What's new in AI?

Aicha OUADAH

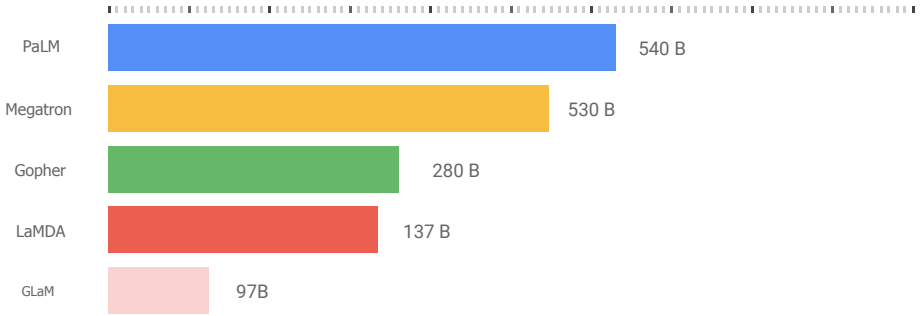
Google I/O Extended

GDG Algiers

- Language models

- Text to image generation

Language Models



Google I/O Extended

GDG Algiers

PaLM , LaMDA

- LaMDA is a model focused on conversation and can power dialogue-based search and voice assistants and other dialogue applications.
- PaLM is a model that follows what Google calls the Pathways AI architecture where a language model is trained to learn how to solve problems.

Google I/O Extended

GDG Algiers

PaLM

Pathways Language Model

new AI architecture, Pathways.

Today's AI models are typically trained to do only one thing. Pathways will enable us to train a single model to do thousands or millions of things.

PaLM

- Social media conversations (multilingual) 50%
- Filtered web pages (multilingual)
- Books (English)
- GitHub (code) 5%
- Wikipedia (multilingual) 4%
- News (English) 1%

PaLM

- It can solve complex math word problems, and answer questions in different languages
- Improvement in Natural Language inference, common-sense reasoning, in context reading comprehension, question answering
- The model can complete code, fix compilation errors, translate code from one language to other

