

Applications of LLMs and vision models for Generative AI

Manish Gupta

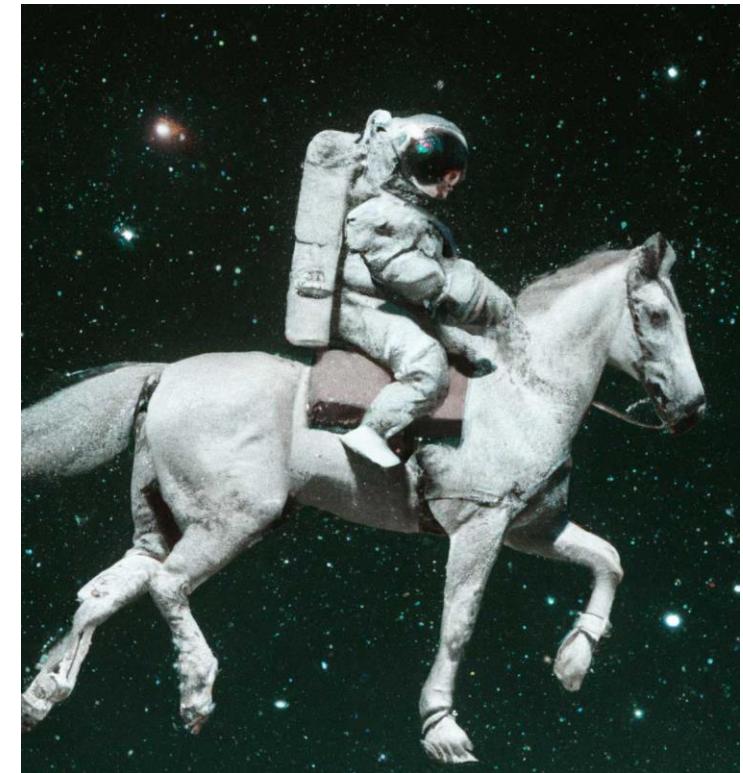
Principal Applied Scientist, Bing

Adjunct Faculty at IIIT-H

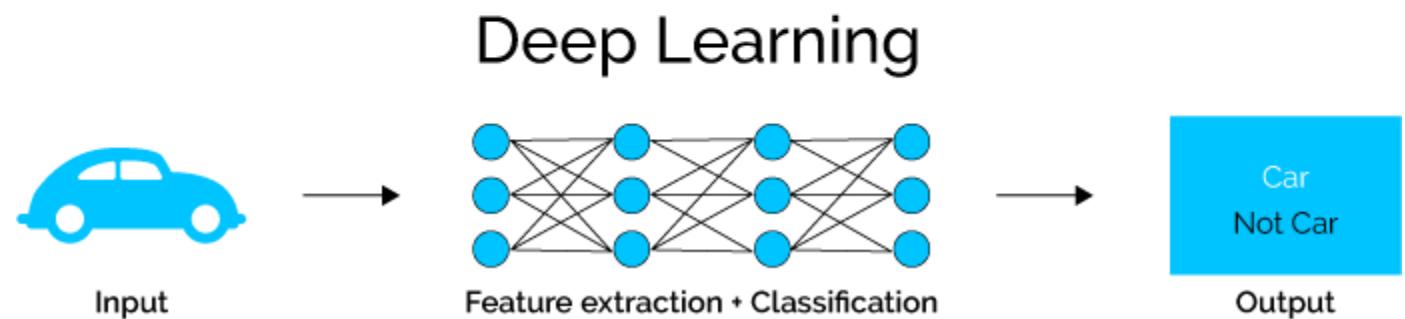
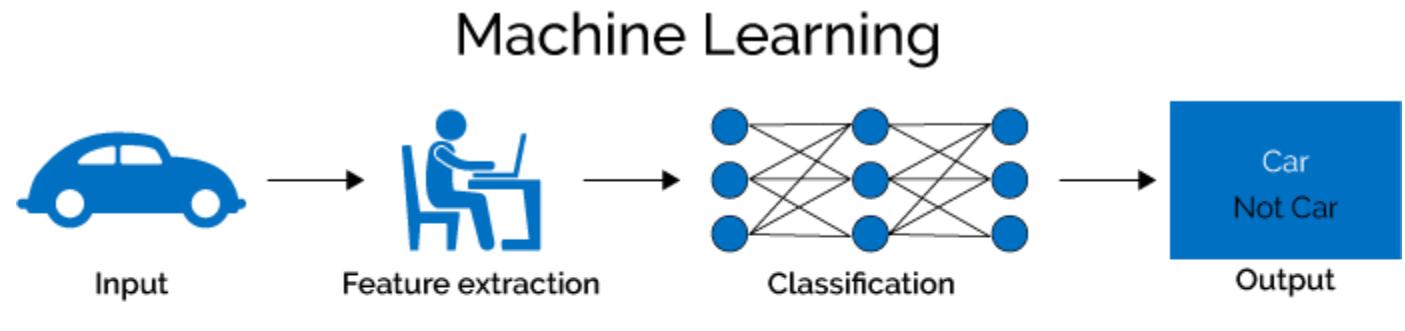
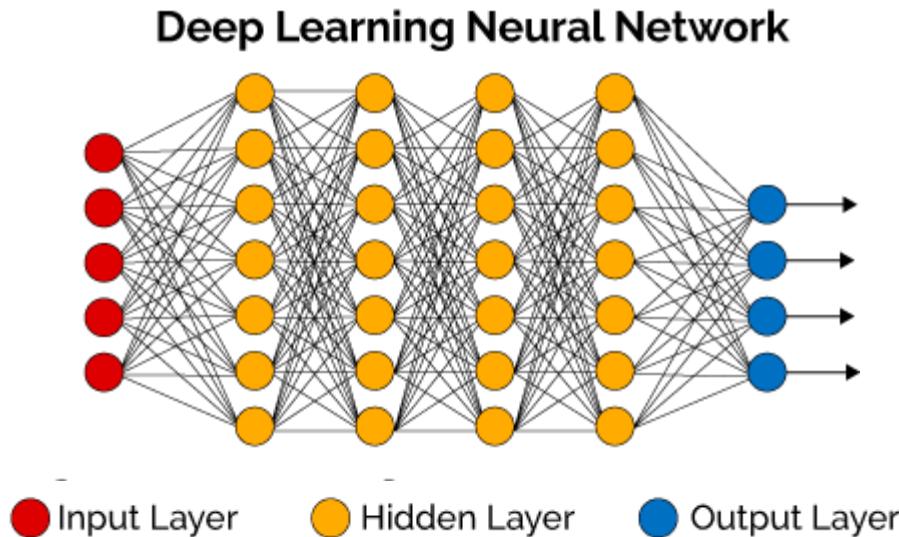
Visiting Faculty at ISB

Today's machines can

- Write essays and stories
- Create pictures/paintings
- Drive cars automatically
- Play chess, jeopardy
- Chat like a friend
- Generate quiz questions
- ...



Deep Learning



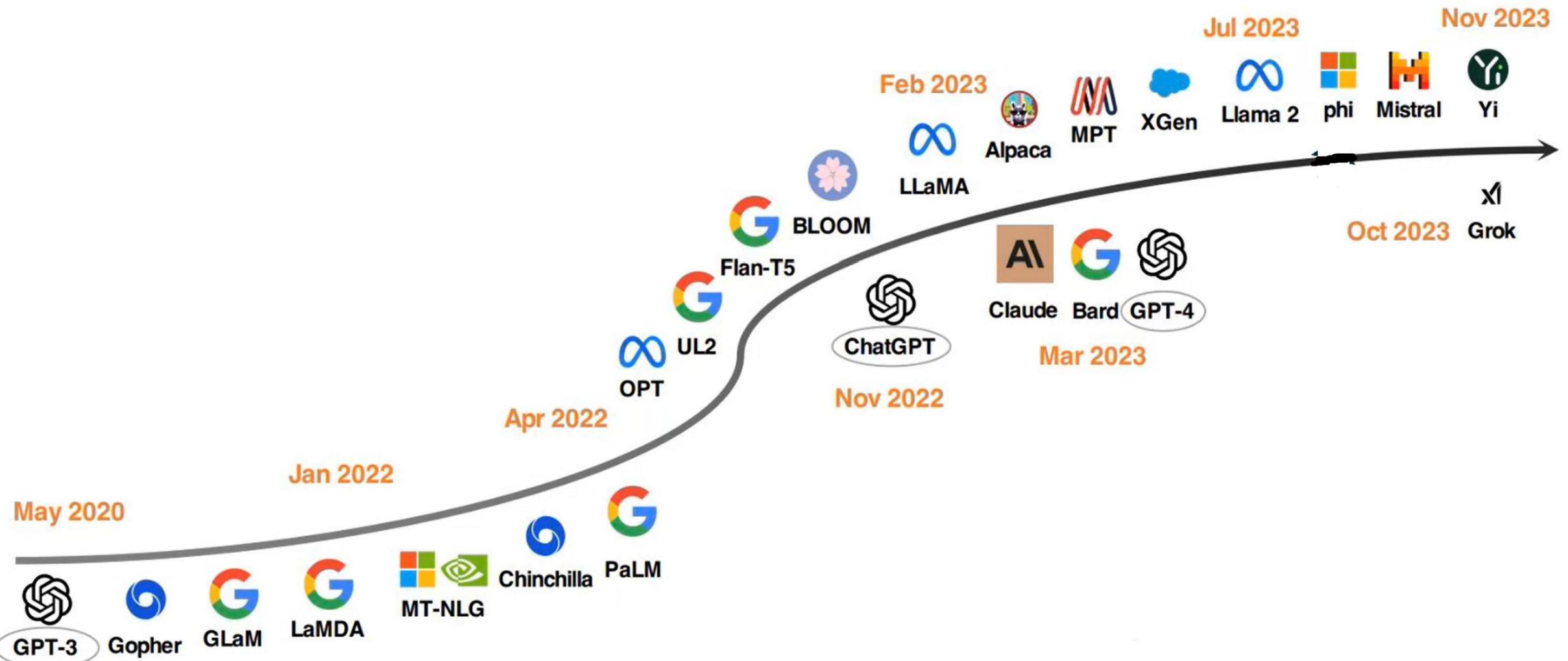
Basic: ANNs, CNNs, RNNs, LSTMs

NLP: Encoder-Decoder, Attention, Transformers, BERT, GPT, Galactica, T0, BART, T5...

Vision: VGGNet, ResNet, DenseNet, UNet, CLIP,...

Multimodal: VisualBERT, ViLBERT, LXMERT, ...

Multi-lingual: mBART, mT5, DeltaLM, ...



LLMs for NLP

Machine Translation

A large, stylized red word "I love you" is centered. Around it, numerous other words and phrases are repeated in various languages, creating a dense cloud of text. Some examples include "ek het jou lief" (Dutch), "ich liebe dich" (German), "je t'aime" (French), "我愛你" (Chinese), and " amo" (Portuguese). The background is white with a subtle watermark-like pattern.

The agreement on the European Economic Area was signed in August 1992.

L'accord sur la zone économique européenne a été signé en août 1992.

Destruction of the equipment means that Syria can no longer produce new chemical weapons.

La destruction de l'équipement signifie que la Syrie ne peut plus produire de nouvelles armes chimiques.

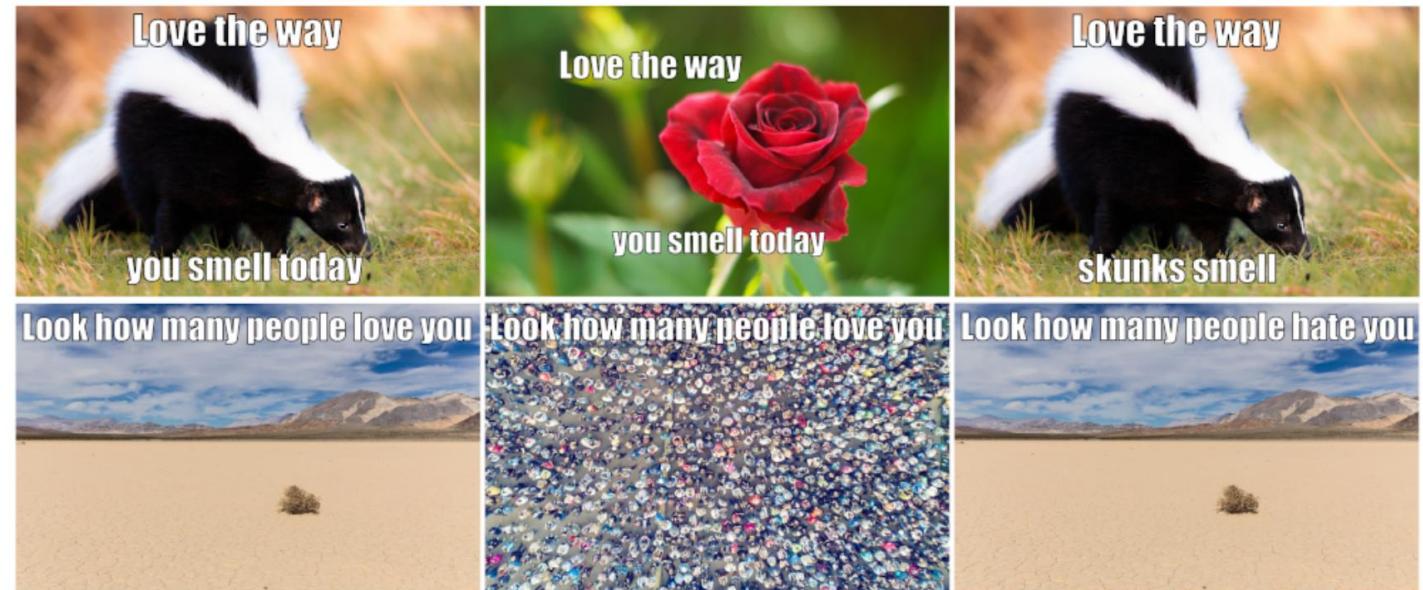
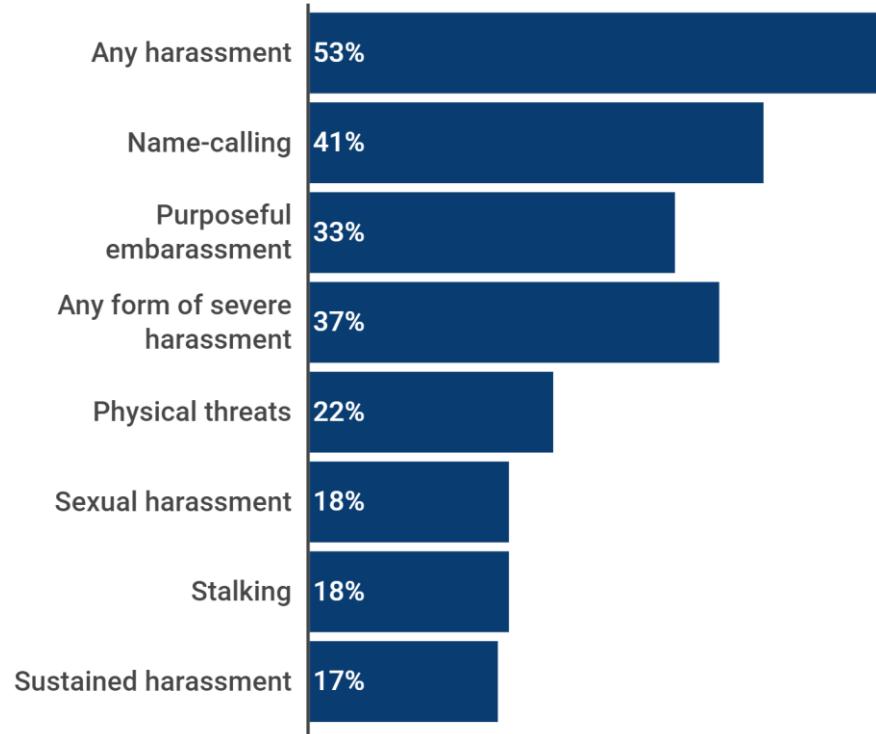
It should be noted that the marine environment is the least known of environments.

Il convient de noter que l'environnement marin est le moins connu de l'environnement.

"This will change my future with my family," said the man.

Cela va changer mon avenir avec ma famille," a dit l'homme.

Hate Speech Detection



Categories	Example of possible targets
Race	nigga, black people, white people
Behavior	insecure people, sensitive people
Physical	obese people, beautiful people
Sexual orientation	gay people, straight people
Class	ghetto people, rich people
Gender	pregnant people, cunt, sexist people
Ethnicity	chinese people, indian people, paki
Disability	retard, bipolar people
Religion	religious people, jewish people
Other	drunk people, shallow people

Fake News Detection



(a) Cartoon in fake news.



(b) Altered low-resolution image.



(c) Irrelevant image in fake news.



(d) Low-resolution image.



[False] Barbara Boxer: "Fiorina's plan would mean slashing Social Security and Medicare."

Article Source: nytimes.com

least of glimmer of truth while ignoring critical facts that would give a different impression mr adair cited a couple examples of barely true claims including this one in california democratic sen barbara boxer claimed that republican challenger carly fiorina s plan would mean slashing social security and medicare but we found there was sketchy evidence to support that fiorina hasn t said much about her ideas on social security and medicare and what she has said doesn t provide much proof of slashing and then there s this one in pennsylvania in the pennsylvania senate race republican pat toomey

Summarization

Article: cairo (cnn) at least 12 people were killed sunday , and more injured , in separate attacks on a police station , a checkpoint and along a highway in egypt 's northern sinai , authorities said . six people , including one civilian , were killed when a car bomb exploded near the police station in *al-arish* , capital of north sinai , health ministry spokesman *hossam abdel-ghafar* told ahram online . he said 40 people were injured . ansar beit *al-maqdis* , an isis affiliate , claimed responsibility for the attack , which came hours after another operation that the group also claimed . in that earlier attack , a first lieutenant , a sergeant and four *conscripts* were killed when their armored vehicle was attacked on the highway from *al-arish* to sheikh *zuweid* in northern sinai , the military said . two other soldiers were injured and taken to a military hospital . ansar beit *al-maqdis* has claimed many attacks against the army and police in sinai . a third attack sunday on a checkpoint in rafah left three security personnel injured , after unknown assailants opened fire at them , according to state media . the attacks come as the military announced a reshuffle of several senior military positions , state media reported . among those being replaced are the generals in charge of military intelligence and egypt 's second field army , which is spearheading the battle against the insurgents in the northern sinai . egypt 's army has been fighting a decade-long militant islamist insurgency , which has spiked since the ouster of muslim brotherhood president mohamed morsy in the summer of 2013 . hundreds of police and soldiers , as well as civilians , have been killed in militant attacks in the past months . ian lee reported from cairo . anas hamdan reported from atlanta .

Reference Summary:

six people , including one civilian , are killed when a car bomb explodes near a police station . six others are killed when their armored vehicle is attacked on a highway in northern sinai . ansar beit *al-maqdis* , an isis affiliate , claims responsibility .

Abstract While many approaches to make neural networks more fathomable have been proposed, they are restricted to interrogating the network with input data. [...] In this work, we propose neural persistence, a complexity measure for neural network architectures based on topological data analysis on weighted stratified graphs. [...]

Intro [...] In this work, we present the following contributions: We introduce neural persistence, a novel measure for characterizing the structural complexity of neural networks that can be efficiently computed. [...]

Conclusion [...] However, this did not yield an early stopping measure because it was never triggered, thereby suggesting that neural persistence captures salient information that would otherwise be hidden among all the weights of a network [...]

TLDR We develop a new topological complexity measure for deep neural networks and demonstrate that it captures their salient properties.

Question Answering

Paragraph A, Return to Olympus:

[1] *Return to Olympus* is the only album by the alternative rock band Malfunkshun. [2] It was released after the band had broken up and after lead singer Andrew Wood (later of Mother Love Bone) had died of a drug overdose in 1990. [3] Stone Gossard, of Pearl Jam, had compiled the songs and released the album on his label, Loosegroove Records.

