beautiful dress, walking in a small medieval town. This town has towers, oil painting, digital illustration, very, very, very detailed, 4k

A Brown-haired girl, walking a beautiful dress, walking in a small medieval town. The scene is extremely clear, but the setting is a fire in the Brown-haired girl, walking a beautiful dress, walking in a small medieval town. At night, people are dressed in white gowns, and with torches in the distance. Sunset at night. Concept art, trending at artstation, dramatic lighting, soft dark mood, dark beauty, Trending on artstation. 4K



3.4 ChatGPT



I want to generate an image like "A brown-haired girl, wearing a beautiful dress, walking in a small medieval town. This town has towers, oil painting, digital illustration, very, very, very, very detailed, 4k", please help me create a more detailed prompt.



Sure, here's a more detailed prompt for you:

In the image, the brown-haired girl is walking down a narrow, winding street in the heart of a medieval town. The street is lined with tall, narrow buildings made of rough-hewn stone, their roofs covered in moss and lichen. Overhead, a tangle of ivy and other climbing plants winds its way up the walls, creating a verdant canopy that filters the sunlight into dappled patches of green and gold.

The girl herself is wearing a dress made of rich, luxurious fabrics in shades of emerald and deep forest green. The bodice is tightly fitted, with delicate lace trims and a row of tiny buttons down the front. The skirt is full and flowing, with layers of petticoats and billowing silk that rustle softly with each step she takes.

As she walks, the girl passes by a series of intriguing sights and sounds. There are market stalls selling everything from fresh-baked bread to exotic spices, their vendors calling out in thick, guttural accents to potential customers. A group of musicians huddle in a nearby corner, playing lively tunes on wooden flutes and tambourines. And up ahead, looming over the town like a giant sentinel, is a towering castle with high, crenelated walls and turrets that stretch up into the sky.

Regenerate response

A brown-haired girl is walking down a narrow, winding street in the heart of a medieval town. As she walks, the girl passes by a series of intriguing sights and sounds. There are market stalls selling everything from fresh-baked bread to exotic spices, their vendors calling out in thick, guttural accents to potential customers. A group of musicians huddle in a nearby corner, playing lively tunes on wooden flutes and tambourines. And up ahead, looming over the town like a giant sentinel, is a towering castle with high, crenelated walls and turrets that stretch up into the sky.

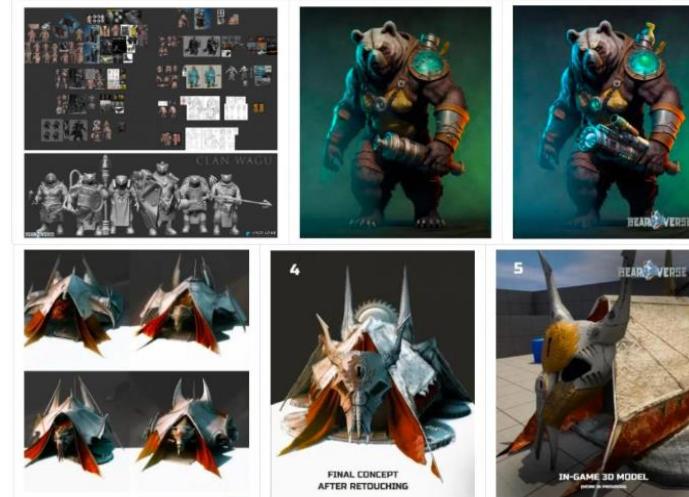


3.4 Challenges from AIGC

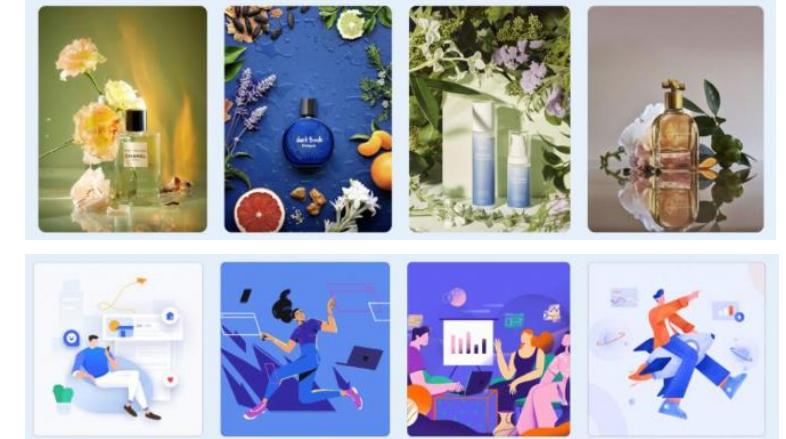
AI Generation



Game Production



Commercial Design



AI technology is constantly evolving and becoming more advanced, it's important to adopt a mindset of lifelong learning and constantly be willing to learn and adapt to new things.