Google I/O Extended

GDG Algiers

LaMDA

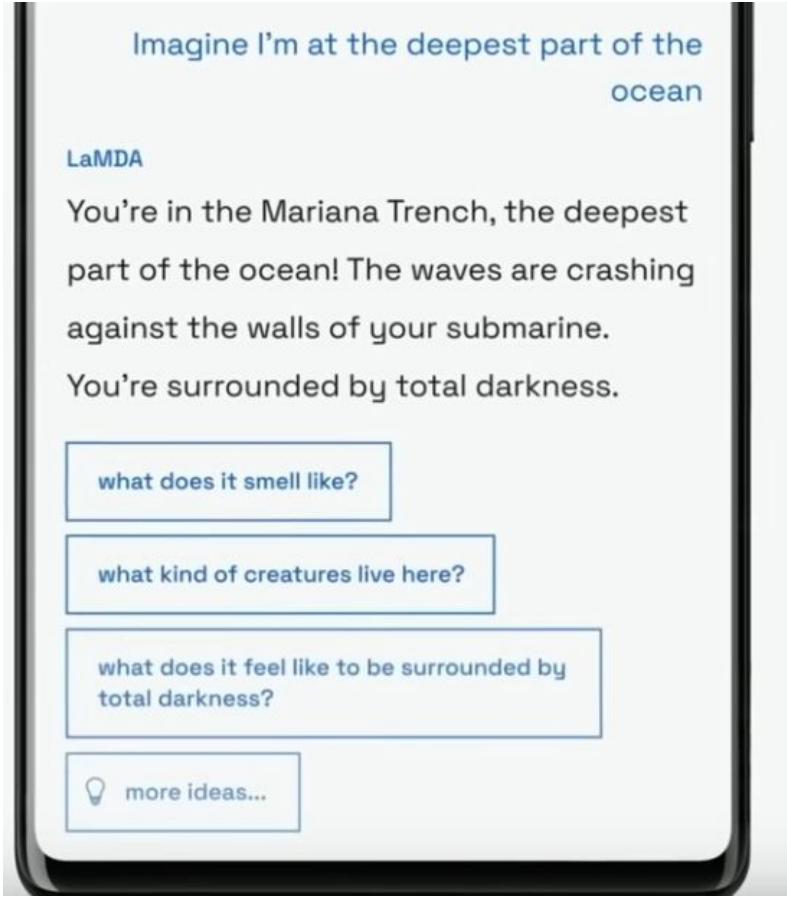
Language Models for Dialog Applications

more natural ways of interacting with technology

LaMDA 2

- Imagen it
- Talk about it
- List it

10:44
List It
May give inaccurate/inappropriate information.
restart
Plant a vegetable garden
• Make a list of what you want to grow
• Research what grows best in your area
• Choose a location in your yard
• Plan your garden layout
• Create a garden plan
• Plant your seeds
How was this list?
Nice
Offensive
Off topic
Not true
get a tip!
regenerate list


Imagine I'm at the deepest part of the ocean
LaMDA
You're in the Mariana Trench, the deepest part of the ocean! The waves are crashing against the walls of your submarine. You're surrounded by total darkness.
what does it smell like?
what kind of creatures live here?
what does it feel like to be surrounded by total darkness?
more ideas...

Text to image generation

DALL·E 2 (openai)
DALL·E Mini (hugging face)
Imagen (google)

DALL·E 2

is a new AI system that can create realistic images and art from a description in natural language.

DALL·E 1

DALL·E 2

"a painting of a fox sitting in a field at sunrise in the style of Claude Monet"

DALL·E 1 vs DALL·E 2

caption matching and photorealism

Google I/O Extended

DALL·E 2 can

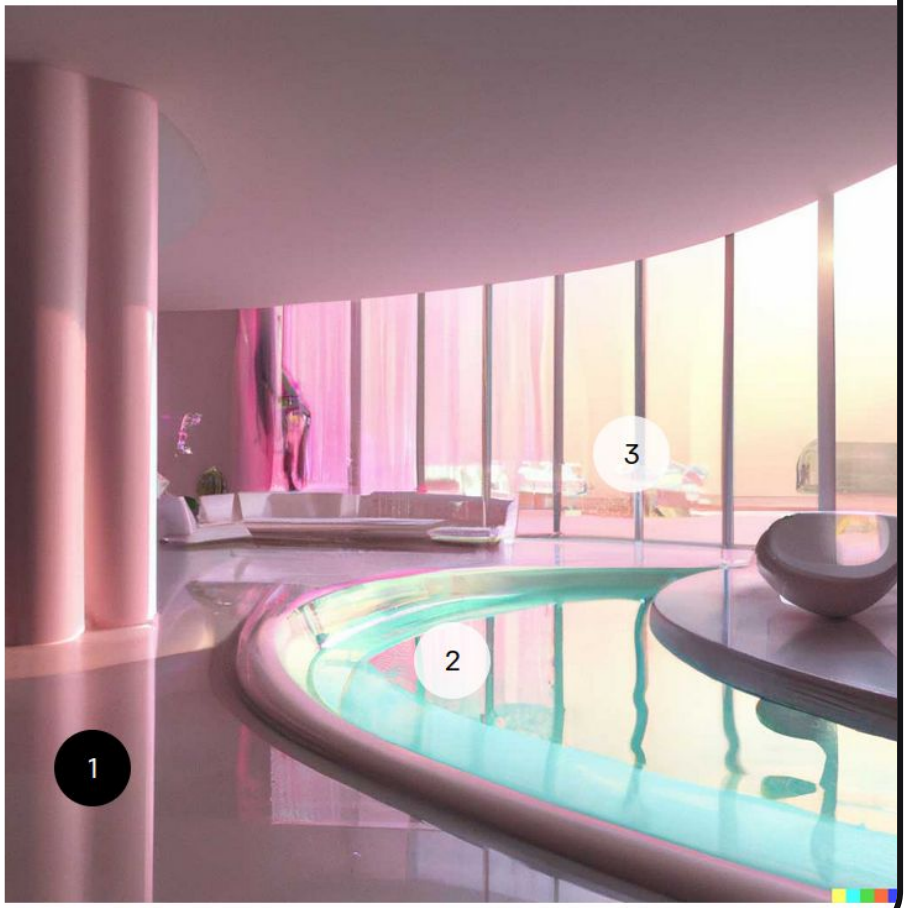
create original, realistic images and art from a text description. It can combine concepts, attributes, and styles.