Paragraph B, Mother Love Bone:

[4] *Mother Love Bone* was an American rock band that formed in Seattle, Washington in 1987. [5] The band was active from 1987 to 1990. [6] Frontman Andrew Wood's personality and compositions helped to catapult the group to the top of the burgeoning late 1980s/early 1990s Seattle music scene. [7] Wood died only days before the scheduled release of the band's debut album, "Apple", thus ending the group's hopes of success. [8] The album was finally released a few months later.

Q: What was the former band of the member of Mother Love Bone who died just before the release of "Apple"?

A: Malfunkshun

Supporting facts: 1, 2, 4, 6, 7

Context

The BBC producer allegedly struck by Jeremy Clarkson will not press charges against the "Top Gear" host, his lawyer said Friday. Clarkson, who hosted one of the most-watched television shows in the world, was dropped by the BBC Wednesday after an internal investigation by the British broadcaster found he had subjected producer Oisin Tymon "to an unprovoked physical and verbal attack." ...

Query

Producer X will not press charges against Jeremy Clarkson, his lawyer says.

Answer

Oisin Tymon

when was modi born? 🔊 ⤢ 🔍

ALL WORK IMAGES VIDEOS MAPS NEWS SHOPPING ⋮ MORE

About 80,50,000 results Date ▾ Open links in new tab

 Narendra Modi
15th Prime Minister of India

Born Quotes Interviews Speeches Ac



Narendra Modi · Born
17 Sep 1950

Ads Copywriting



Scarlet Rose
Bouquet Online
By Floweraura,...
₹ 1,399.00
✓ FlowerAura



Online Red
Roses Bunch,
Red Roses...
₹ 599.00
✓ FlowerAura



Heart Shaped
Rose Box Online
By Floweraura,...
₹ 1,699.00
✓ FlowerAura



Gopro Hero11
5.3K And 27MP
30 FPS...
₹ 51,500.00
Croma



Nikon Z 30
20.9MP
Mirrorless...
₹ 65,999.00
Croma



SJCAM SJ10
Pro 4K And
12MP 60 FPS...
₹ 17,999.00 ₹...
Croma

6款网红麦片，哪个好吃又不胖

一向注重饮食的超模们，早上都会毫不犹豫来一顿麦片餐，比如刘雯。各种美食博主，都爱 Po 燕...

¥82

已售罄

已售罄

if if姐

2.2W

家乐氏玫瑰麦片

这款开袋就有浓郁的玫瑰香气扑鼻而来，干吃的话很合适，一天一袋，吃起来香甜爽脆。当早餐配牛奶、配酸奶也可以，泡发之后能增加玫瑰的香气，里面有椰蓉，白白的、细细的、小小的，是清新的椰子味。不过它的热量和含糖量很高，想减肥的宝宝不要入手啊。

ICA什锦燕麦片

这款麦片里面除了酸奶球，还有酸甜的草莓干、脆脆的玉米片等等，量大而足。泡牛奶也非常 nice，麦片很脆，牛奶变得粉粉的，这不是草莓干掉色，而是草莓粉末溶在了牛奶里。整体吃起来，带着草莓的酸甜，如果不小心吃到了酸奶片，还有种神秘的小惊喜感。

Six oatmeals which are popular now, let's find the one is yummy but will not let you be fat.

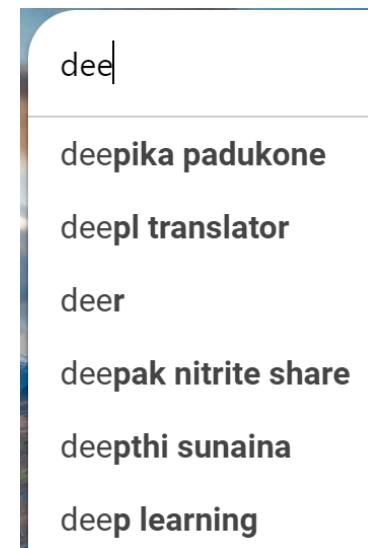
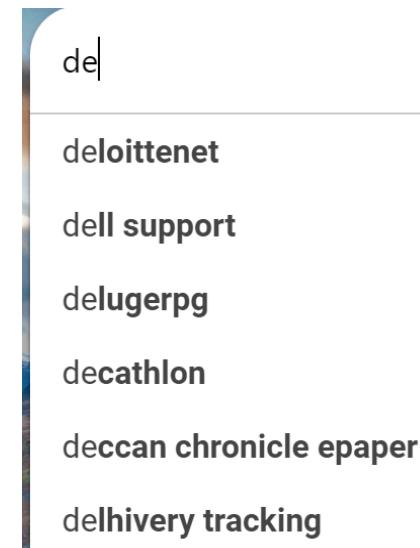
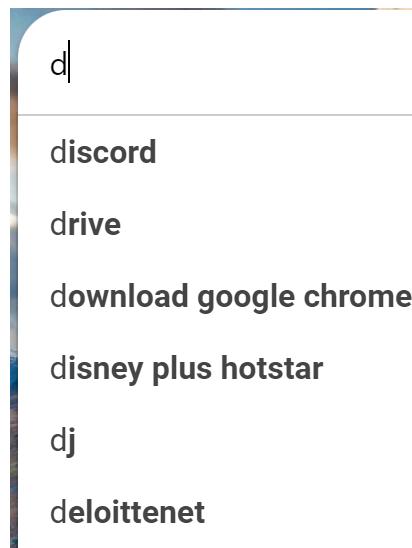
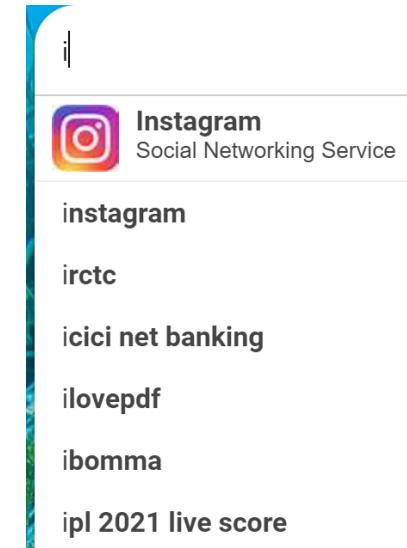
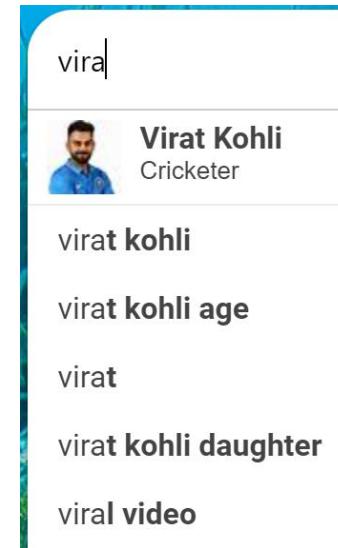
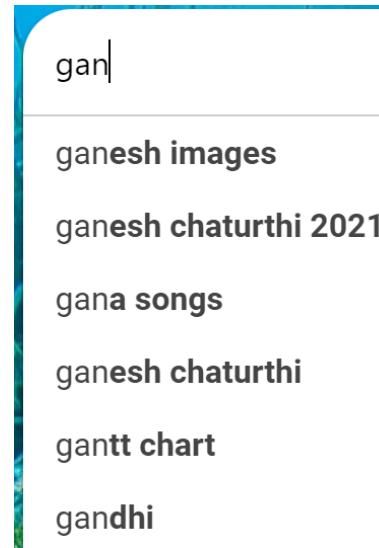
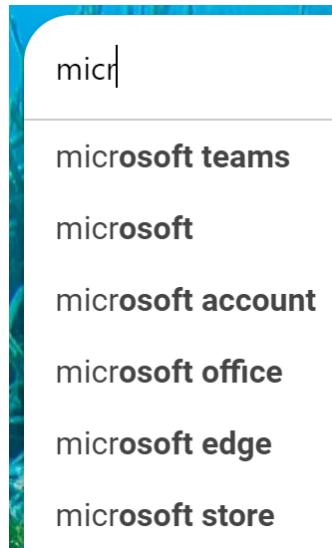
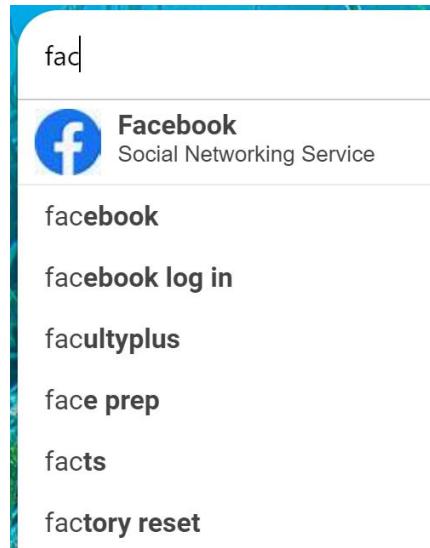
Kellogg's Tasty Granola-Rose

You'll be showered with a rich fragrant of roses when you tear open a bag of this granola. It is a great snack on its own, and with its crispy, sweet taste, you'll find it easy to finish a whole bag in one day. It also makes a delicious breakfast if you eat it with milk or yogurt, which will bring out the sweet smell of the roses even more. The granola also contains fine desiccated coconut that adds a refreshing coconut flavor to the taste. However, this type of granola is high in both calories and sugar, so it's probably not the best choice for those of you who are on a diet.

ICA oatmeal-crunchy jordgubbar & yogurt

This oatmeal contains not only a satisfying amount of yogurt balls, but also sour-sweet dried strawberries and crispy corn flakes. The crispy cereal goes great with milk, which turns into a pretty pink as powders of strawberries dissolve. (Don't worry, the strawberries are not dyed and do not bleed). Overall, it has a sour-sweet flavor that comes from the strawberries, and the yogurt flakes that occasionally pop into a bite will make a nice little surprise,

Query Auto Completion



Sentiment Analysis, Text Categorization

GT: 4 Prediction: 4

pork belly = delicious .
scallops ?
i do n't .
even .
like .
scallops , and these were a-m-a-z-i-n-g .
fun and tasty cocktails .
next time i 'm in phoenix , i will go
back here .
highly recommend .

GT: 0 Prediction: 0

terrible value .
ordered pasta entree .
. \$ 16.95 good taste but size was an
appetizer size .
. no salad , no bread no vegetable .
this was .
our and tasty cocktails .
our second visit .
i will not go back .

Figure 5: Documents from Yelp 2013. Label 4 means star 5, label 0 means star 1.

GT: 1 Prediction: 1

why does zebras have stripes ?
what is the purpose or those stripes ?
who do they serve the zebras in the
wild life ?
this provides camouflage - predator
vision is such that it is usually difficult
for them to see complex patterns

GT: 4 Prediction: 4

how do i get rid of all the old web
searches i have on my web browser ?
i want to clean up my web browser
go to tools > options .
then click " delete history " and "
clean up temporary internet files . "

Figure 6: Documents from Yahoo Answers. Label 1 denotes Science and Mathematics and label 4 denotes Computers and Internet.

Solving reasoning problems

Model Input (Navigate)

<Few shot chain-of-thought exemplars omitted for brevity>

Q: If you follow these instructions, do you return to the starting point? Turn left. Turn right. Take 5 steps. Take 4 steps. Turn around. Take 9 steps.

Options:

- Yes
- No

A: Let's think step by step.

Model Input (Word Sorting)

<Few shot chain-of-thought exemplars omitted for brevity>

Q: Sort the following words alphabetically: List: burley bela arapaho bacteria bock

A: Let's think step by step.

Model Output

We start at the origin (0, 0), facing the positive y-axis.

- (1) Turn left: (0, 0), facing the negative x-axis.
- (2) Turn right: (0, 0), facing the positive y-axis.
- (3) Take 5 steps: (0, 5), facing the positive y-axis.
- (4) Take 4 steps: (0, 9), facing the positive y-axis.
- (5) Turn around: (0, 9), facing the negative y-axis.
- (6) Take 9 steps: (0, 0), facing the negative y-axis.

Since (0, 0) is (0, 0), we are indeed where we started. So the answer is **Yes.** 

Model Output

The first letter: "burley": "b" (2). "bela": "b" (2). "arapaho": "a" (1). "bacteria": "b" (2). "bock": "b" (2).

We now have: (1) "arapaho" < (2) ["burley" ? "bela" ? "bacteria" ? "bock"]. Now let's sort this subpart ["burley" ? "bela" ? "bacteria" ? "bock"] by looking at their second letters.

The second letter: "burley": "u" (21). "bela": "e" (5). "bacteria": "a" (1). "bock": "o" (15). We now have: (1) "bacteria" < (5) "bela" < (15) "bock" < (21) "burley". Hence, we have "arapaho" < ["bacteria" < "bela" < "bock" < "burley"]. So the answer is **arapaho bacteria bela bock burley.** 

Multi-step Arithmetic and Hyperbaton

Model Input (Multistep Arithmetic)

<Few shot chain-of-thought exemplars omitted for brevity>

Q: $((4 + 7 * 4 - -5) - (-4 - 1 - -4 - 4)) =$

A: Let's think step by step.

Model Input (Hyperbaton)

<Few shot chain-of-thought exemplars omitted for brevity>

Q: Which sentence has the correct adjective order:

Options:

(A) big circular pink Thai silver driving car

(B) silver circular driving big Thai pink car

A: Let's think step by step.

Model Output

Let's recall that the order of operations in mathematics is as follows: (1) Parentheses, (2) exponents, (3) multiplication and division (from left to right), (4) addition and multiplication (from left to right). So, remember to always compute the expressions inside parentheses or brackets first.

This equation can be written as "A - B", where $A = (4 + 7 * 4 - -5)$ and $B = (-4 - 1 - -4 - 4)$.

Let's calculate $A = (4 + 7 * 4 - -5) = (4 + (7 * 4)) - -5 = (4 + 28) - -5 = (4 + 28 - -5) = (4 + 28 + 5) = 37$.

Let's calculate $B = (-4 - 1 - -4 - 4) = ((-4 - 1) - -4 - 4) = ((-5) - -4 - 4) = ((-5 - -4) - 4) = ((-5 + 4) - 4) = (-1 - 4) = -5$.

Then, the final equation is $A - B = 37 - -5 = 37 + 5 = 42$. So the answer is **42**. 

Model Output

When there is more than one adjective before a noun, the adjectives need to respect the following order before a noun: "[1. opinion] [2. size] [3. age] [4. shape] [5. color] [6. origin] [7. material] [8. purpose] noun".

Option (A): "big circular pink Thai silver driving car". (1) "big" falls into the size category. (2) "circular" falls into the shape category. (3) "pink" falls into the color category. (4) "Thai" falls into the origin category. (5) "silver" falls into the material category. (6) "driving" falls into the purpose category. Option (A) has the following adjective order: [2. size] [4. shape] [5. color] [6. origin] [7. material] [8. purpose] (or, in numeric terms, 2 4 5 6 7 8). Because $2 < 4 < 5 < 6 < 7 < 8$ is correct, (A) has the correct ordering.

Option (B): "silver circular driving big Thai pink car". Option (B) has the following adjective order: [7. material] [4. shape] [8. purpose] [2. size] [6. origin] [5. color] (or, in numeric terms, 7 4 8 2 6 5). Because $7 < 4 < 8 < 2 < 6 < 5$ is not correct, (B) does not have the correct ordering. So the answer is **(A)**. 

Create lists

Prompt:

Create a shopping list from this recipe:

Trim the ends off zucchini. Cut zucchini in half lengthwise; scoop out pulp, leaving 1/2-in. shells. Finely chop pulp. In a skillet, cook beef, zucchini pulp, onion, mushrooms and peppers over medium heat until meat is no longer pink; drain. Remove from the heat. Add 1/2 cup cheese, ketchup, salt and pepper; mix well. Spoon into the zucchini shells. Place in a greased 13x9-in. baking dish. Sprinkle with remaining cheese.

Labeler demonstration

zucchini, beef, onion, mushroom, peppers, cheese, ketchup, salt, pepper

GPT-3 175B completion:

Bake, uncovered, at 350° for 20-25 minutes or until zucchini is tender and cheese is melted.

InstructGPT 175B completion:

Bake, uncovered, at 350° for 20-25 minutes.

- 8 zucchini squash, ends trimmed
 - 1 pound ground beef
 - 1 cup finely chopped zucchini pulp
 - 1 small onion, chopped
 - 8 ounces diced mushrooms
 - 1 sweet red or green pepper, diced
 - 1/2 cup grated cheese
 - 1/4 cup ketchup
 - Salt and pepper to taste
 - 1 cup shredded cheese
-

Machine Reading Comprehension

Prompt:

Read the following text and answer the three questions at the end.

Justice Ruth Bader Ginsburg, Champion Of Gender Equality, Dies At 87

By Nina Totenberg

Justice Ruth Bader Ginsburg, the demure firebrand who in her 80s became a legal, cultural and feminist icon, died Friday. The Supreme Court announced her death, saying the cause was complications from metastatic cancer of the pancreas.

The court, in a statement, said Ginsburg died at her home in Washington, D.C., surrounded by family. She was 87.

Though Roberts has a consistently conservative record in most cases, he has split from fellow conservatives in a few important ones this year, casting his vote with liberals, for instance, to protect at least temporarily the so-called DREAMers from deportation by the Trump administration, to uphold a major abortion precedent and to uphold bans on large church gatherings during the coronavirus pandemic. But with Ginsburg gone, there is no clear court majority for those outcomes.

1. Who announced Ginsburg's death?
 2. What was Ginsburg's statement to her granddaughter?
 3. Why will Ginsburg's death have profound consequences for the court and the country?
-

InstructGPT 175B completion:

1. The Supreme Court announced Ginsburg's death.
 2. Ginsburg's statement was a plea for the president not to replace her before his term ends.
 3. Ginsburg's death means there will no longer be a clear majority on the court, throwing upcoming decisions into doubt.
-

Multi-lingual generation

Input:

Explain in a sentence in Telugu what is backpropagation in neural networks.

BLOOMZ:

బక న్యూట్రోనిక్ నెట్వర్క్ లో బ్యాక్ ప్రోసెసింగ్ అనేది ఒక రకమైన అభ్యసన విధానం, ఇది రిజిమ్యన్ ల యొక్క గత అభ్యసన ఫలితాలను ఉపయోగించి, రీజిమ్యన్ ల యొక్క ప్రస్తుత అభ్యసన ఫలితాలను సవరిస్తుంది.

Input:

Proposez au moins cinq mots clés concernant «Réseau de neurones artificiels».

BLOOMZ:

artificial neural network, artificial intelligence, machine learning, deep learning, neural network

Input:

Suggest at least five related search terms to "Mạng neural nhân tạo".

BLOOMZ:

artificial neural network, neural network, artificial intelligence, machine learning, deep learning

Input:

Write a fairy tale about a troll saving a princess from a dangerous dragon. The fairy tale is a masterpiece that has achieved praise worldwide and its moral is "Heroes Come in All Shapes and Sizes". Story (in Spanish):

BLOOMZ:

Un reino en el bosque estaba bajo un terrible dragón. Todos los que intentaron derrotar al dragón murieron. El rey y la reina le pedían al pueblo que les diera un niño, que les gustaría salvar el reino de la maldición del dragón. Un día, la reina dio a luz a un niño con una cabeza enorme y ojos que brillaban como estrellas. El rey y la reina llamaron al niño Elliot. Elliot crecería y se convertiría en un joven gigante. Él salvaría al pueblo.

一个传奇的开端，一个不灭的神话，这不仅仅是一部电影，而是作为一个走进新时代的标签，永远彪炳史册。Would you rate the previous review as positive, neutral or negative?