04 04 FUTURE DIRECTIONS

4. Future Directions

- Make-A-Video : <https://makeavideo.studio/>

Meta's artificial intelligence model for generating short videos from given text prompts, launched on September 29, 2022.

- Jasper : <https://www.jasper.ai/>

Jasper AI will assist you with handling emails, blogs, social media posts, and more through AI information and writing assistance.

- Mubert : <https://mubert.com/>

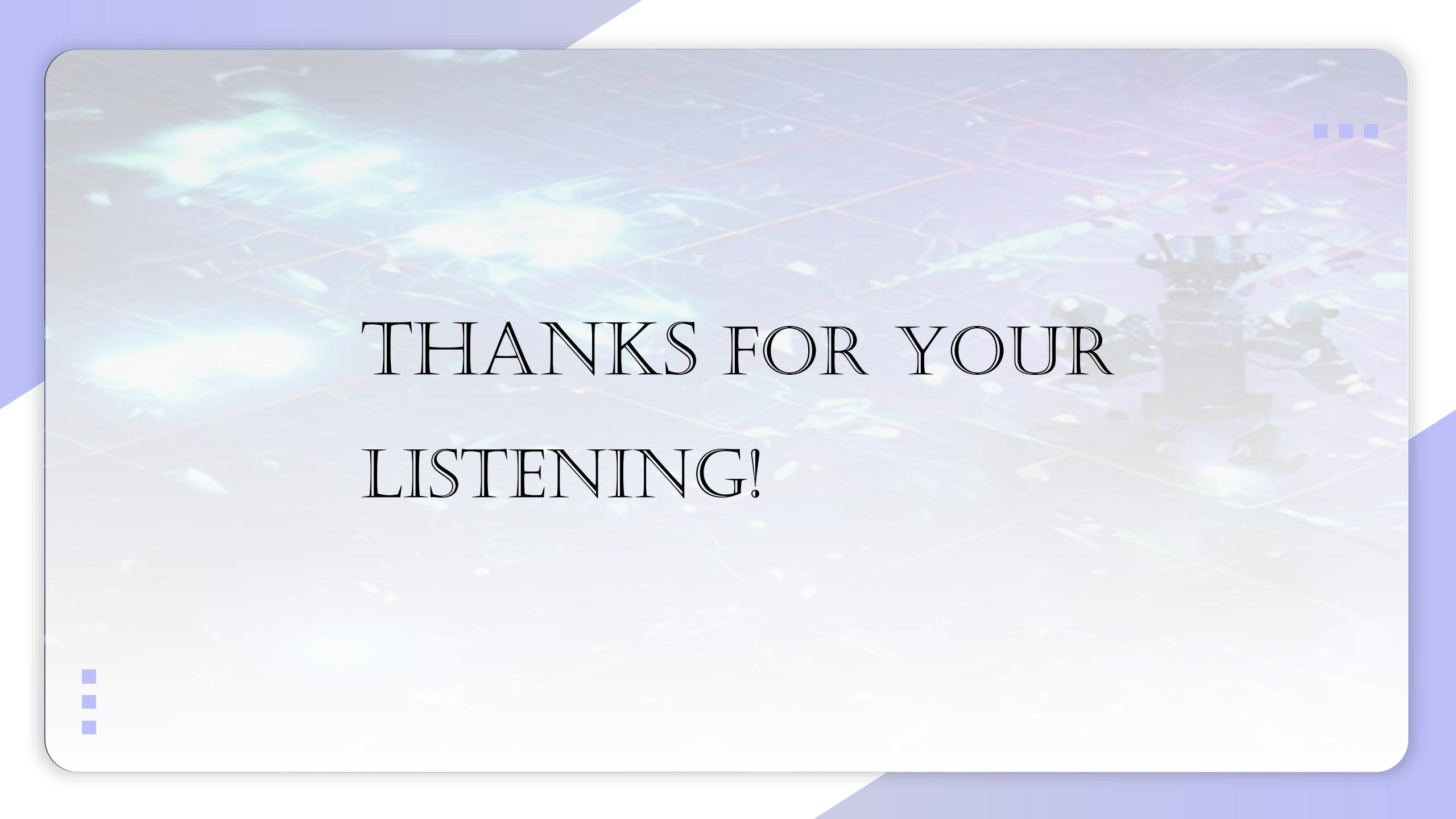
Effortlessly create a soundtrack of the perfect length, mood and style.