Input : “A penguin reading a book peacefully while having coffee in the morning”

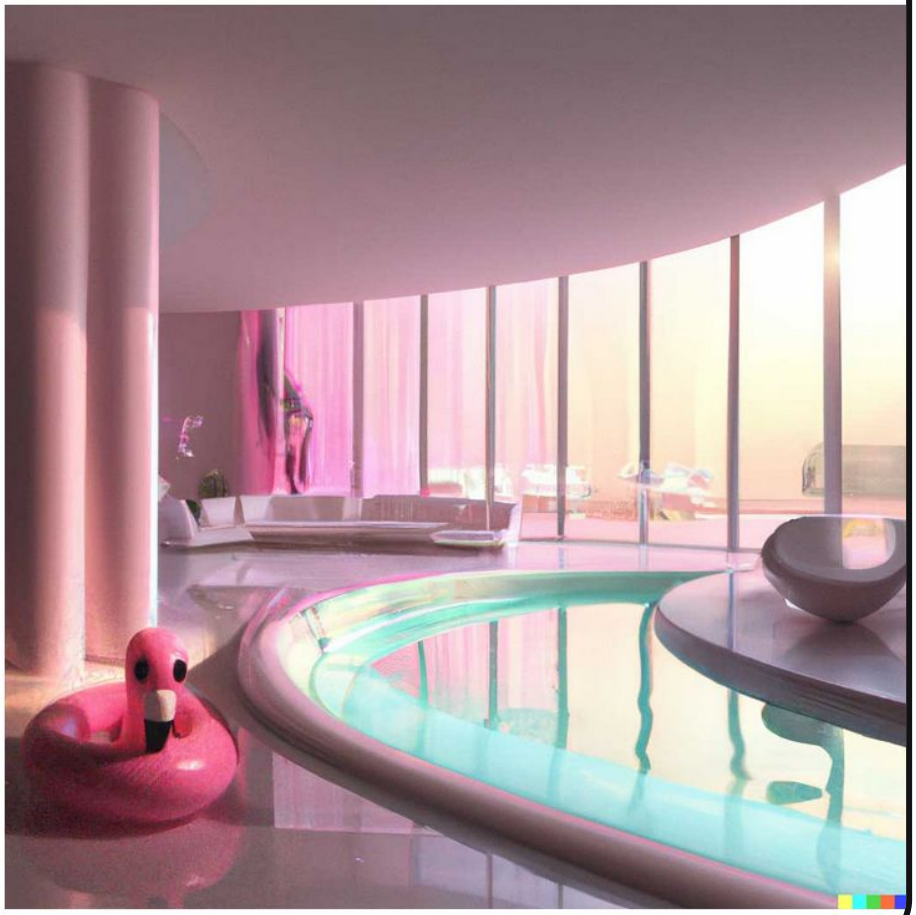
DALL·E 2 can

make realistic edits to existing images from a natural language caption. It can add and remove elements while taking shadows, reflections, and textures into account.