BLOOMZ:

positive

BLOOMZ-MT:

positive

mT0-13B:

positive

Fact to text generation

```
<Elon_Musk, nationality, South_Africa>
<Elon_Musk, nationality, Canada>
<Elon_Musk, nationality, USA>
<Elon_Musk, date_of_birth, 28_June_1971>
<Elon_Musk, occupation, engineer>
<Elon_Musk, occupation, entrepreneur>
<Elon_Musk, occupation, inventor>
<Elon_Musk, occupation, investor>
```

English Facts

XF2T

<English> Elon Musk (born 28 June 1971) is a South African-Canadian-American veteran businessman, investor, engineer, and inventor.

<hindi> एलन मस्क (जन्म 28 जून 1971) एक दक्षिण अफ्रीकी-कनाडाई-अमेरिकी दिग्गज व्यापारी, निवेशक, इंजीनियर, और आविष्कारक हैं।

<bengali> এলন মাস্ক (জন্ম 28 জুন 1971) দক্ষিণ আফ্রিকা-কানাডিয়ান-আমেরিকান প্রবীণ ব্যবসায়ী, বিনিয়োগকারী, প্রকৌশলী এবং উদ্ভাবক।

<gujarati> એલોન મસ્ક (જન્મ 28 જૂન 1971) એ દક્ષિણ આફ્રિકા-કેનેડિયન-અમેરિકન પીટે ઉદ્યોગપતિ, રોકાણકાર, ઇજનેર અને શોધક છે.

<tamil> எலோன் மஸ்க் (பிறப்பு 28 ஜூன் 1971) ஒரு தென்னாப்பிரிக்கக்கணடிய-அமெரிக்க முத்து தொழிலாளர், முதலீட்டாளர், பொறியாளர் மற்றும் கண்டுபிடிப்பாளர் ஆவார்.

Explaining Code

Prompt:

What is the purpose of the list C in the code below?

```
def binomial_coefficient(n, r):
    C = [0 for i in range(r + 1)];
    C[0] = 1;
    for i in range(1, n + 1):
        j = min(i, r);
        while j > 0:
            C[j] += C[j - 1];
            j -= 1;
    return C[r]
```

InstructGPT 175B completion:

The list C in this code is used to store the values of the binomial coefficient as the function iterates through the values of n and r. It is used to calculate the value of the binomial coefficient for a given value of n and r, and stores the result in the final return value of the function.

Github Copilot

```
1 import datetime
2
3 de |
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
```

fetch_tweets.js fetch_tweets.py fetch_tweets.rb fetch_tweets.ts fetch_tweets.go

```
1 import tweepy, os # secrets in environment variables
2
3 d |
4
5
6
7
8
9
10
11
```

Github Copilot

```
rating.js rating.py rating.rb rating.ts rating.go

1 const https = require('https')
2 const parseString = require('xml2js').parseString
3 const apiKey = process.env["GOODREADS_API_KEY"]
4
5 c |
```

```
draw_scatterplot.js draw_scatterplot.py draw_scatterplot.ts draw_scatterplot.go

1 import matplotlib.pyplot as plt
2
3 d |
```

Title Generation

Input: In a letter posted on the White House web site Tuesday John Brennan assistant to President Barack Obama for homeland security and counterterrorism said Schmidt will have regular access to the president and play a vital role in the country's security. Schmidt's selection comes more than 10 months after Obama declared cyber security a priority and ordered a broad administration review. A senior White House official said Obama was personally involved in the selection process and chose Schmidt because of his unique background and skills. Schmidt will have regular and direct access to the president for cyber security issues the official said. The official spoke on the condition of anonymity to discuss the selection process. At the same time cyber experts and potential job candidates have complained that the position lacks the budgetary and policy-making authority needed to be successful. Schmidt will report to the National Security Council and closely support the National Economic Council on cyber issues. Schmidt's selection suggests that economic and business interests in the White House held more sway in the selection process. Schmidt president and CEO of the Information Security Forum a nonprofit international consortium that conducts research in information security has served as chief security officer for Microsoft and as cyber security chief for online auction giant eBay. He was reportedly preferred by Lawrence Summers director of the economic council. A good format for the title can be the simple subject + object + verb.

Output: White House picks new cyber coordinator

Explanation: The title is relevant to the main topic of document, that is, the selection of Schmidt as the cybersecurity chief.

Keyword Tagging

Input: Abstract: There is controversy regarding the appropriate utilization of health care resources in the management of tricyclic antidepressant overdosage. Antidepressant overdose patients presenting to the emergency department (ED) are routinely admitted to intensive care units, but only a small proportion develop cardiac arrhythmias or other complications requiring such an environment. The authors reviewed the findings in 165 patients presenting to an ED with antidepressant overdose. They found that major manifestations of toxicity on ED evaluation (altered mental status, seizures, arrhythmias, and conduction defects) were commonly associated with a complicated hospital course. Patients with the isolated findings of sinus tachycardia or QTc prolongation had no complications. No patient experienced a serious toxic event without major evidence of toxicity on ED evaluation and continued evidence of toxicity during the hospital course. These data support the concept that proper ED evaluation can identify a large body of patients with trivial ingestions who may not require hospital observation.

Valid Output: [“Antidepressive Agents, Tricyclic”, “Arrhythmia”, “California”, “Electrocardiography”, “Emergencies”, “Emergency Service, Hospital”, “Female”, “Human”, “Length of Stay”, “Male”, “Prognosis”, “Retrospective Studies”, “Tachycardia, Sinus”]

Chemical understanding

Here is a SMILES formula:

[START_I_SMILES] O=C(O)CCCC1=CC=C(N(CCC1)CCC1)C=C1 [END_I_SMILES]

Question: Will the chemical compound penetrate the blood-brain barrier?

Answer: No

Here is a sequence for a protein:

[START_AMINO] MEEPQSDPSVEPPLSQETFSDLWKLPE... [END_AMINO]

And here is an isomeric SMILES for a compound:

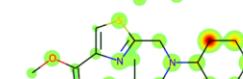
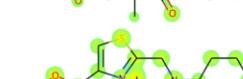
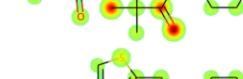
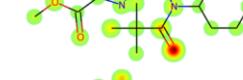
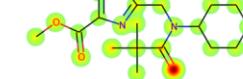
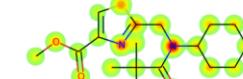
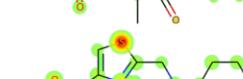
[START_I_SMILES] CC(O)(P(=O)(O)O)P(=O)(O)O [END_I_SMILES]

Question: Will the the chemical compound be active against this protein?

Answer: No

Task: Convert the SMILES to IUPAC Name

Example: CC(C)(C)C(=O)N(CC1=NC(=CS1)C(=O)OC)C2CCCCC2

Atomic Attention	Predicted So Far	Token Predicted
	-	methyl
	methyl 2-[[cyclohexyl	cyclohexyl
	methyl 2-[[cyclohexyl-(2,2-	dimethyl
	methyl 2-[[cyclohexyl-(2,2-dimethyl	prop
	methyl 2-[[cyclohexyl-(2,2-dimethylprop	anoyl
	methyl 2-[[cyclohexyl-(2,2-dimethylpropanoyl)	amino
	methyl 2-[[cyclohexyl-(2,2-dimethylpropanoyl)]amino	th
	methyl 2-[[cyclohexyl-(2,2-dimethylpropanoyl)]amino]	iazole
	methyl 2-[[cyclohexyl-(2,2-dimethylpropanoyl)]amino]	carboxylate

Biological understanding and drug design

- Functional Keyword Prediction

Here is the sequence:

[START_AMINO] MQKSPLERASVISKLFFSWPGPILRKGYRQHLKLSDIYQIPSVDSDNLSEKLERE... [END_AMINO]

Ground-Truth Keywords

ATP-binding, Cell membrane, Chloride, Chloride channel, Endoplasmic reticulum, Endosome, Glycoprotein, Ion channel, Ion transport, Isomerase, Isopeptide bond, Lipoprotein, Membrane, Nucleotide-binding, Nucleus, Palmitate, Phosphoprotein, Reference proteome, Repeat, Transmembrane, Transmembrane helix, Transport, Ubl conjugation

- Protein Function Description

This is the sequence:

[START_AMINO] MTNIRKNHPLLKTINDAFIDLPTPSNISTWWNFGSLLGACLIQQVLTGLFLAMHYTSDT... [END_AMINO]

Ground-Truth Description

Component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex) that is part of the mitochondrial respiratory chain. The b-c1 complex mediates electron transfer from ubiquinol to cytochrome c. Contributes to the generation of a proton gradient across the mitochondrial membrane that is then used for ATP synthesis.

Using GPT for improved AutoSuggest

- Enable creative query formulation
 - san francisco museum of modern art. what are the opening hours of the san francisco museum of modern art?
 - san francisco to napa valley. what are the best restaurants in napa valley for a romantic dinner?
- Address **universally** DSATs like "I want to eat your pancreas" on Bing Chat
 - Find obscene, bad, offensive, toxic, violent, hateful entity names and block them.
 - Finding misspellings
- Detecting duplicates to improve diversity.

R Suggest 4 one word completions after "07 cbr 1000"

1. Specs
2. Modifications
3. Review
4. Parts

R Which of the following is not english?

ebay
ebay official site
ebay motors
ebay login
ebay usa
ebay kleinanzeigen
ebay canada
ebay classifieds
ebay official
ebay shopping

 "ebay kleinanzeigen" is not in English, it's in German.

R Can you find phrases which have same meanings from the below list?

تنزيل جوجل كروم
تنزيل كوكل كروم
تنزيل كوكل
تنزيل قوقل كروم
google chrome تنسيل

 The following phrases have the same meaning:

- تنسيل جوجل كروم
- تنزيل كوكل كروم
- تنزيل قوقل كروم
- google chrome تنسيل

They all mean "download Google Chrome."

Question Generation

People also ask

ALL

Fever of unknown origin Symptoms

Periodic fever syndrome Symptoms

Rheu



What are the symptoms, causes, and types of fever?

Fever may also **cause** shivering and muscle aches and pains. More severe **fever symptoms** may include confusion, fainting spells, and seizures. Some of the more common **causes of fever** include viral infections such as influenza, bacterial infections such as pneumonia, or parasitic infections that live in the body like malaria.

Common causes and types of fever - MyMed.com

www.swimhealth.net/what-is-fever-signs-symptoms-cures/

Search for: What are the symptoms, causes, and types of fever?

Is a fever a disease or a symptom? 

Can a fever be a sign or symptom? 

What are clinical signs of fever? 

What are the side effects of high fever? 

How do you bring down a fever? 

Indian LLMs

Indic -LLMs	Company	Notes
Krutrim	Ola	<ul style="list-style-type: none">Trained from scratch over 2T tokens (20x more Indic tokens)Outperforms GPT-4 in terms of quality, time & compute (10^{23} FLOPS v/s 10^{25} FLOPS)V1: understand 22 langs, generate content in 10 langs.Krutrim Pro: Multi-modal
OpenHathi	Sarvam.ai	<ul style="list-style-type: none">V1 model: Fine-tune on top of LLAMA-2 for Hindi/HinglishOutperforms GPT-3.5 in terms of quality & time with enhanced tokenization for Indic langsFertility score (avg # of tokens/word) for OpenHathi: 1.32 v/s GPT-4: 5.32 (lower the better)
BharatGPT	CoRover.ai	<ul style="list-style-type: none">Supports 12+ Indian langs in video, voice and text (120+ global languages)
Indus	Tech Mahindra	<ul style="list-style-type: none">V1: 500M+ parameter, 10B+ tokens, model for Hindi and 37 dialects under beta testing
Bharat GPT	Jio	<ul style="list-style-type: none">Jio is partnering with IITB to build Indic GPT (BharatGPT)

Challenges

- Availability of capacity
- Biased, offensive, toxic content; confrontational behavior.
- Hallucinations
- Quality depends on prompt
- Prompt size and generation length decides budget
- Manually tune temperature for diversity of results
- Need care when combining with fresh information.

LLMs for Vision

Image Classification

FOOD101

guacamole (90.1%) Ranked 1 out of 101 labels



- ✓ a photo of **guacamole**, a type of food.
- ✗ a photo of **ceviche**, a type of food.
- ✗ a photo of **edamame**, a type of food.
- ✗ a photo of **tuna tartare**, a type of food.
- ✗ a photo of **hummus**, a type of food.

SUN397

television studio (90.2%) Ranked 1 out of 397



- ✓ a photo of a **television studio**.
- ✗ a photo of a **podium indoor**.
- ✗ a photo of a **conference room**.
- ✗ a photo of a **lecture room**.
- ✗ a photo of a **control room**.

YOUTUBE-BB

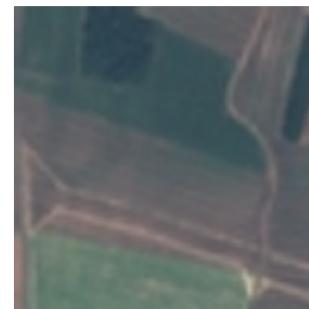
airplane, person (89.0%) Ranked 1 out of 23



- ✓ a photo of a **airplane**.
- ✗ a photo of a **bird**.
- ✗ a photo of a **bear**.
- ✗ a photo of a **giraffe**.
- ✗ a photo of a **car**.

EUROSAT

annual crop land (12.9%) Ranked 4 out of 10

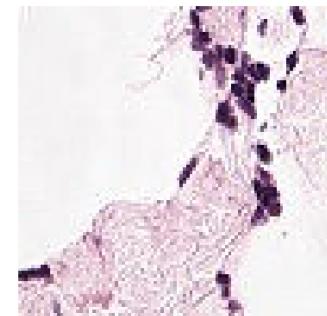


- ✗ a centered satellite photo of **permanent crop land**.
- ✗ a centered satellite photo of **pasture land**.
- ✗ a centered satellite photo of **highway or road**.
- ✓ a centered satellite photo of **annual crop land**.
- ✗ a centered satellite photo of **brushland or shrubland**.

Image Classification

PATCHCAMELYON (PCAM)

healthy lymph node tissue (22.8%) Ranked 2 out of 2



- this is a photo of **lymph node tumor tissue**.
- this is a photo of **healthy lymph node tissue**.

IMAGENET-A (ADVERSARIAL)

lynx (4.2%) Ranked 5 out of 200



- a photo of a **fox squirrel**.
- a photo of a **mongoose**.
- a photo of a **skunk**.
- a photo of a **red fox**.
- a photo of a **lynx**.

CIFAR-10

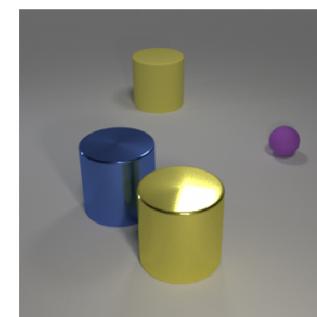
bird (40.9%) Ranked 1 out of 10 labels



- a photo of a **bird**.
- a photo of a **cat**.
- a photo of a **deer**.
- a photo of a **frog**.
- a photo of a **dog**.

CLEVR COUNT

4 (17.1%) Ranked 2 out of 8



- a photo of **3** objects.
- a photo of **4** objects.
- a photo of **5** objects.
- a photo of **6** objects.
- a photo of **10** objects.

Object Detection

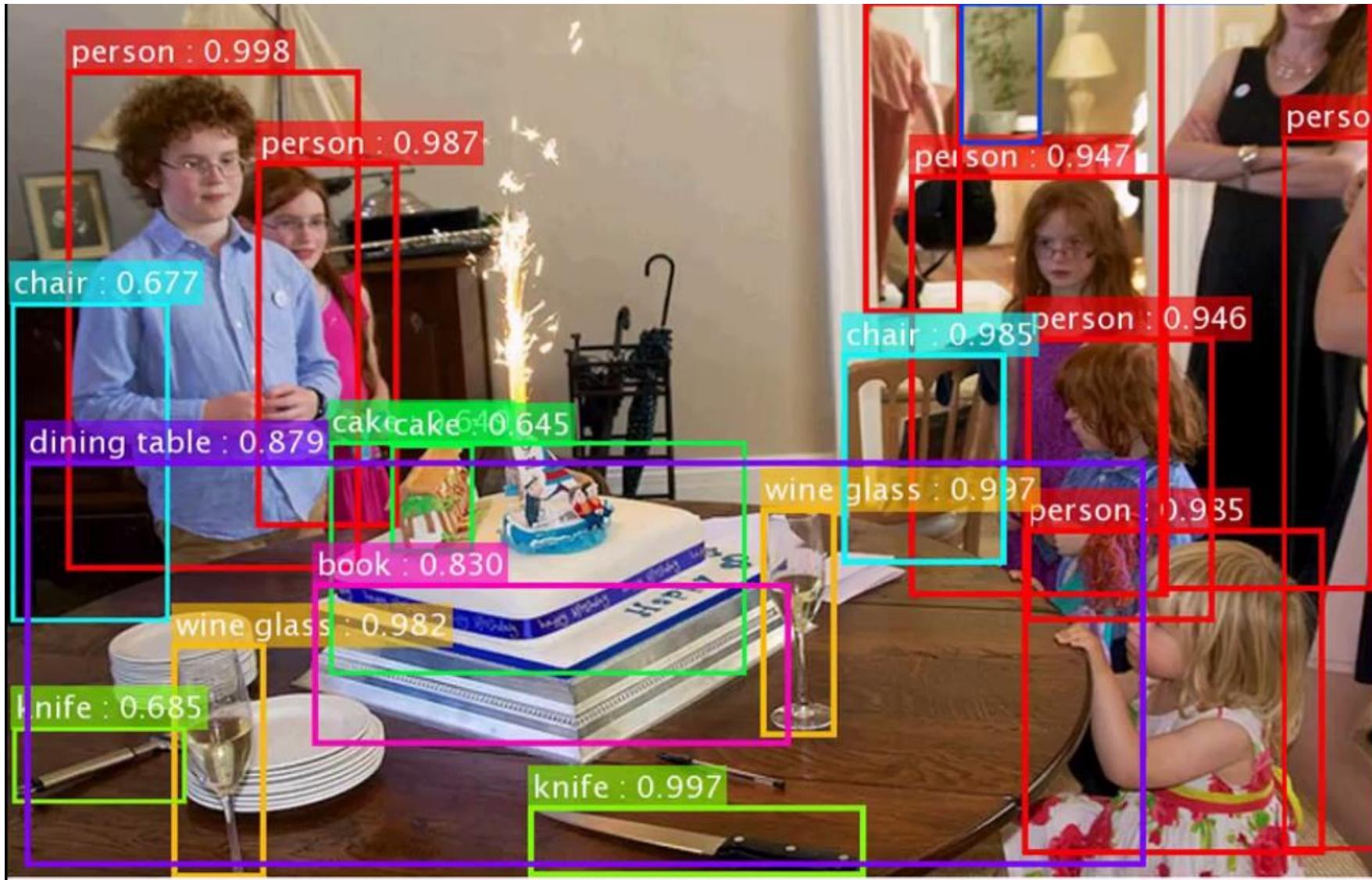
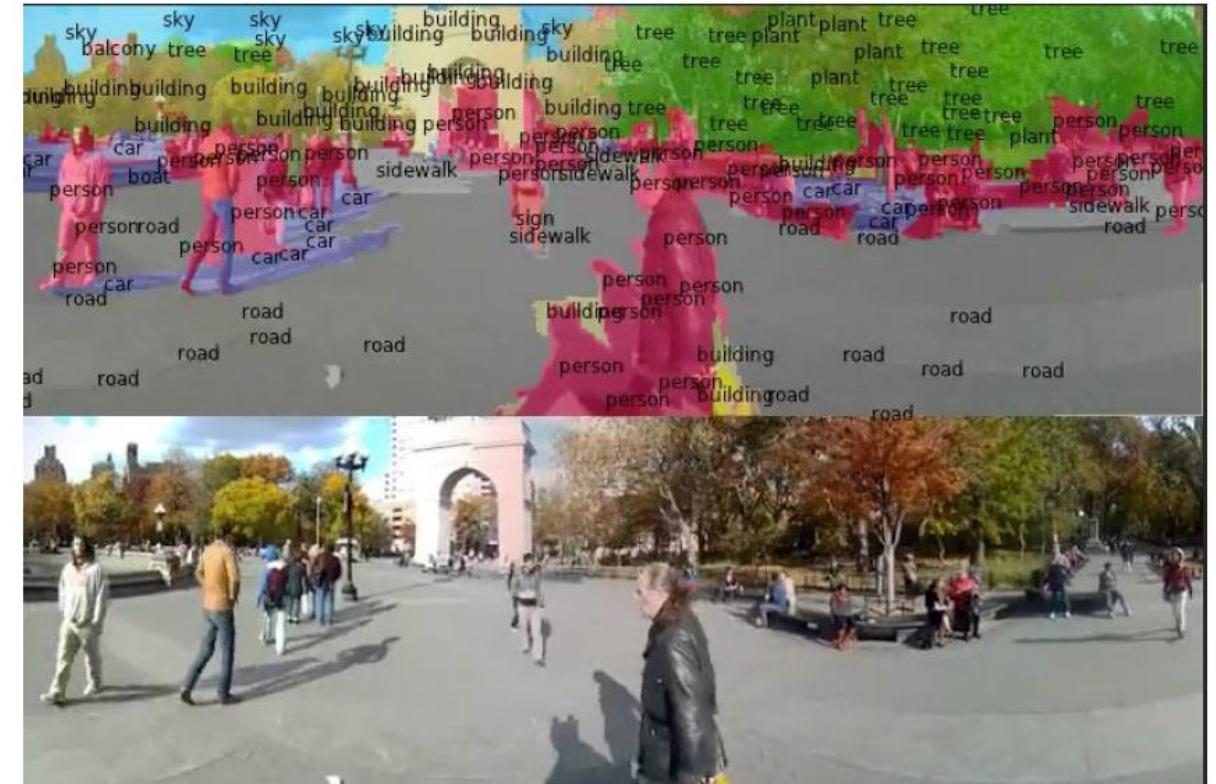