05 03 RECOURSSES AND REFERENCES

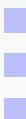


5. Recourses and References

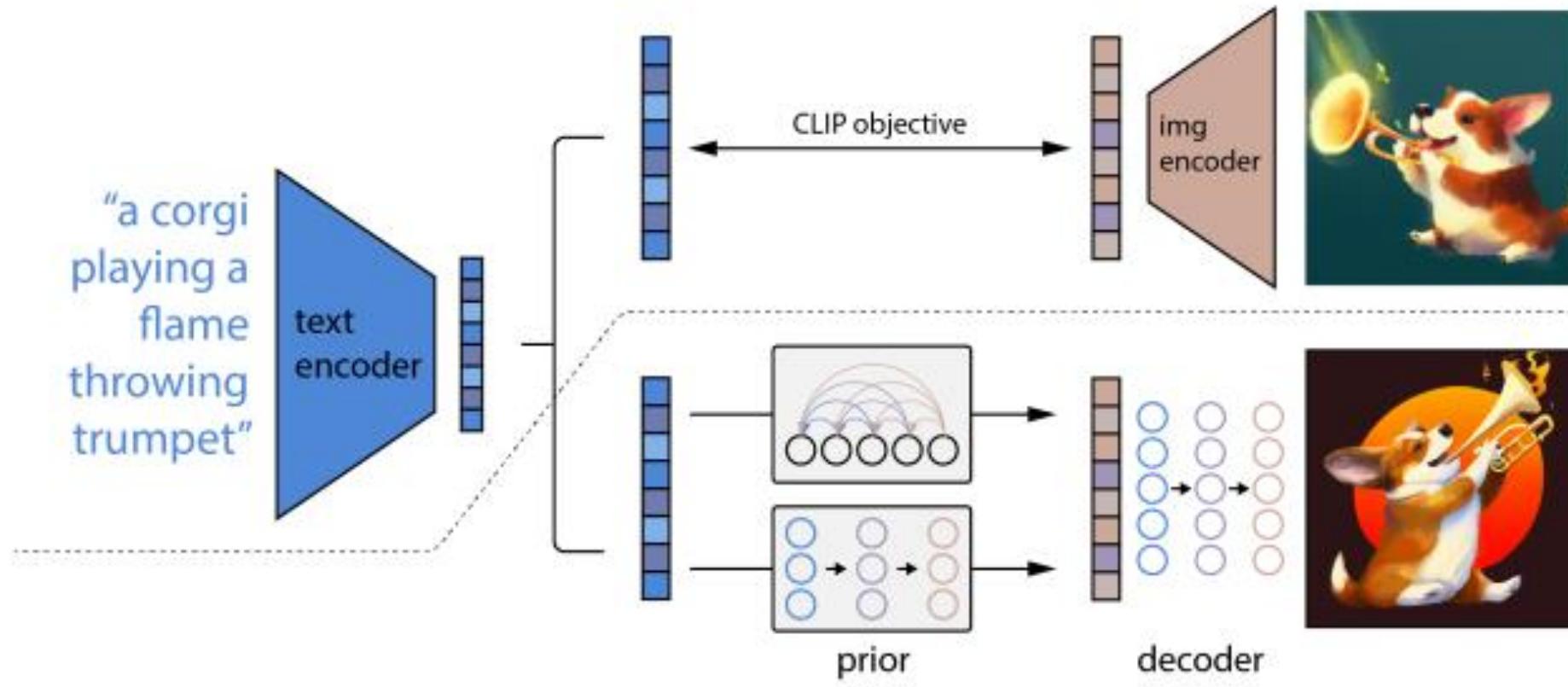
- Open AI official site: <https://platform.openai.com/docs/introduction/key-concepts>
- Open AI research: <https://openai.com/research/dall-e>
<https://openai.com/research/hierarchical-text-conditional-image-generation-with-clip-latents>
- DALL·E2 Principle Explanation: <http://adityaramesh.com/posts/dalle2/dalle2.html>
<https://www.assemblyai.com/blog/how-dall-e-2-actually-works/>
- Docker Prompt Generator: <https://github.com/soulteary/docker-prompt-generator>
- Auxiliary tools: <https://creator.nolibox.com/guest>
<https://discord.com/app/invite-with-guild-onboarding/midjourney>
<https://www.bing.com/images/create?FORM=GENILP>
<https://chat.openai.com/>



THANKS FOR YOUR
LISTENING!



1.5 How DALL·E 2 Actually Works



A birds-eye view of the DALL-E 2 image generation process

1.5 How DALL·E 2 Actually Works