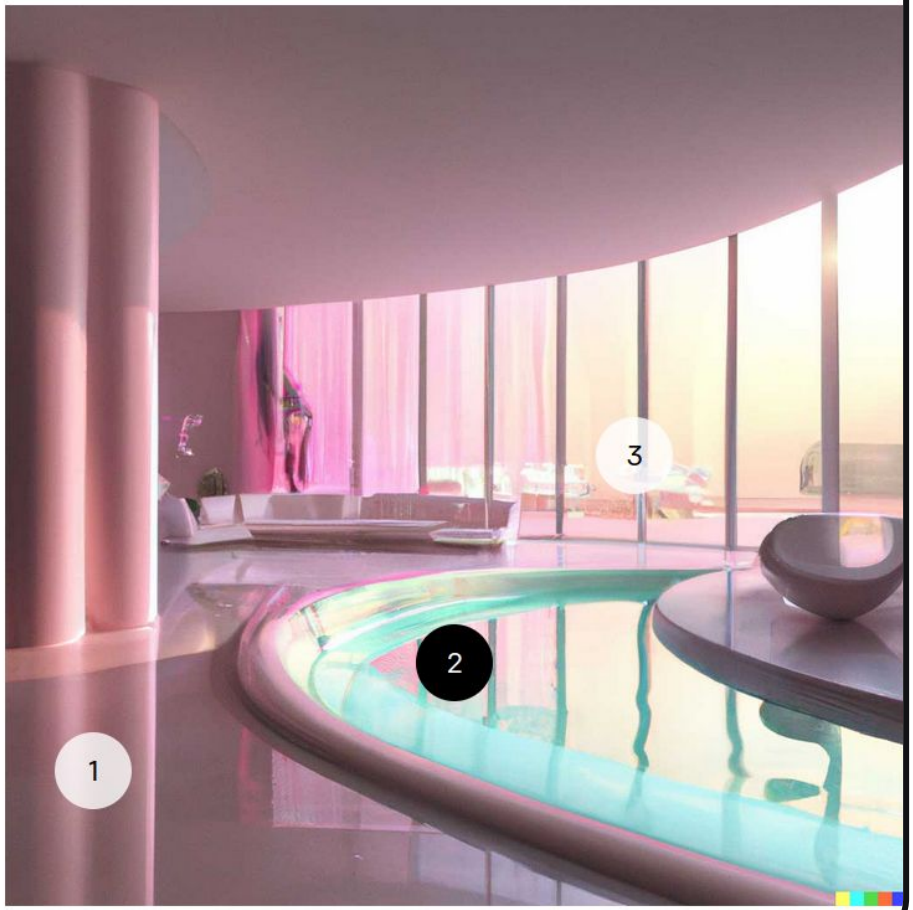
ORIGINAL IMAGE



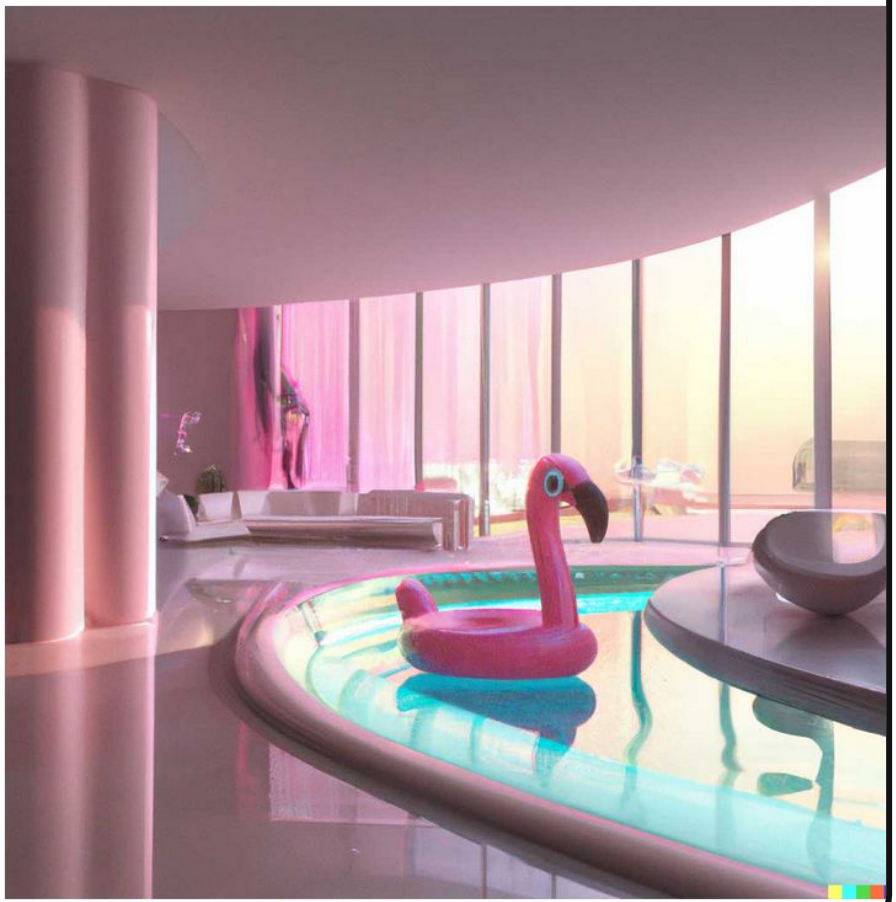
DALL·E 2 EDITS



ORIGINAL IMAGE

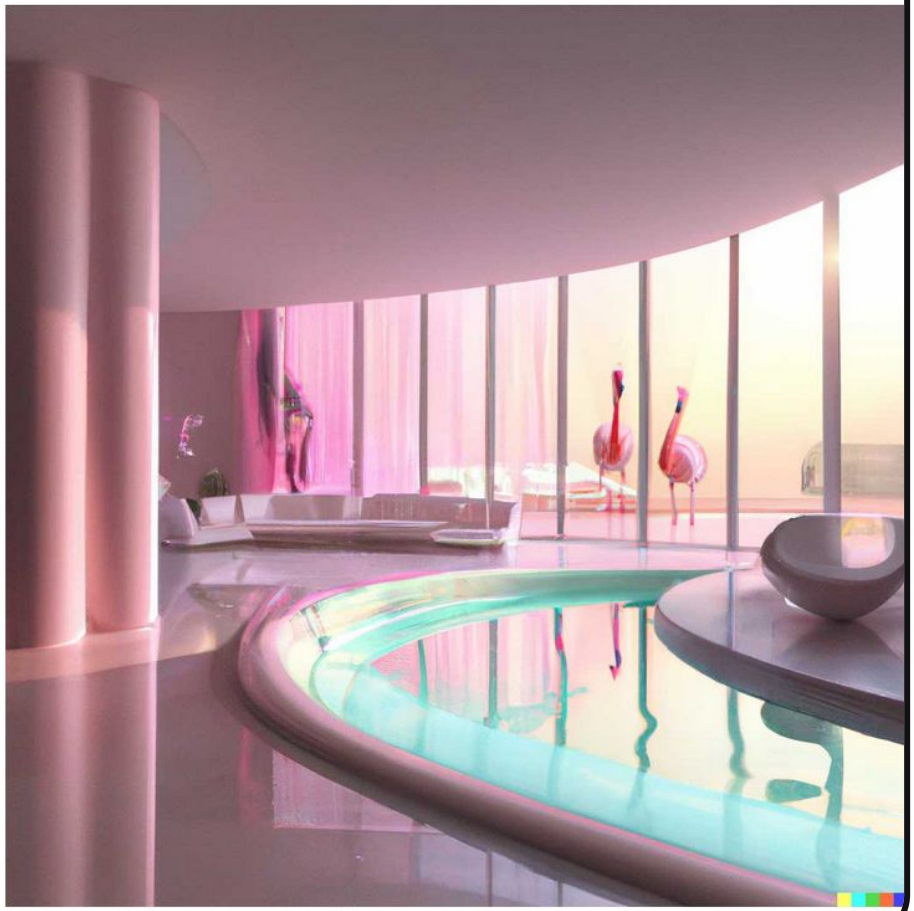


DALL·E 2 EDITS



ORIGINAL IMAGE

DALL·E 2 EDITS



Google I/O Extended

GDG Algiers

DALL·E 2 can

can take an image and create different variations of it inspired by the original.