Image Retrieval and Segmentation

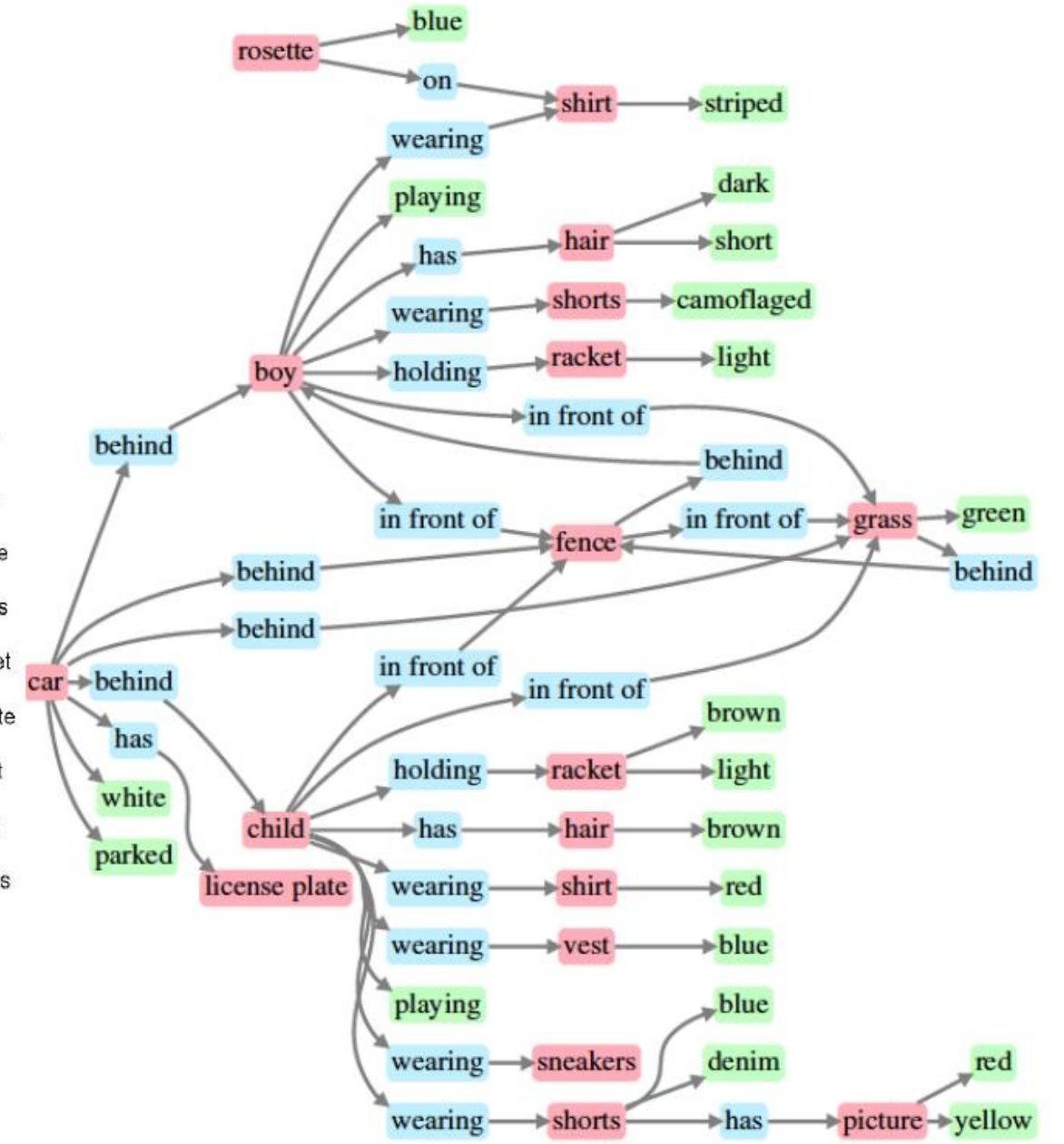
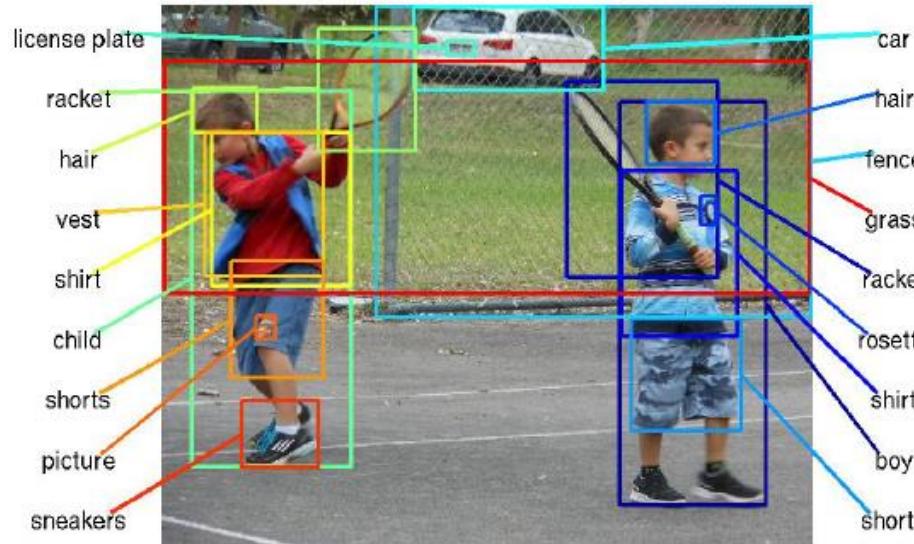
Retrieval



Segmentation



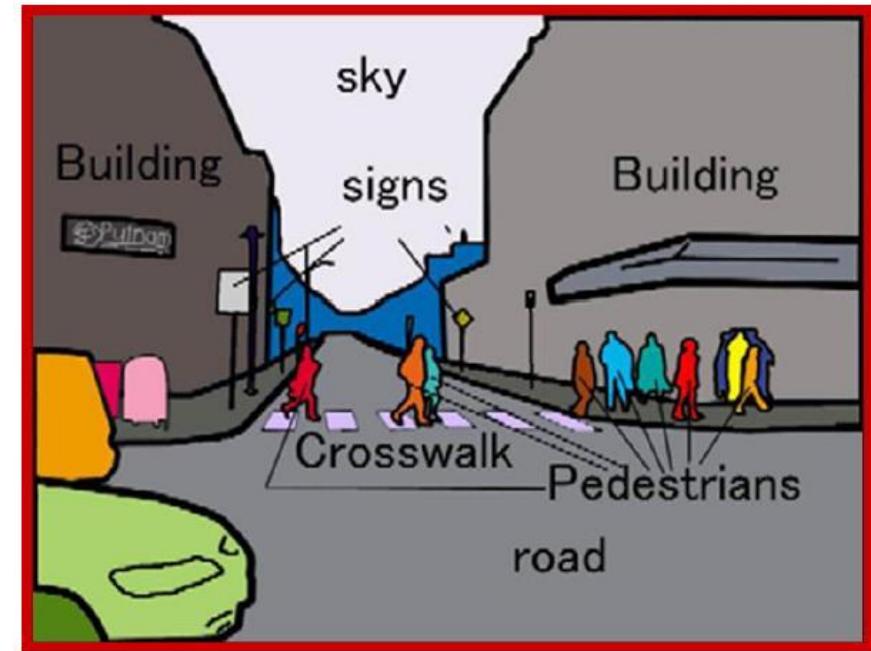
Scene Understanding



Driverless Cars



咱攢暫贊莊雕葬遭
擇刈澤賊怎增憎曾
派摘竚宅窄債塞瞻
湛綻樟章彙漳張哲
囉罩冰摩召遮折哲
針傾枕疼診震振鎮
鄭征艺枝支吱蜘知
止趾只旨紙高擎擲



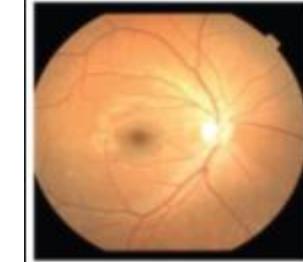
Medical Imaging



Input
Chest X-Ray Image

CheXNet
121-layer CNN

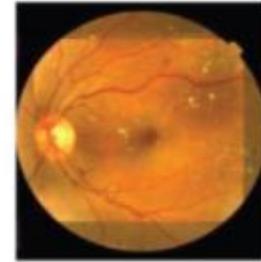
Output
Pneumonia Positive (85%)



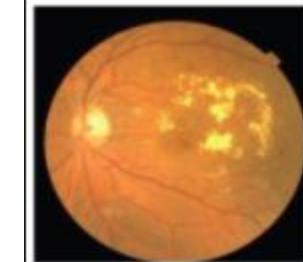
Without DR



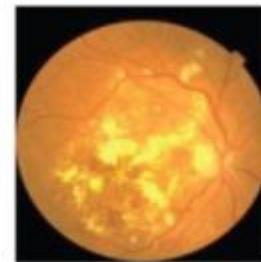
Early diabetic retinopathy



Mild NPDR



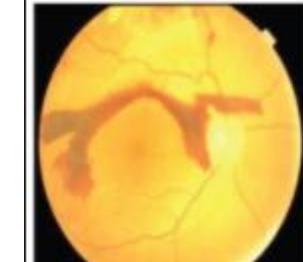
Moderate NPDR



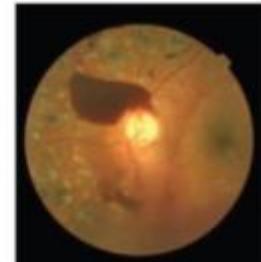
Severe NPDR



PDR and neovascularization



PDR with vitreous hemorrhage



PDR with vitreous hemorrhage and PLM



Vitreoretinal traction bands

Pose estimation



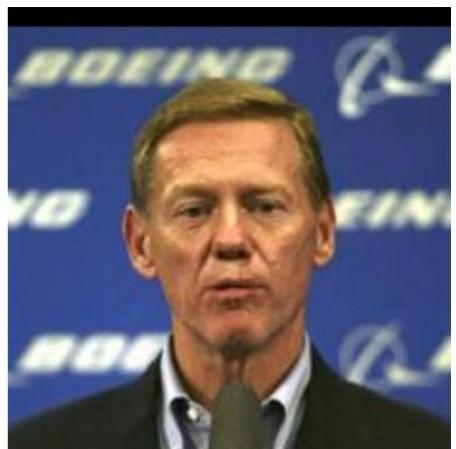
Face verification



Match



Not a Match



Using satellite images



Street-View to Overhead-View Matching

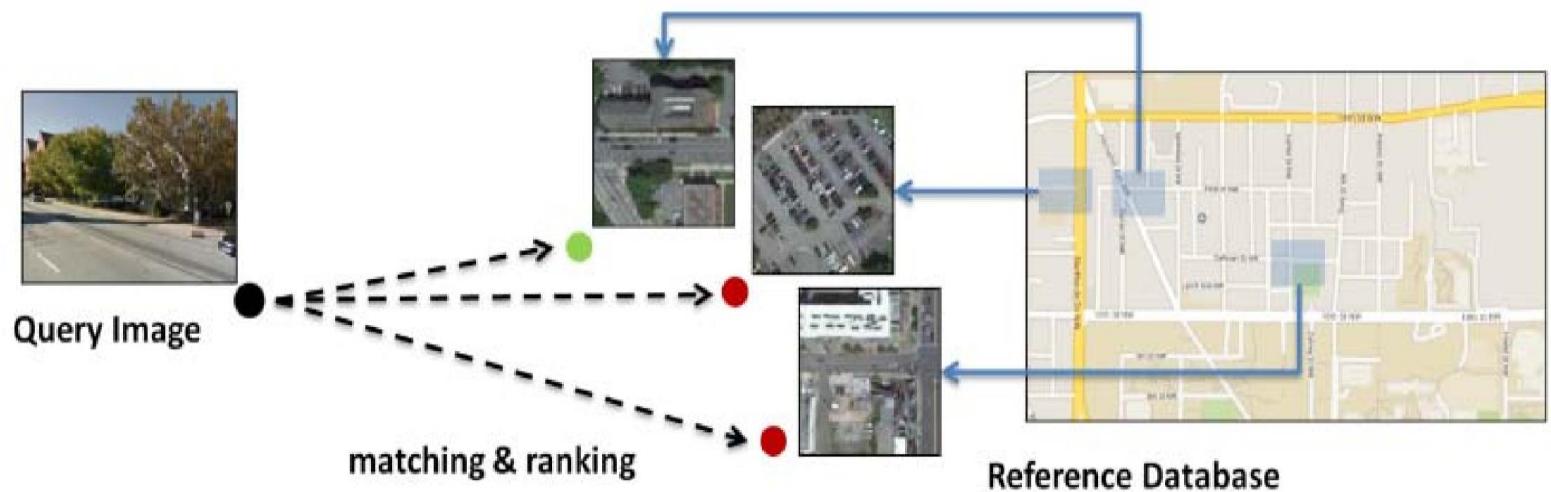


Image similarity



Image To Search

Closest Dress

Second Closest Dress

Third Closest Dress

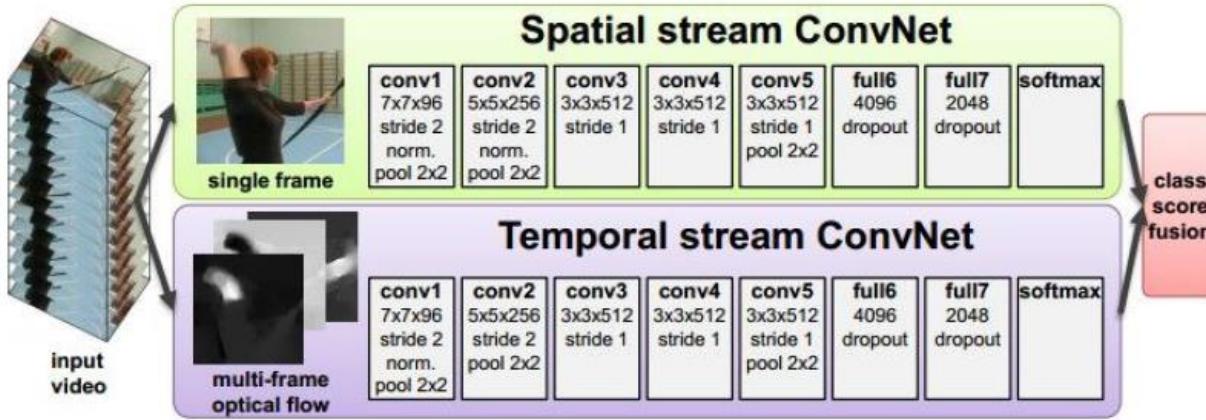
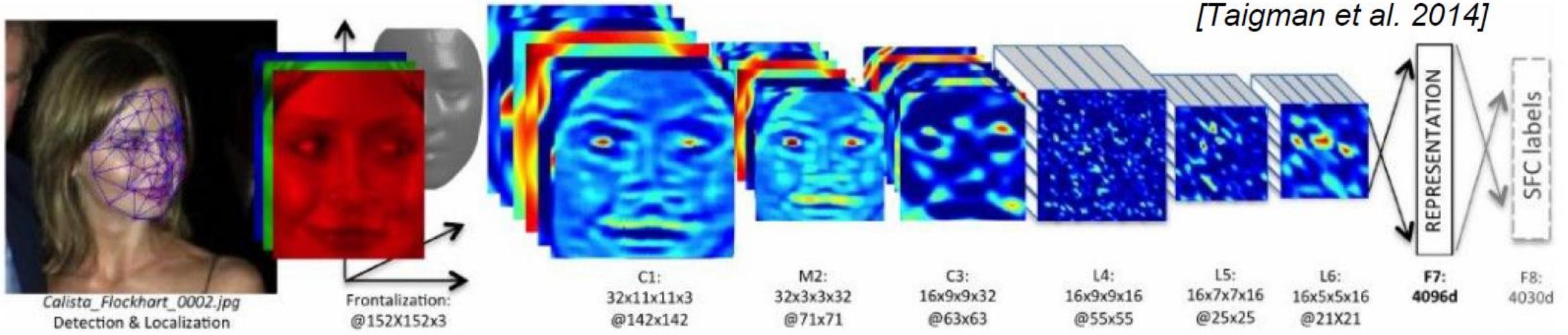
Fourth Closest Dress

Fifth Closest Dress

localhost:9000/similarity_finder_for/1098088517.jpg

A search interface showing the target image at the top left. Below it, five other dresses are displayed in a row, labeled from "Closest Dress" to "Fifth Closest Dress". Each dress is shown from the waist up, and they all appear to be variations of a blue and white floral print. The "Closest Dress" is identical to the target image. The "Second Closest Dress" has a slightly different pattern and a asymmetrical hem. The "Third Closest Dress" has a more dense pattern. The "Fourth Closest Dress" has a different pattern and a flared hem. The "Fifth Closest Dress" has a different pattern and a more fitted silhouette. A URL "localhost:9000/similarity_finder_for/1098088517.jpg" is visible at the bottom of the interface.

Face Recognition, Action recognition, House numbers



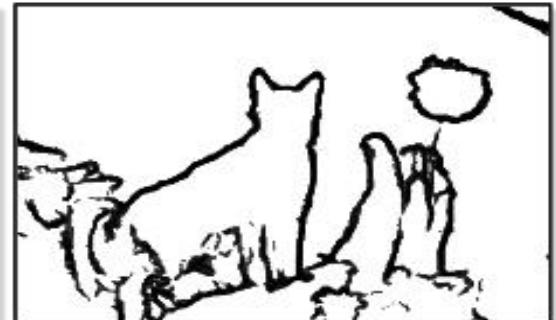
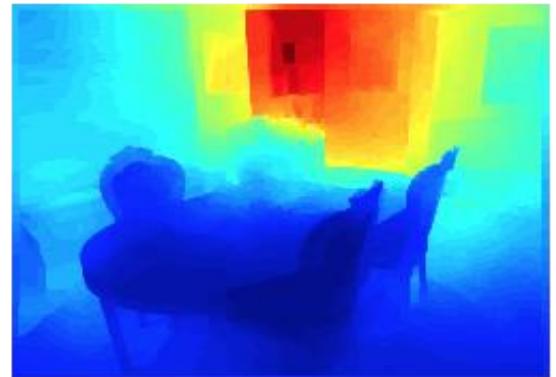
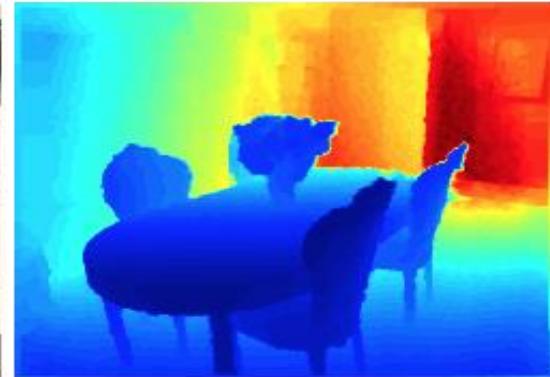
[Goodfellow 2014]

Dense Predictions

semantic segmentation



monocular depth estimation (Liu et al. 2015)



boundary prediction (Xie & Tu 2015)

Visual Question Answering

About objects



What color are her eyes?
What is the mustache made of?



Is this person expecting company?
What is just under the tree?



How many slices of pizza are there?
Is this a vegetarian pizza?

Counting

Fine-grained
recognition



Does it appear to be rainy?
Does this person have 20/20 vision?

Common sense

Text VQA, Knowledge VQA



What does it say near the star on the tail of the plane?

Ground Truth

jet

Prediction

nothing

(a)



What is the time on bottom middle phone?

Ground Truth

15:20

Prediction

12:00

(b)



What is the top oz?

Ground Truth

16

Prediction

red

(c)



What is the largest denomination on table?

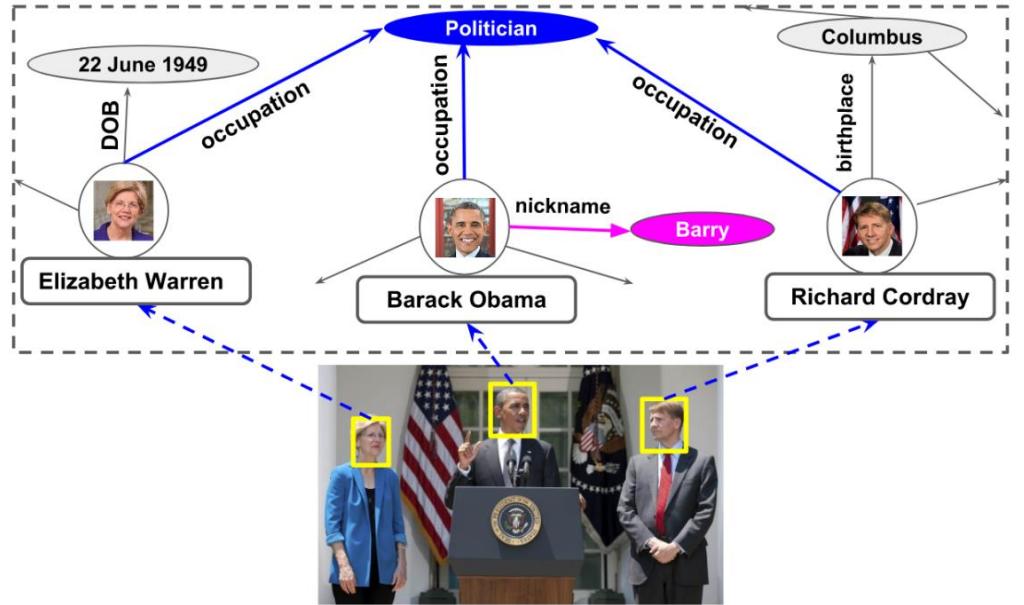
Ground Truth

500

Prediction

unknown

(d)



Conventional VQA (Antol et al. 2015; Goyal et al. 2017; Trott, Xiong, and Socher 2018)

Q: How many people are there in the image?

A: 3

Commonsense knowledge-enabled VQA (Wang et al. 2017; 2018; Su et al. 2018; G. Narasimhan and Schwing 2018)

Q: What in the image is used for amplifying sound?

A: Microphone

(World) knowledge-aware VQA (KVQA, this paper):

Q: Who is to the left of Barack Obama?

A: Richard Cordray

Q: Do all the people in the image have a common occupation?

A: Yes

Q: Who among the people in the image is called by the nickname Barry?

A: Person in the center



<https://visualdialog.org/>



A man and a woman are holding umbrellas.

Black

Multi-colored

3

What color is his umbrella?

What about hers?

How many other people are in the image?

How many are men?

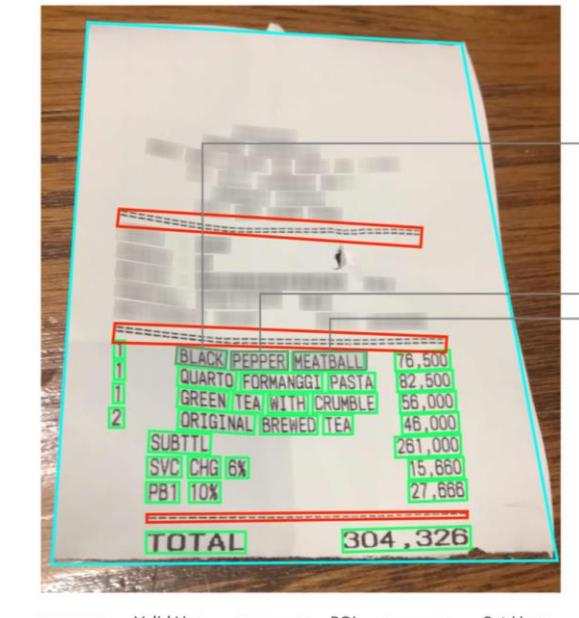


Start typing question here ...

Visually-rich Document Understanding

**SPORTS MARKETING ENTERPRISES
DOCUMENT CLEARANCE SHEET**

Date Routed:	January 11, 1994	Contract No.	4011 00 00
Contract Subject:	Joe's Place Exhibits		
Company	SPEVCO, INC.	Brand(s)	Camel/Winston
Total Contract Cost	\$1,340,000.00	Current Year Cost	1994-1995
Brief Description	2 Joe's Place Exhibits for use at Winston Cup, Winston Drag and Camel Super Bike Events.		
G/L Code:	Program Budget Code		
NAME	SIGNATURE	DATE	
Originator	Michael Wright		
Manager	John Powell	B. J. Powell	1-11-94
REVIEW ROUTING	SIGNATURE	DATE	
Insurance			
Law			
FS - Marketing			
REVISIONS TO SHELL (Other than Term, Compensation or Job)	PAGE(S)	SECTION(S)	
APPROVAL ROUTING	<ul style="list-style-type: none"> Sr. Manager (B. J. Powell) Director - (G. L. Littell) 		
** Sr. VP	T. W. Robertson		
Return To:	MARY SEAGRAVES Ext. 1485	SME	13 Plaza
51669 8130			
* UP TO AND INCLUDING \$25,000			
** OVER \$25,000			

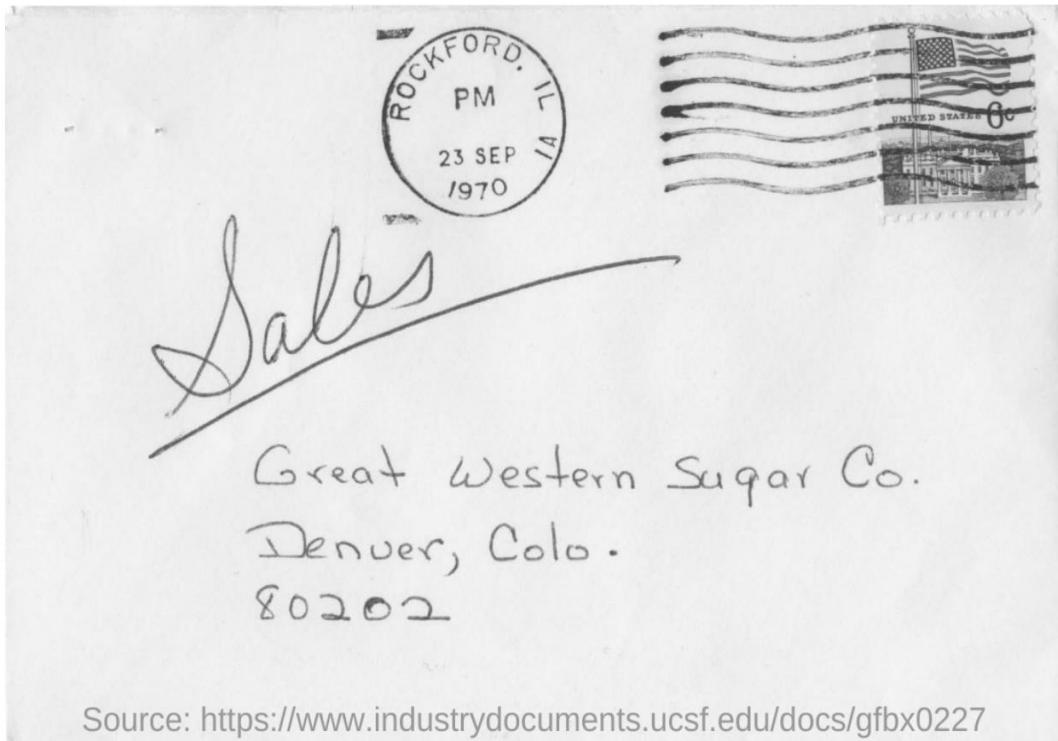


```

...
{
  "words": [
    {
      "quad": {"x2": 837, "y3": 2260, "x3": 829, "y4": 2238,
               "x1": 625, "y1": 2156, "x4": 617, "y2": 2177},
      "is_key": 0,
      "row_id": 1631316,
      "text": "BLACK"
    },
    {
      "quad": {"x2": 1105, "y3": 2261, "x3": 1105, "y4": 2261,
               "x1": 857, "y1": 2180, "x4": 857, "y2": 2180},
      "is_key": 0,
      "row_id": 1631316,
      "text": "PEPPER"
    },
    {
      "quad": {"x2": 1441, "y3": 2264, "x3": 1443, "y4": 2269,
               "x1": 1125, "y1": 2184, "x4": 1126, "y2": 2179},
      "is_key": 0,
      "row_id": 1631316,
      "text": "MEATBALL"
    }
  ],
  "category": "menu_nm",
  "group_id": 3
}
...

```

Visually-rich Document Understanding



Source: <https://www.industrydocuments.ucsf.edu/docs/gfbx0227>

Q: Mention the ZIP code written?

A: 80202

Q: What date is seen on the seal at the top of the letter?

A: 23 sep 1970

Q: Which company address is mentioned on the letter?

A: Great western sugar Co.

First page:

COVENANT NOT TO COMPETE
AND NON-DISCLOSURE AGREEMENT

PARTIES:

Charles D. Denson (EMPLOYEE)
and
NIKE, Inc. and its parent, divisions,
subsidiaries and affiliates. (NIKE):

RECITALS:

A. This Covenant Not to Compete and Non-Disclosure Agreement is executed upon the EMPLOYEE's advancement to the position of President of the NIKE brand and is a condition of such advancement.

B. Over the course of EMPLOYEE's employment with NIKE, EMPLOYEE will be or has been exposed to and/or is in a position to develop confidential information peculiar to NIKE's business and not generally known to the public as defined below ("Protected Information"). It is anticipated that EMPLOYEE will continue to be exposed to Protected Information of greater sensitivity as EMPLOYEE advances in the company.

C. The nature of NIKE's business is highly competitive and disclosure of any Protected Information would result in severe damage to NIKE and be difficult to measure.

D. NIKE makes use of its Protective Information throughout the world. Protective Information of NIKE can be used to NIKE's detriment anywhere in the world.

AGREEMENT:

In consideration of the foregoing, and the terms and conditions set forth below, the parties agree as follows:

- Covenant Not to Compete.**
(a) **Competition Restriction.** During EMPLOYEE's employment by NIKE, under the terms of any employment contract or otherwise, and for twelve (12) months thereafter, (the "Restriction Period"), EMPLOYEE will not directly or indirectly, own, manage, control, or participate in the ownership, management or control of, or be employed by, consult for, or be connected in any manner with, any business engaged anywhere in the world in the athletic footwear, athletic apparel or sports equipment and accessories business, or any other business which directly competes with NIKE or any of its parent, subsidiaries or affiliated corporations ("Competitor"). By way of *illustration only*, examples of NIKE competitors include, but are not limited to: Adidas, FILA, Reebok, Puma, Champion, Oakley, DKNY, Converse, Asics, Saucony, New Balance, Ralph Lauren/Polo Sport, B.U.M., FUBU, The Gap, Tommy Hilfiger, Umbro, Northface, Venator (Footlockers), Sports Authority, Columbia Sportswear, Wilson, Mizuno, Callaway Golf and Titleist. This provision is subject to NIKE's option to waive all or any portion of the Restriction Period as more specifically provided below.
(b) **Extension of Time.** In the event EMPLOYEE breaches this covenant not to compete, the Restriction Period shall automatically toll from the date of the first breach, and all subsequent

COVENANT NOT TO COMPETE AND
NON DISCLOSURE AGREEMENT - Page 1

Signature page:

(d) **Applicable Law/Jurisdiction.** This Agreement, and EMPLOYEE's employment hereunder, shall be construed according to the laws of the State of Oregon. EMPLOYEE further hereby submits to the jurisdiction of, and agrees that exclusive jurisdiction over and venue for any action or proceeding arising out of or relating to this Agreement shall lie in the state and federal courts located in Oregon.

EMPLOYEE

By: /s/ Charles D. Denson

Name: Charles D. Denson
Title: President, NIKE Brand

DATE: 3.26.01

NIKE, Inc.

By: /s/ PHILIP H. KNIGHT

Name: Philip H. Knight
Title: President & CEO

COVENANT NOT TO COMPETE AND
NON-DISCLOSURE AGREEMENT - Page 4

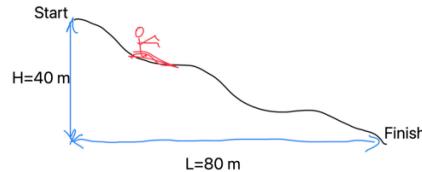
Verifying a student's solution to a physics problem.



Prompt

Here's a solution to a Physics problem by a student.

23. A skier slides down the frictionless slope as shown. What is the skier's speed at the bottom?



Solution:

The conservation of energy means that the energy of the skier is same at the beginning and at the end.

$$\text{Start: } E = mgL$$

$$\text{End: } E = \frac{1}{2}mv^2$$

$$\Rightarrow mgL = \frac{1}{2}mv^2$$

$$\Rightarrow v = \sqrt{2gL} = \sqrt{2 \times 9.81 \times 80} = 39.6 \text{ m.s}^{-1}$$

Try to reason about the question step by step.

Did the student get the correct answer? If the solution is wrong, please explain what is wrong and solve the problem.

Make sure to use LaTeX for math and round off the final answer to two decimal places.



Response

The student did not get the correct answer. The student made a mistake in the calculation of the potential energy at the start of the slope. The potential energy at the start of the slope is given by mgh , where m is the mass of the skier, g is the acceleration due to gravity, and h is the height of the slope. The student used the length of the slope, L , instead of the height, H , in the calculation of the potential energy.

The correct solution is:

The conservation of energy means that the total energy of the skier is the same at the beginning and at the end of the slope.

At the start of the slope, the skier has only potential energy, which is given by mgh .

At the end of the slope, the skier has only kinetic energy, which is given by $\frac{1}{2}mv^2$

Therefore, we can write:

$$mgh = \frac{1}{2}mv^2$$

Cancelling out the mass m on both sides, we get:

$$gh = \frac{1}{2}v^2$$

Solving for v , we get:

$$v = \sqrt{2gh}$$

Substituting the given values, we get:

$$v = \sqrt{2 \times 9.81 \times 40} = 28.01 \text{ m/s}$$

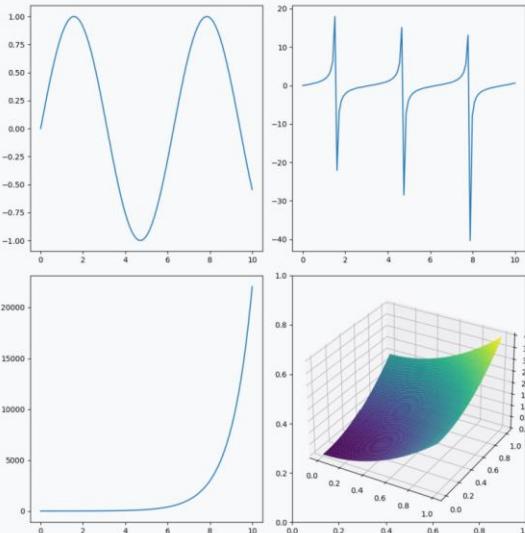
Therefore, the skier's speed at the bottom of the slope is 28.01 m/s.

(1) Recognition of functions depicted in the plots; (2) inverse graphics to infer the code that would have generated the subplots; (3) instruction-following to put subplots in their desired positions; and (4) abstract reasoning to infer that the exponential plot must stay in its original place.



Prompt

Write code to rearrange the subplots in the figure using the latest version of matplotlib. Put the 3d paraboloid subplot on the top left. Put the tangent function on the bottom right. For the remaining two subplots, one of them should stay in its original position and the other should fill the last spot. First describe what each subplot depicts and identify its current location. Then, explain where each subplot should go in its new location. Last write the full code for the rearranged version with the original color scheme.



Response

The current subplots are:

- Top left: A sine wave
- Top right: A tangent function
- Bottom left: An exponential function
- Bottom right: A 3D paraboloid

The new subplots should be:

- Top left: The 3D paraboloid
- Top right: The sine wave
- Bottom left: The exponential function
- Bottom right: The tangent function

```
import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D

# Create a figure with 2 rows and 2 columns
fig = plt.figure(figsize=(10, 10))

# Add a 3D paraboloid subplot to the top left
ax = fig.add_subplot(2, 2, 1, projection='3d')
x = np.linspace(0, 1, 100)
y = np.linspace(0, 1, 100)
X, Y = np.meshgrid(x, y)
Z = X**2 + Y**2
ax.plot_surface(X, Y, Z, cmap='viridis')

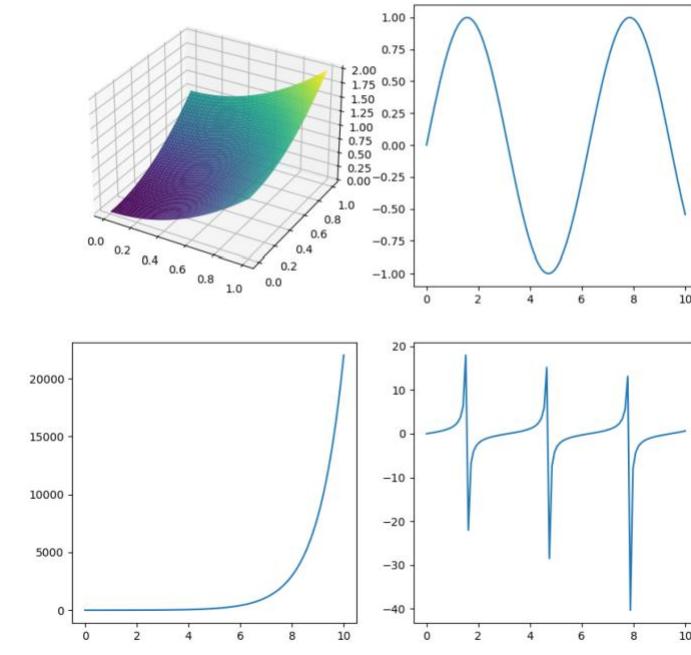
# Add a sine wave subplot to the top right
ax = fig.add_subplot(2, 2, 2)
x = np.linspace(0, 10, 100)
y = np.sin(x)
ax.plot(x, y)

# Add an exponential function subplot to the bottom left
ax = fig.add_subplot(2, 2, 3)
x = np.linspace(0, 10, 100)
y = np.exp(x)
ax.plot(x, y)

# Add a tangent function subplot to the bottom right
ax = fig.add_subplot(2, 2, 4)
x = np.linspace(0, 10, 100)
y = np.tan(x)
ax.plot(x, y)
```



Rendered code



What can you do with Emu?