Google I/O Extended

GDG Algiers

DALL·E 2 VARIATIONS

DALL·E 2 VARIATIONS

DALL·E 2 VARIATIONS

DALL·E 2 :

Diffusion Models are **generative models, meaning that they are used to generate data similar to the data on which they are trained**

Google I/O Extended

GDG Algiers

DALL·E 2
Risks and limitations:

- Adult content, illegal activities, violence, specific identities
- text in images, long text input

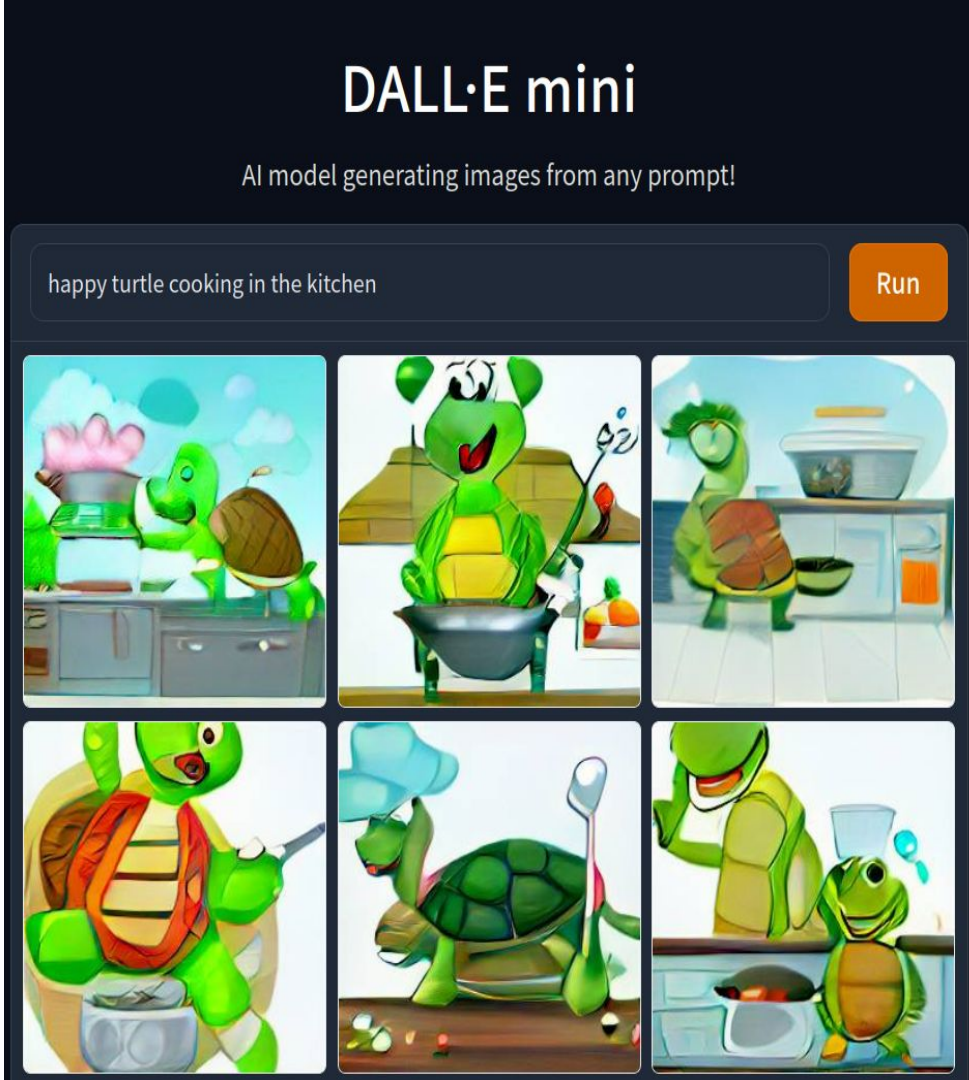
Google I/O Extended

DALL·E Mini

Open source, public use.

Generate images from a text prompt.

Input : “happy turtle cooking in the kitchen”



DALL·E Mini

- is trained with pairs : images and captions
- it has 2 main components, language and image model
- the language model (BART), transforms the text input in natural language to tokens readble by the next model (encoding)
- the image model (VQGAN) is the decoder , it takes the BART output (encoding) and decode it to an image, then its output (generated image) is compared to the image input to adjust it.