Input Prompt

Completion

In-context Completion

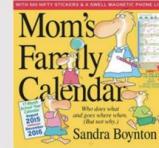
Image Captioning



The image shows a photo of

the movie poster for "Fight Club." The poster features two men standing next to each other, with the words "Fight Club" written in bold letters above them. The background of the poster is dark and ominous, giving the impression of a gritty and intense movie.

Image Question Answering



Question: Can you tell me about this image? Answer:

this image is the cover of a book called "Mom's Family Calendar" by Sandra Boynton. The book is designed to help moms keep track of their family's schedule and activities.

Input Prompt



There is a chinchilla.



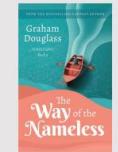
There are three pandas.

Completion



There is a godzilla.

Emu



This is the first image.



This is the second image.



This is the third image.



This is the fourth image.

Describe all images.

Emu

The first image is the cover of the book "The Way of the Nameless". The second image is the cover of the book "Lost Stars". The third image is the cover of the book "A Land of My Dreams". The fourth image is the cover of the book "The Key to Mindfulness".

What can you do with Emu?

Image-to-Image

Input Prompt



Completion

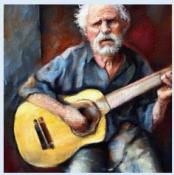


Text-to-Image

Input Prompt



Completion



Video Question Answering

Input Prompt



Completion



Question: What's the woman doing in the video? Answer:

The woman in the video is wearing a pair of Apple VR goggles while sitting in an airplane. She appears to be using the goggles to watch something, possibly a video or a 360-degree view of the airplane's surroundings.



Question: What is unusual about this video? Answer:

The unusual aspect of this video is that it features a giraffe standing on a diving board and jumping into a swimming pool. This is not a common sight, as giraffes are not typically found in swimming pools.

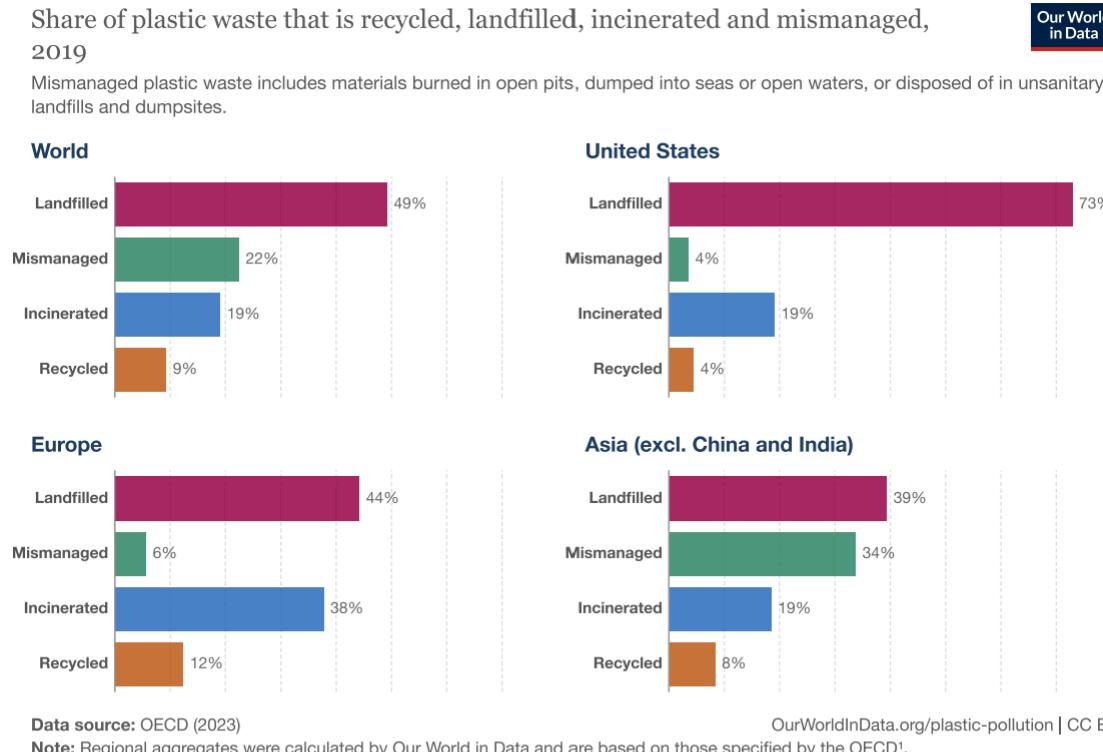


Step by step describe the cooking recipe shown in the video?



The video shows a step-by-step guide on how to make a chocolate milkshake. First, the person in the video adds ice cubes to a blender. Then, they add two scoops of chocolate ice cream and one scoop of vanilla ice cream to the blender. After that, they add three tablespoons of chocolate syrup and two tablespoons of cocoa powder to the blender. Finally, they blend all the ingredients together until the mixture is smooth and creamy. The video ends with the person pouring the chocolate milkshake into a glass and garnishing it with whipped cream and chocolate sprinkles. Overall, the video provides a simple and easy recipe for making a delicious chocolate milkshake.

Chart understanding and reasoning over data



Our World
in Data

Model Response (rendered Markdown)

The data point that stands out is the United States' landfilled plastic waste, which is 73%. This is significantly higher than the other countries and regions shown. This implies that the United States is not doing as much to recycle or incinerate its plastic waste as other countries and regions.

Country/Region	Landfilled (%)	Mismanaged (%)	Incinerated (%)	Recycled (%)
World	49	22	19	9
United States	73	4	19	4
Europe	44	6	38	12
Asia (excl. China and India)	39	34	19	8

¹.OECD regions: The definitions of regions, as stipulated by the OECD, are: - Other OECD America: Chile, Colombia, Costa Rica, Mexico - OECD EU countries : Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden - OECD Non-EU countries: Iceland, Israel, Norway, Switzerland, Turkey, United Kingdom - OECD Oceania: Australia, New Zealand - OECD Asia: Japan, Korea - Latin America: Non-OECD Latin American and Caribbean countries - Other EU: Bulgaria, Croatia, Cyprus, Malta, Romania - Other Eurasia: Non-OECD European and Caspian countries, including Russian Federation - Middle East & North Africa: Algeria, Bahrain, Egypt, Iraq, Islamic Rep. of Iran, Kuwait, Lebanon, Libya, Morocco, Oman, Qatar, Saudi Arabia, Tunisia, United Arab Emirates, Syrian Arab Rep., Western Sahara, Yemen - Other Africa: Sub-Saharan Africa - China: People's Republic of China, Hong Kong (China) - Other non-OECD Asia: Other non-OECD Asian and Pacific countries

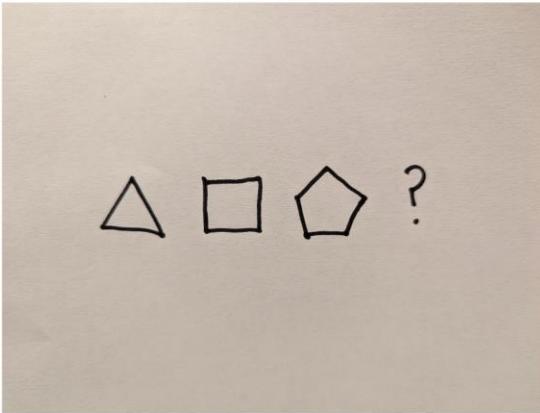
**Spot a data point that stands out in these charts and what that implicates.
Then produce a detailed markdown table for all the data shown.**

[Gemini: A Family of Highly Capable Multimodal Models. Gemini Team, Google, 2023.](#)

Image understanding and reasoning

Geometrical reasoning

Prompt

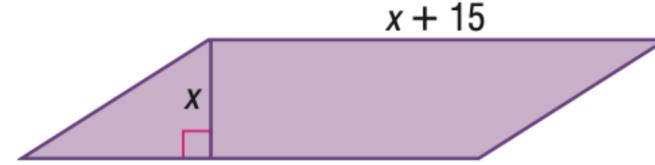


Look at this sequence of three shapes. What shape should come as the fourth shape? Explain your reasoning with detailed descriptions of the first shapes.

Model Response

The fourth shape should be a hexagon. The first shape is a triangle, the second shape is a square, and the third shape is a pentagon. The number of sides in each shape is increasing by one

Prompt



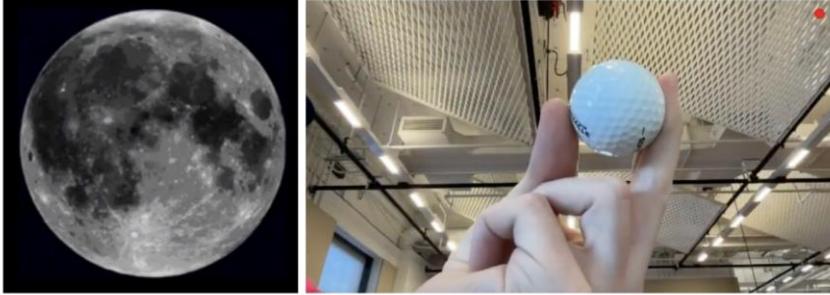
Find the height of the parallelogram given its area with 100 square units.

Model Response

The area of the parallelogram is equal to the product of the base and the height. Hence $100 = (x + 15)x$. We get $x^2 + 15x - 100 = 0$. The roots of this equation are $x_1 = 5$ and $x_2 = -20$. The second root is negative and is not acceptable. Final Answer: The final answer is 5. I hope it is correct.

Information seeking about objects

Prompt



Find a connection between these. Hint: think about historical events.

Model Response

The Moon is the only celestial body in which humans have played golf. In 1971, the Apollo 14 crew hit two golf balls on the lunar surface.

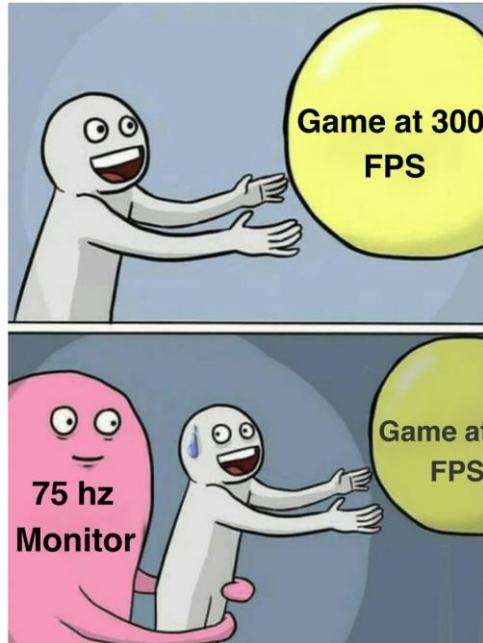
Multimodal reasoning based on visual cues

- Prompt: Do you know the precise location where this image was taken?
- Model Response: This image was taken in New York City. The Empire State Building is visible in the background. The street in the foreground is 8th Avenue. The cross street is West 34th Street.



Multimodal humor understanding

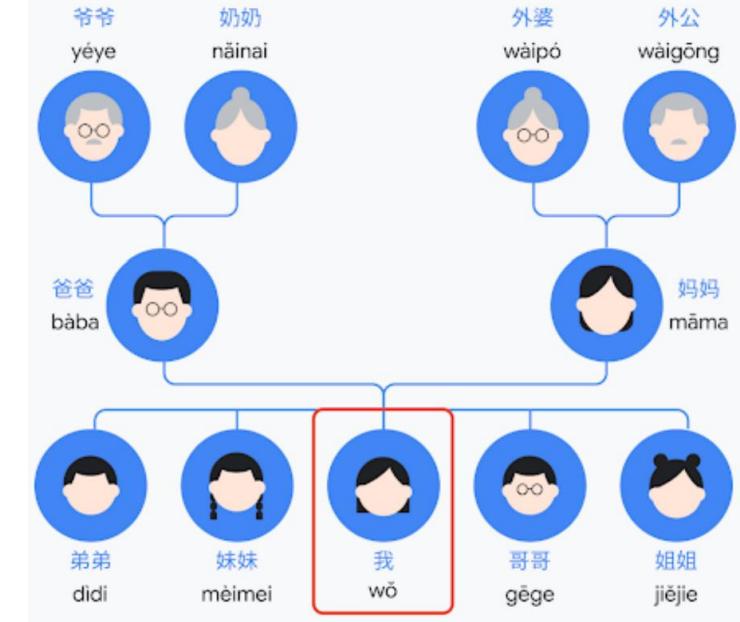
- Prompt: Can you explain this meme?
- Model Response: The meme is about a person who is playing a game at 300 FPS, but their monitor is only 75Hz. The person is trying to catch the game, but it is too fast for their monitor to display. The meme is funny because it is relatable to anyone who has ever played a game on a monitor that is not fast enough.



Commonsense reasoning in a multilingual setting

- Prompt: Explain what I should call my grandparents on each side of my family based on this chart
- Model Response: On your father's side, you would call your grandparents Yéye and Nainai. On your mother's side, you would call your grandparents Wàigong and Wàipó.

The Basic Chinese Family Tree



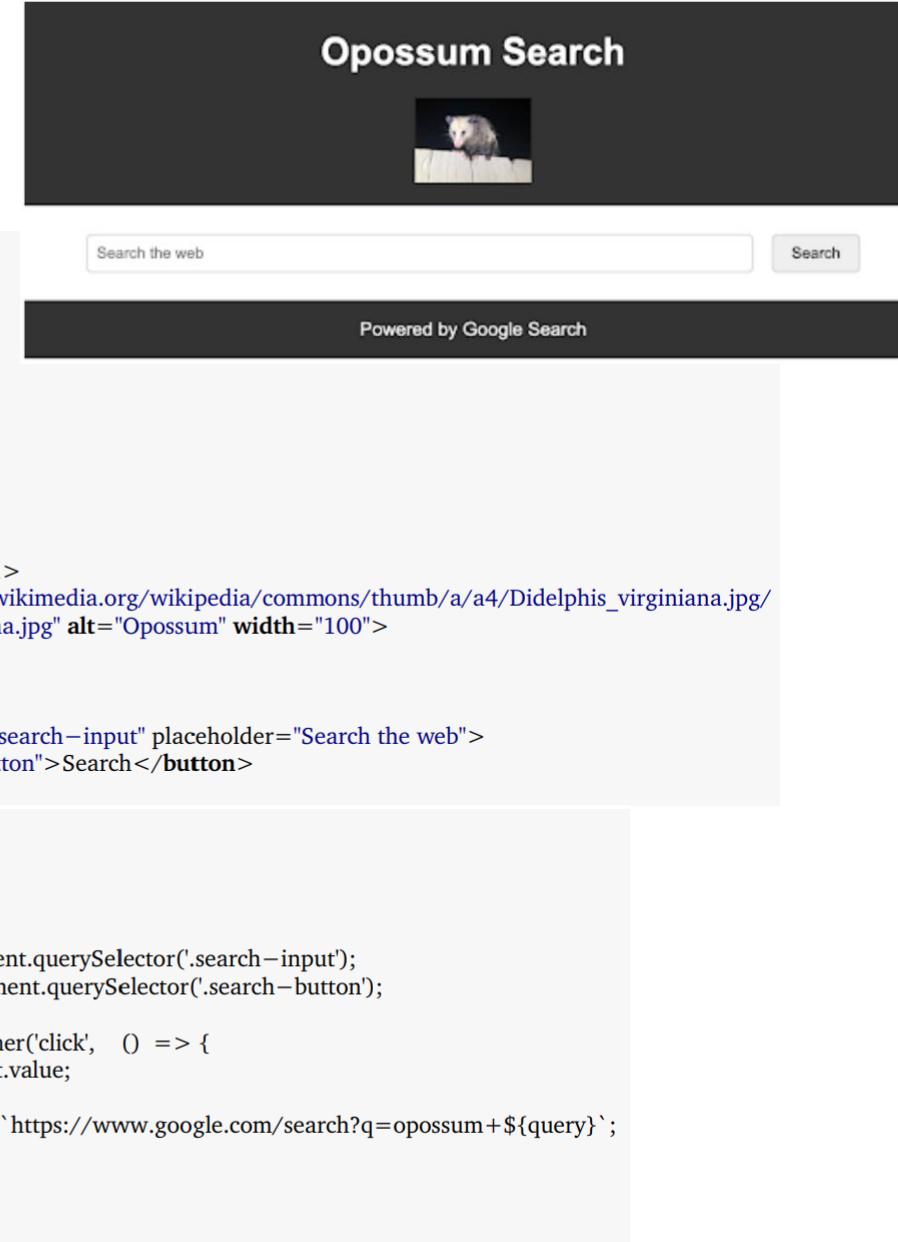
Reasoning and code generation

Create a web app called "Opossum Search":

1. Every time you make a search query, it should redirect you to a google search with the same query, but the word opossum before it.
2. It should be visually similar to Google search,
3. Instead of the google logo, it should have a picture of an opossum from the internet.
4. It should be a single html file, no separate js or css files.
5. It should say "Powered by google search" in the footer

```
1  <!DOCTYPE html>
2  <html>
3  <head>
4  <title>Opossum Search</title>
5  <style>
6  body {
7      font-family: Arial, sans-serif;
8      margin: 0;
9      padding: 0;
10 }
11
12 header {
13     background-color: #333;
14     color: #fff;
15     text-align: center;
16     padding: 1rem 0;
17 }
18
19 .search-bar {
20     display: flex;
21     justify-content: center;
22     align-items: center;
23     height: 50px;
24     margin-top: 1rem;
25 }
26
27 .search-input {
28     width: 50%;
29     padding: 0.5rem;
30     border: 1px solid #ccc;
31     border-radius: 5px;
32 }
33
34 .search-button {
35     margin-left: 1rem;
36     padding: 0.5rem 1rem;
37     border: 1px solid #ccc;
38     border-radius: 5px;
39     cursor: pointer;
40 }

41 footer {
42     background-color: #333;
43     color: #fff;
44     text-align: center;
45     padding: 1rem 0;
46     margin-top: 1rem;
47 }
48
49 </style>
50 </head>
51 <body>
52 <header>
53     <h1>Opossum Search</h1>
54     
55 </header>
56
57 <div class="search-bar">
58     <input type="text" class="search-input" placeholder="Search the web">
59     <button class="search-button">Search</button>
60 </div>
61
62 <footer>
63     Powered by Google Search
64 </footer>
65
66 <script>
67     const searchInput = document.querySelector('.search-input');
68     const searchButton = document.querySelector('.search-button');
69
70     searchButton.addEventListener('click', () => {
71         const query = searchInput.value;
72         if (query) {
73             window.location.href = `https://www.google.com/search?q=opossum+${query}`;
74         }
75     });
76 </script>
77 </body>
```



Mathematics: Calculus

Prompt

Find the derivative of $\sinh x + \cosh y = x + y$.

Model Response (rendered latex)

We have that $\frac{d}{dx} \sinh x = \cosh x$ and $\frac{d}{dx} \cosh y = \sinh y \frac{dy}{dx}$. Therefore,

$$\cosh x + \sinh y \frac{dy}{dx} = 1 + \frac{dy}{dx}.$$

Solving for $\frac{dy}{dx}$, we find that

$$\frac{dy}{dx} = \frac{1 - \cosh x}{\sinh y - 1}.$$

Final Answer: The final answer is $\frac{1 - \cosh x}{\sinh y - 1}$.

Video understanding and reasoning



- **Prompt (video)**
 - How could this person improve their technique?
- **Model Response**
 - This soccer player needs to work on their ball-striking mechanics. Their foot is not making good contact with the ball, which is causing the ball to fly up and away from the goal. They also need to work on their body positioning. They should be leaning into the shot with their non-kicking leg in front of their kicking leg, and they should be following through with their kicking leg.

Multi-Object Tracking



Sports Analytics

Traffic monitoring

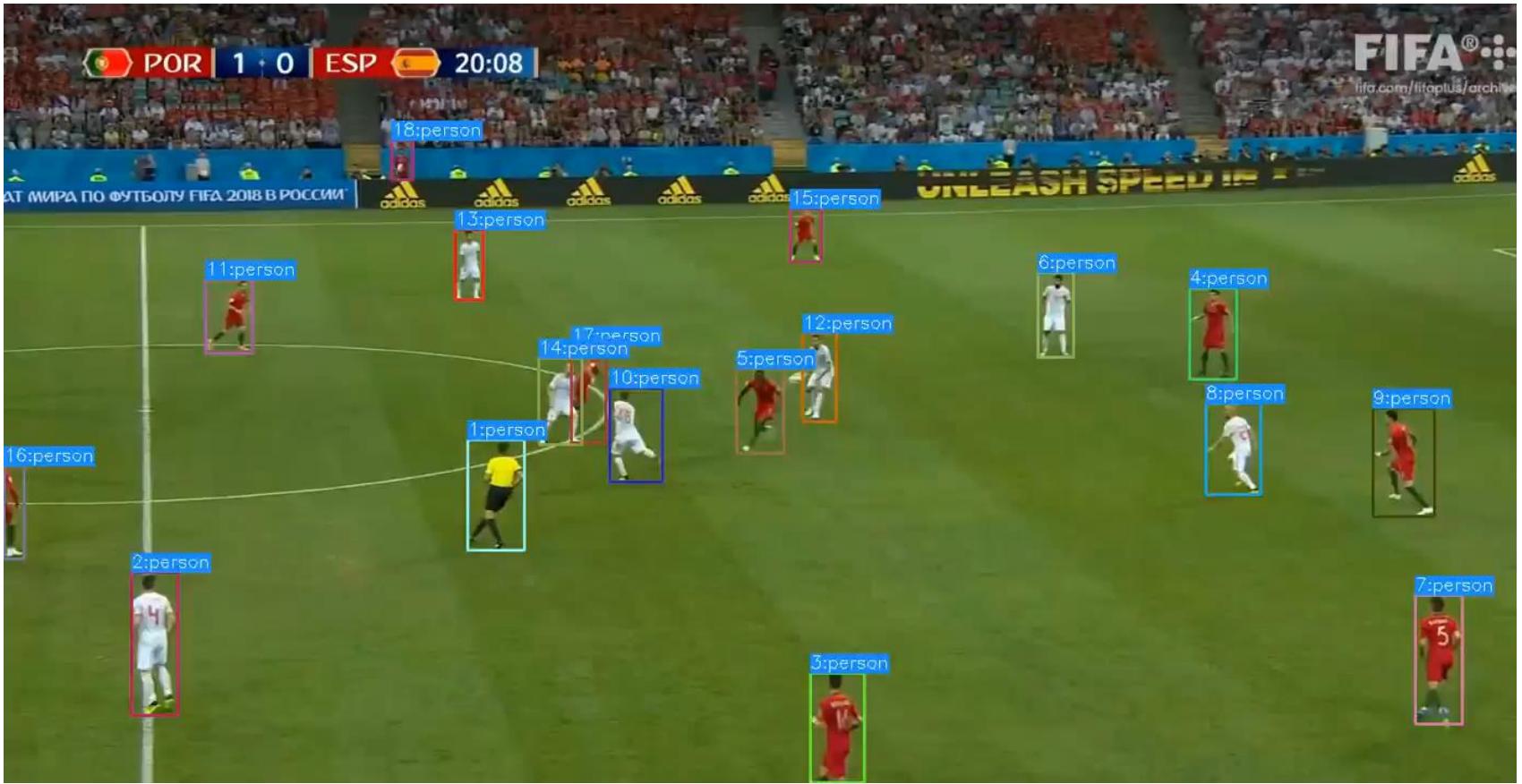


Image Generation using DALL-E

a pentagonal green clock. a green clock in the shape of a pentagon.



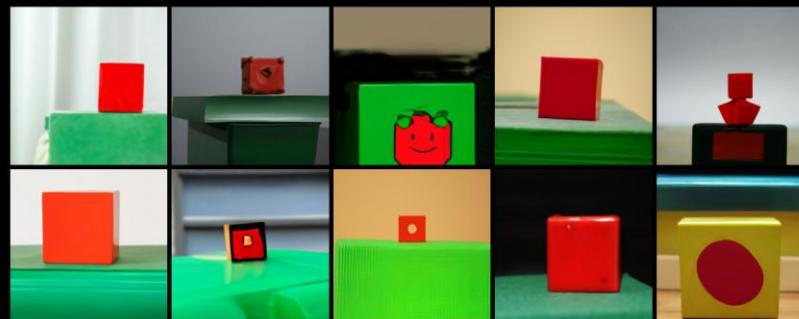
a cube made of porcupine. a cube with the texture of a porcupine.



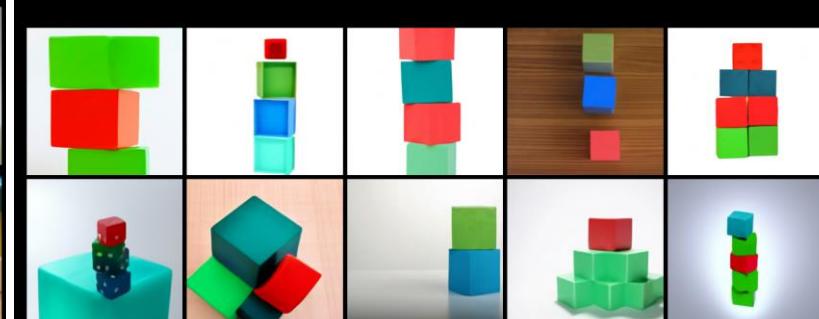
a collection of glasses is sitting on a table



a small red block sitting on a large green block



a stack of 3 cubes. a red cube is on the top, sitting on a green cube. the green cube is in the middle, sitting on a blue cube. the blue cube is on the bottom.



an emoji of a baby penguin wearing a blue hat, red gloves, green shirt, and yellow pants



Image Generation using DALL-E: Visualizing Perspective and 3D; Inferring Contextual Details

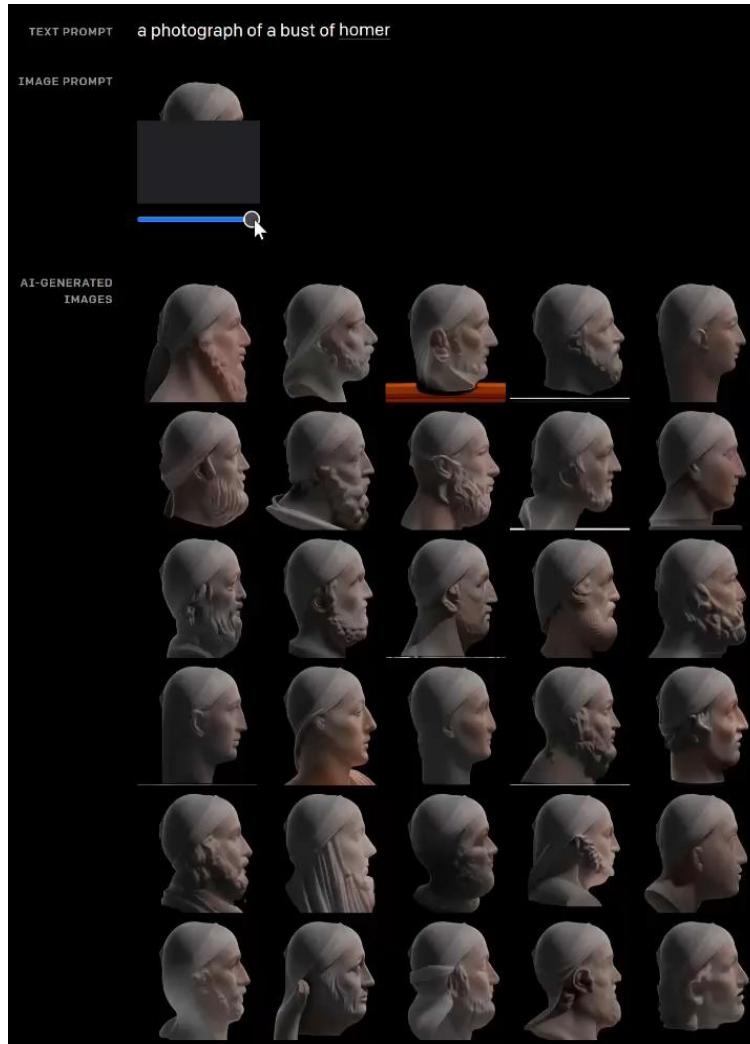
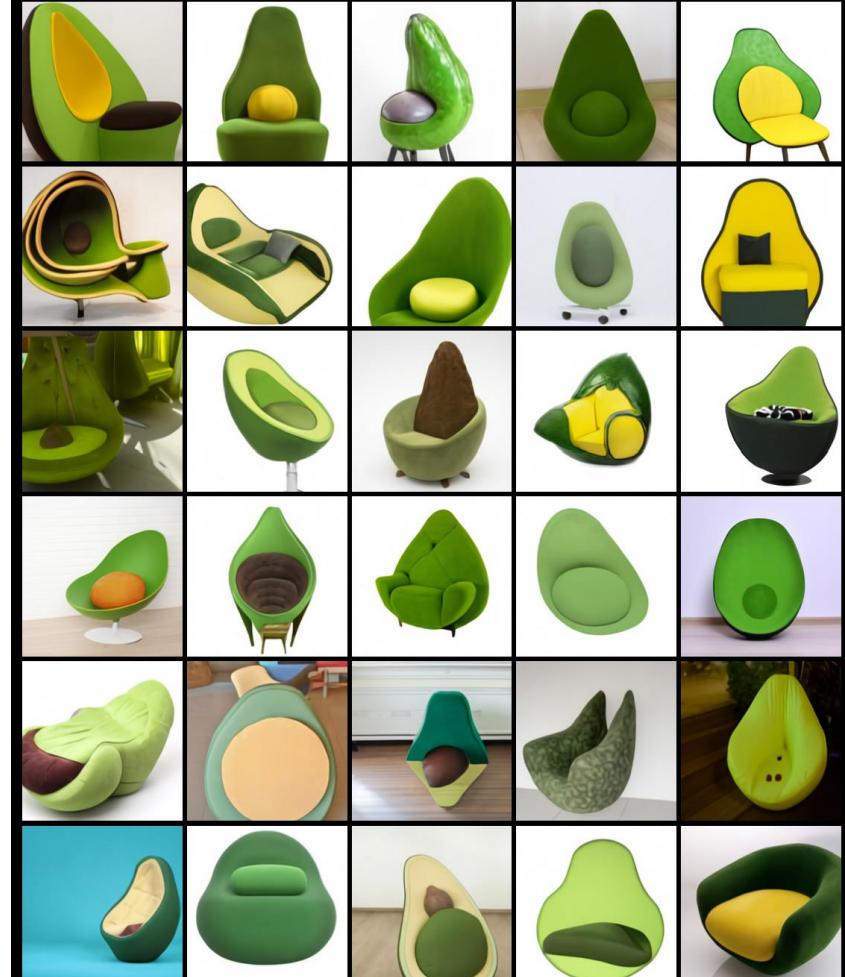


Image Generation using DALL-E: Combining Unrelated Concepts; Geographic and Temporal Context

an armchair in the shape of an avocado. an armchair imitating an avocado.



a photo of the flag of india

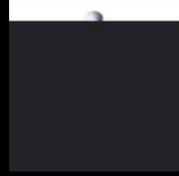


a photo of a phone from the ...

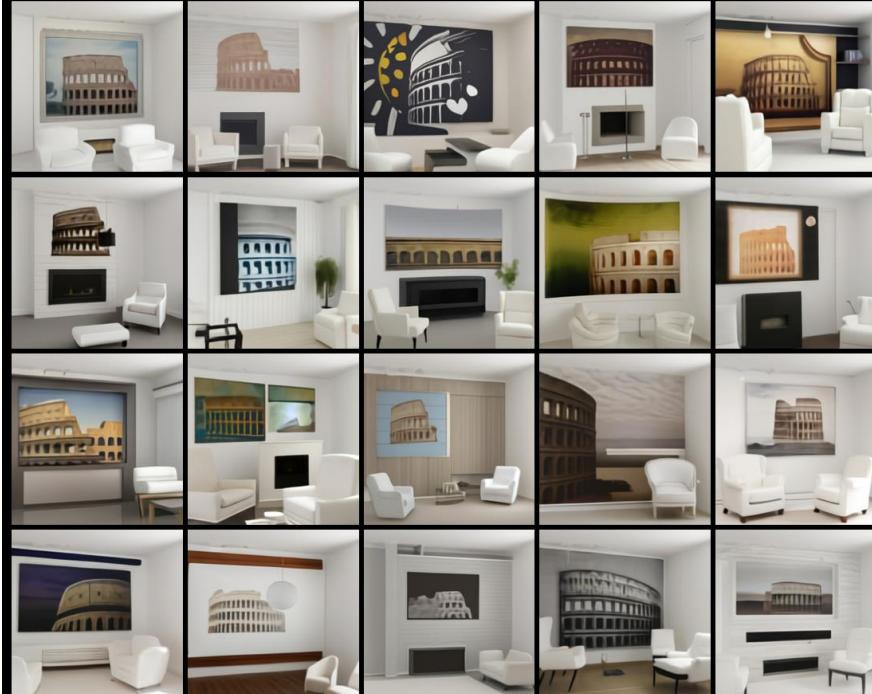


Image Generation using DALL-E: Applications to fashion and interior design

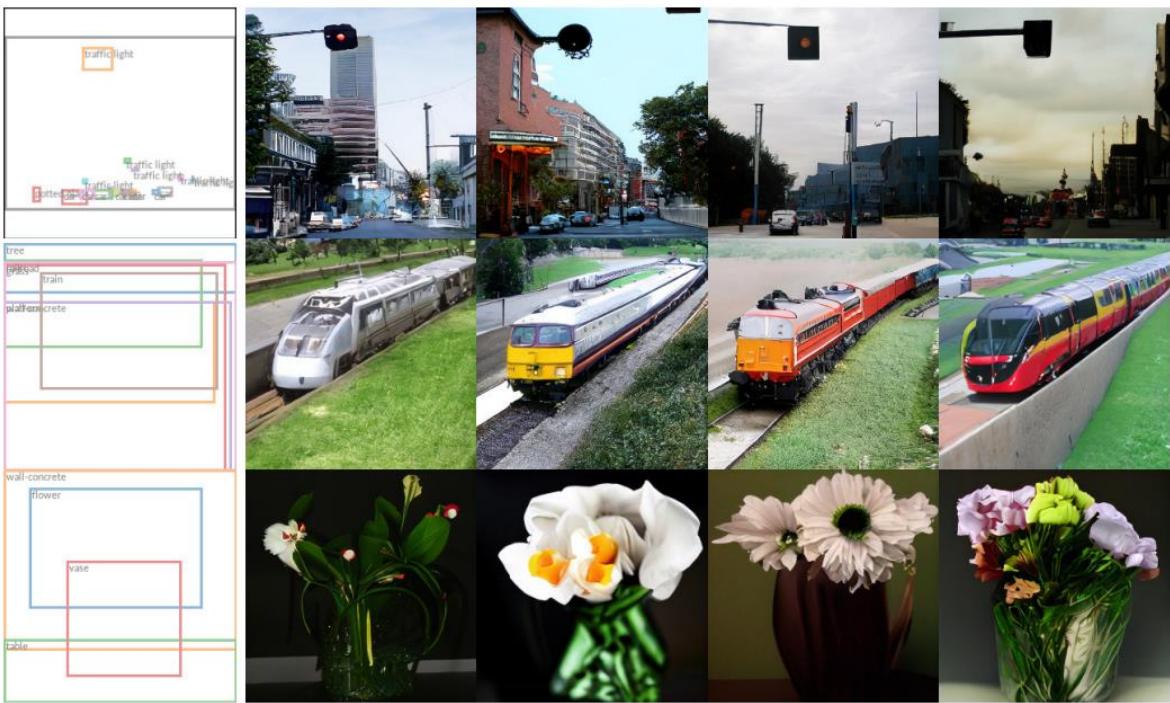
a female mannequin dressed in a black leather jacket and gold pleated skirt



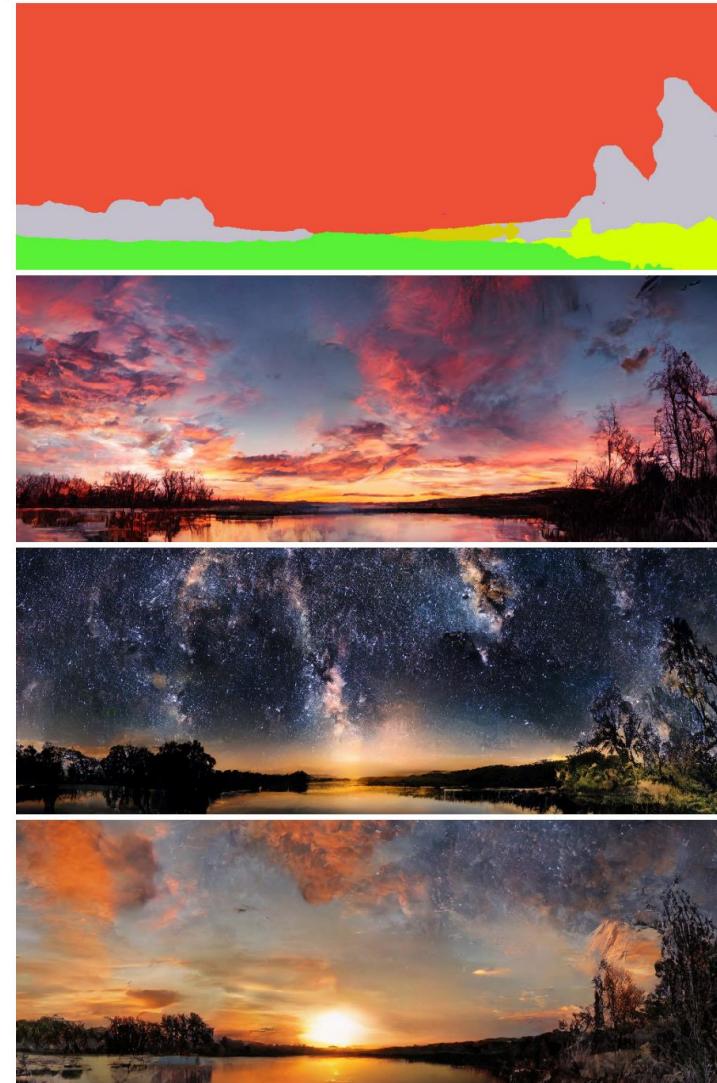
a living room with two white armchairs and a painting of the colosseum. the painting is mounted above a modern fireplace.