Imagen

Is an AI system that creates photorealistic images from input text

Google I/O Extended

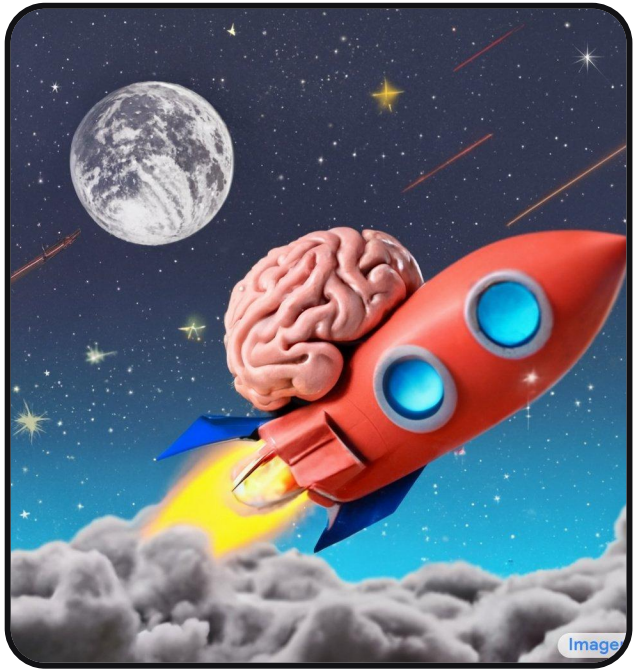
Imagen

Input : “a photo of a fuzzy panda wearing sunglasses and red shirt riding a bike on top of a mountain.”

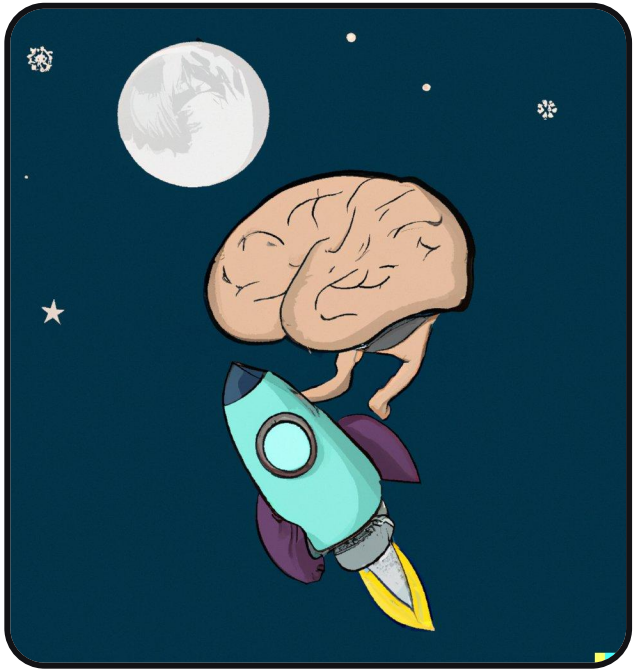
Google I/O Extended

GDG Algiers

A brain riding a rocket ship heading towards the moon.



Imagen



DALL-E 2

source

WHY?

Imagen, DALL·E 1, Glide, DALL·E 2

Google I/O Extended

GDG Algiers

Live Demo (DALL·E Mini)

Volunteer

Google I/O Extended

GDG Algiers

“9ed ma tmed l GDG 9ed ma ymdlek .”

Redha Amirouche
Ex-GDG Algiers Lead

Google I/O Extended

GDG Algiers

“Enjoy your journey with GDG.”

Keep doing cool things that matter

Google I/O Extended

GDG Algiers

Thank you!

Aicha OUADAH

AI Student @USTHB
Intern @LAMIH UPHF

@AichaOuadah

Resources

- https://ai.googleblog.com/
- https://openai.com/
- https://imagen.research.google/