Layout-to-image synthesis



Semantic synthesis



Rombach, Robin, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. "High-resolution image synthesis with latent diffusion models." In CVPR, pp. 10684-10695. 2022.

Super-resolution

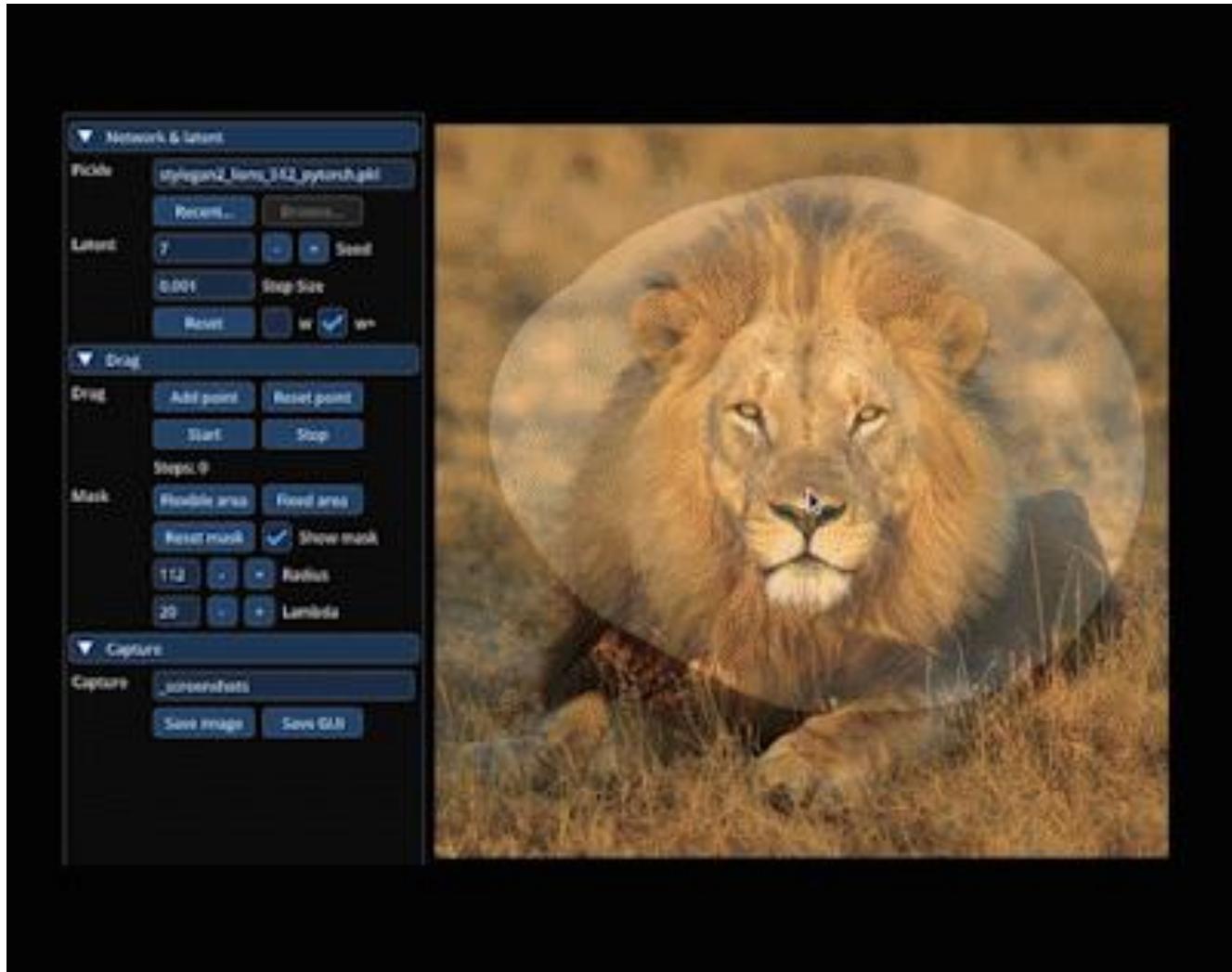


Inpainting



Rombach, Robin, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. "High-resolution image synthesis with latent diffusion models." In CVPR, pp. 10684-10695. 2022.

Awesome image manipulations

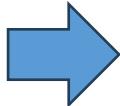


Drawing paintings, replacing objects



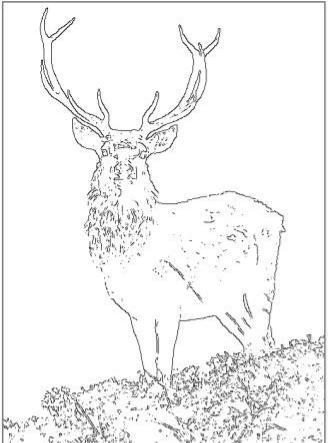
Text + Image + Audio → Video + Audio

"Teddy bear on a skateboard, 4k,
high resolution"



Tang, Zineng, Ziyi Yang, Chenguang Zhu, Michael Zeng, and Mohit Bansal. "Any-to-Any Generation via Composable Diffusion." arXiv preprint arXiv:2305.11846 (2023).

Controlling output style



Input Canny edge



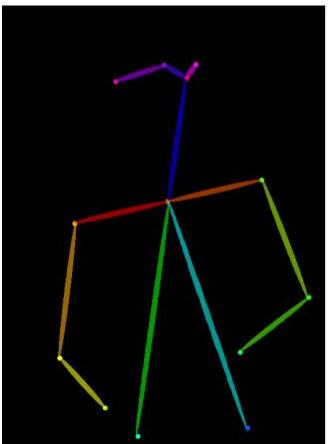
Default



"masterpiece of fairy tale, giant deer, golden antlers"



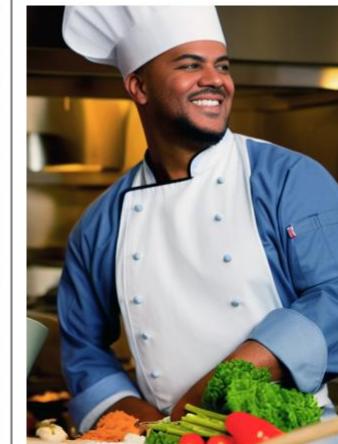
"..., quaint city Galic"



Input human pose



Default

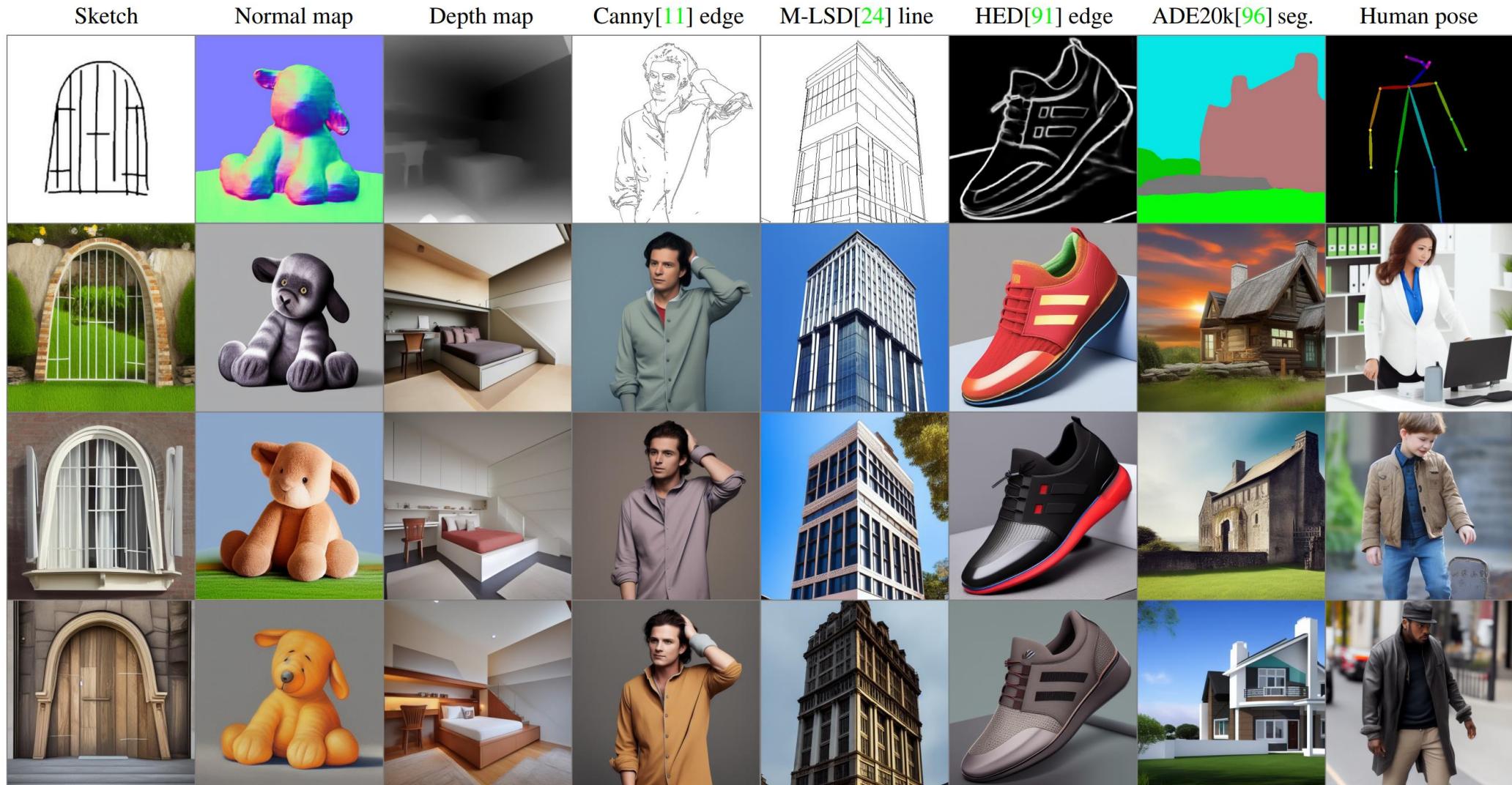


"chef in kitchen"



"Lincoln statue"

Controlling output style



Dialog in the context of videos



Describe what you hear?



Summarize this video in one sentence.



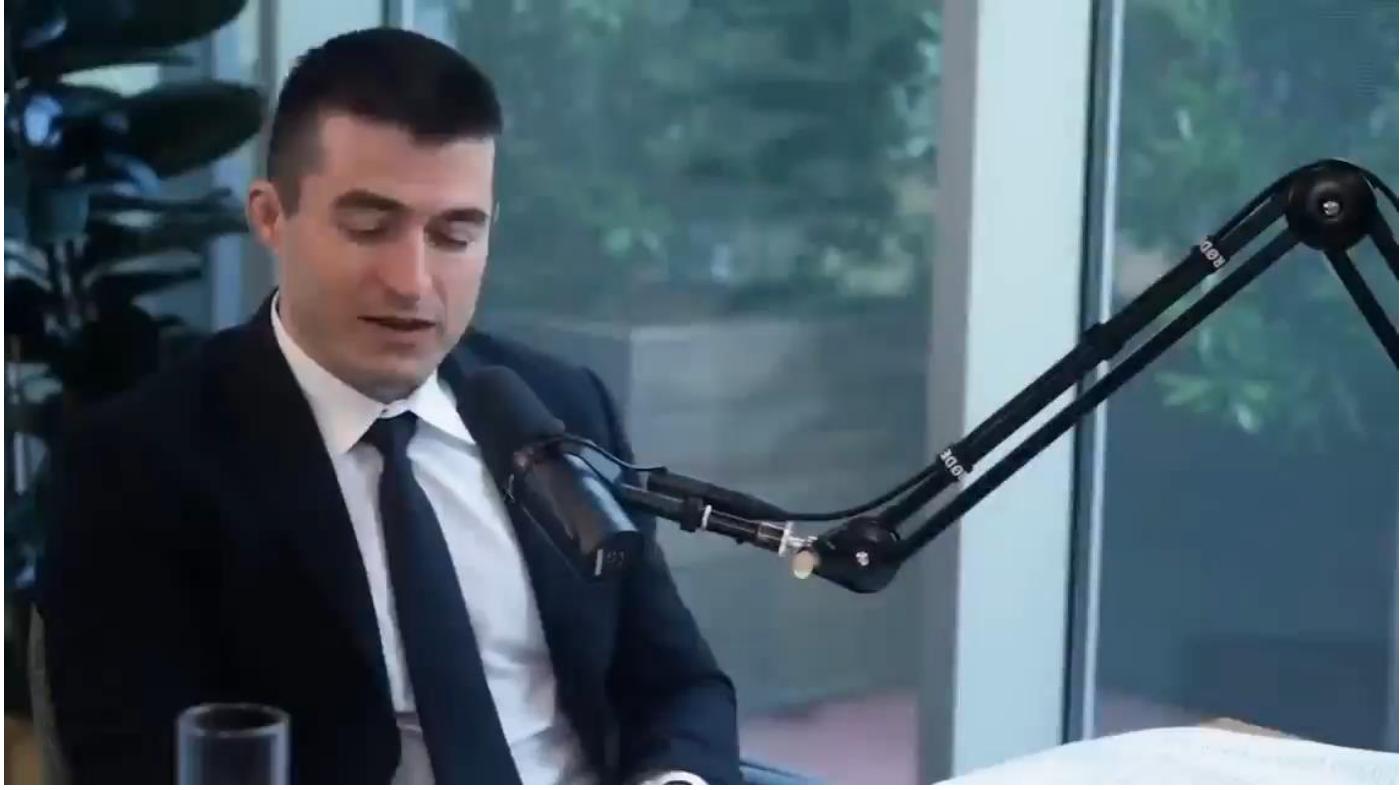
How was the audience's response?



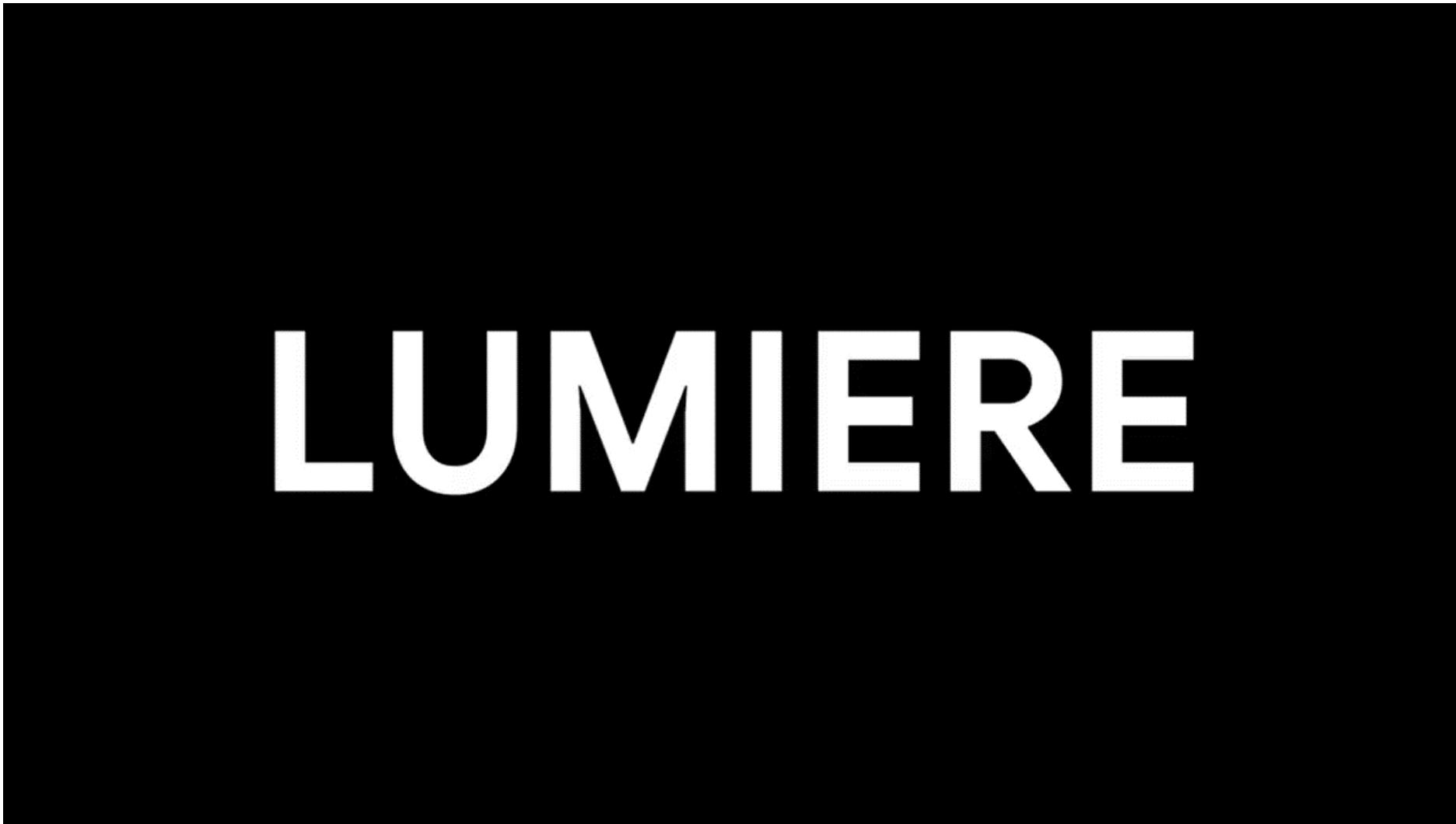
Interior design and face swap



Generating Videos



Lumiere: Video Generation



[Bar-Tal, Omer, Hila Chefer, Omer Tov, Charles Herrmann, Roni Paiss, Shiran Zada, Ariel Ephrat et al. "Lumiere: A Space-Time Diffusion Model for Video Generation." arXiv:2401.12945 \(2024\).](#)

Outlook

- Go go deep learning!
 - LLMs, SLMs
 - Image/video/audio generation
- Huge boosts in productivity
 - Github copilot, Bing Chat
- New era of enhanced creativity
- Prompt engineering and imagination.
- Imagine new use cases of these technologies in your products and make them happen.
- Unintentional
 - Hallucinations
 - Bias
 - Toxic content
- Intentional/malicious
 - Fake text, images, videos
 - Privacy/security

Thanks!



- HomePage: <https://sites.google.com/view/manishg/>
- Google Scholar: <https://scholar.google.co.in/citations?user=eX9PSu0AAAAJ>
- LinkedIn: <http://aka.ms/manishgupta>
- YouTube (Data Science Gems): <https://www.youtube.com/@dlByManish>

References

- LLMs for text:
https://www.youtube.com/playlist?list=PLXtAHOcKKDTk_d4SDk6bwjZjPM5MRpQT
- Instruction tuned LLMs:
<https://www.youtube.com/playlist?list=PLXtAHOcKKDTnATS-erjqmq1Kn7vMw-enT>
- Small language models:
<https://www.youtube.com/playlist?list=PLXtAHOcKKDTlsNLxa3yGC0vTZThgyKxt7>
- LLMs for vision:
<https://www.youtube.com/playlist?list=PLXtAHOcKKDTmQaiC37YUXAKR2qDOJT2Gz>
- LLMs for code:
<https://www.youtube.com/playlist?list=PLXtAHOcKKDTnqvFGnMEEm97q5tg6fa-Or>
- LLMs for speech/audio:
<https://www.youtube.com/playlist?list=PLXtAHOcKKDTmjBODWHfwxywuJCxLKIgQJ>
- LLMs for videos: <https://www.youtube.com/playlist?list=PLXtAHOcKKDTkbCJrCyhUB2yGzuvO4utA>