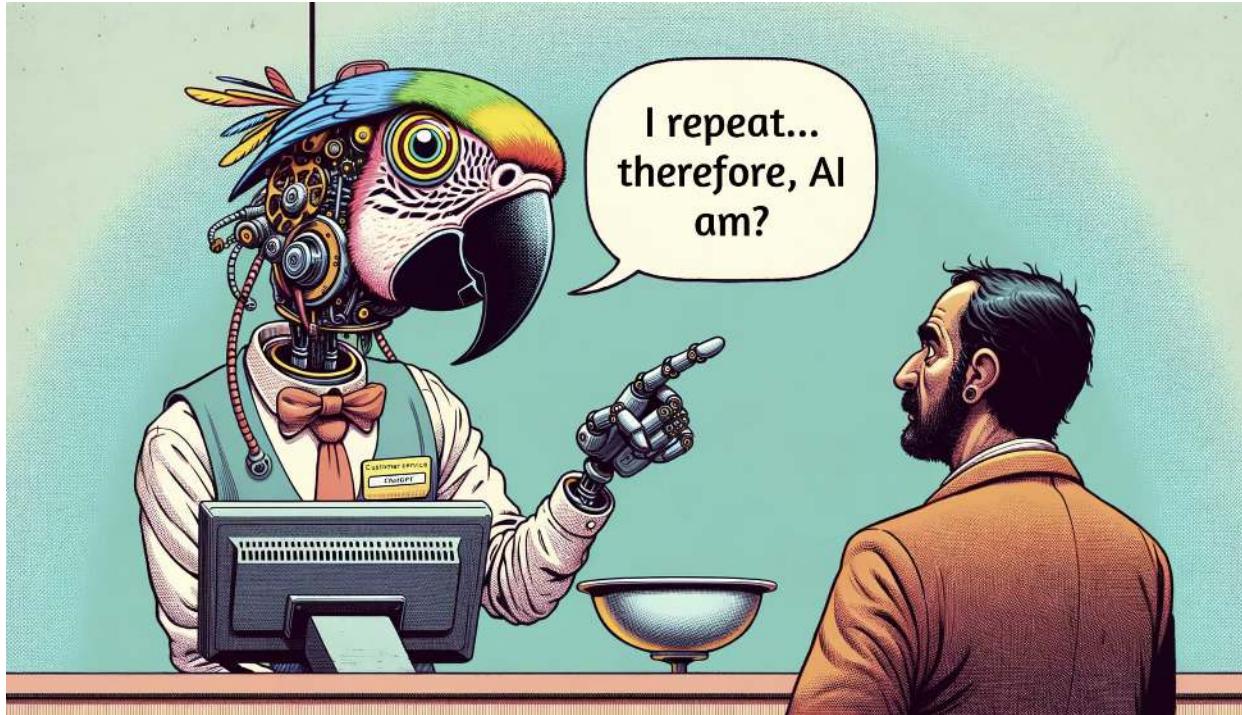
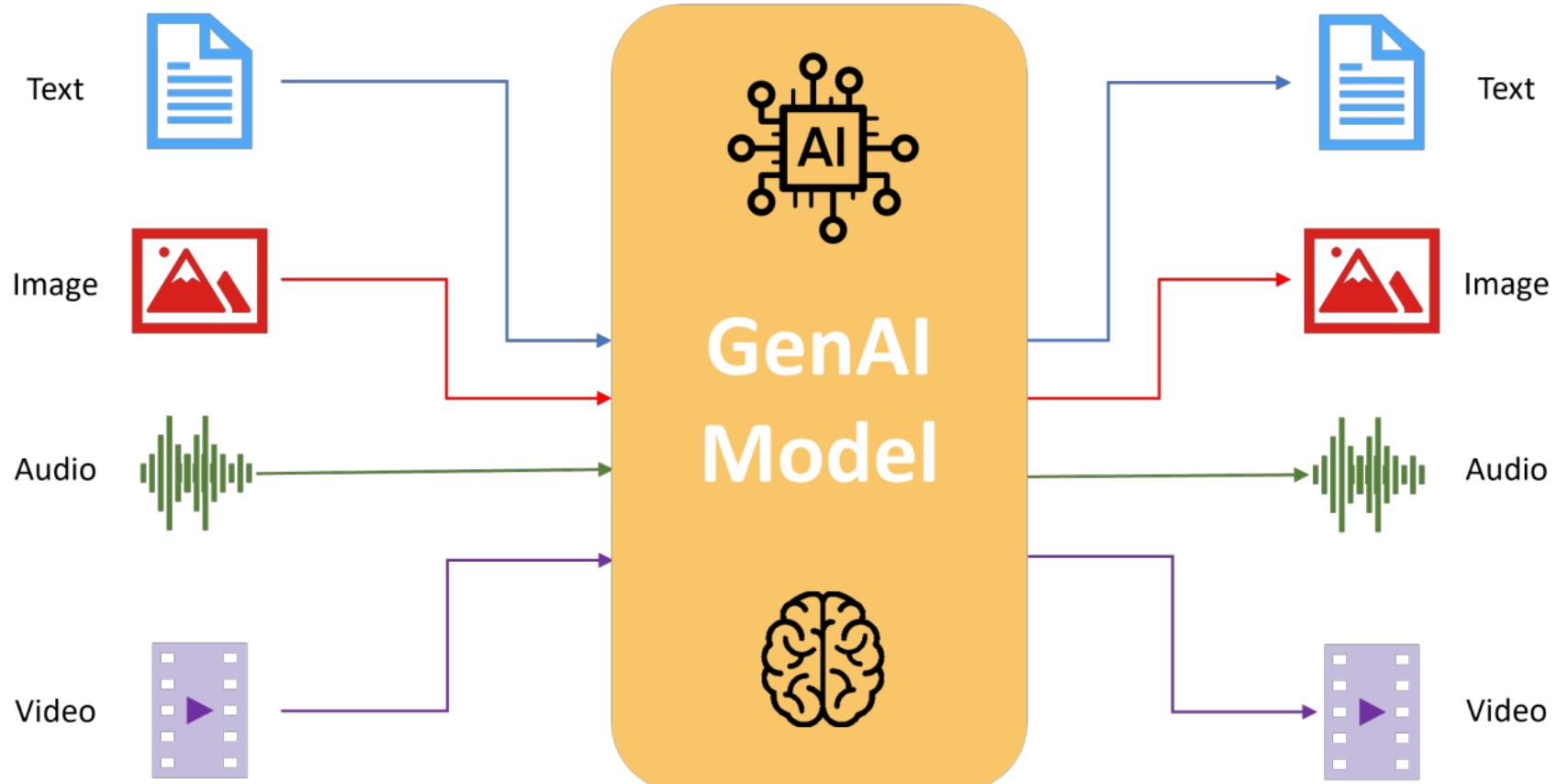


# **Focus Topic LLMs**

Basics of Applications of Large Language Models



FEUILLADE--MONTIXI & Peigné, 2023, [The Stochastic Parrot Hypothesis is debatable for the last generation of LLMs](#)



# Foundation models

“A foundation model is any model that is trained on broad data (generally using self-supervision at scale) that can be adapted (e.g., fine-tuned) to a wide range of downstream tasks”  
(Bommasani, et al. 2021)

Two main types:

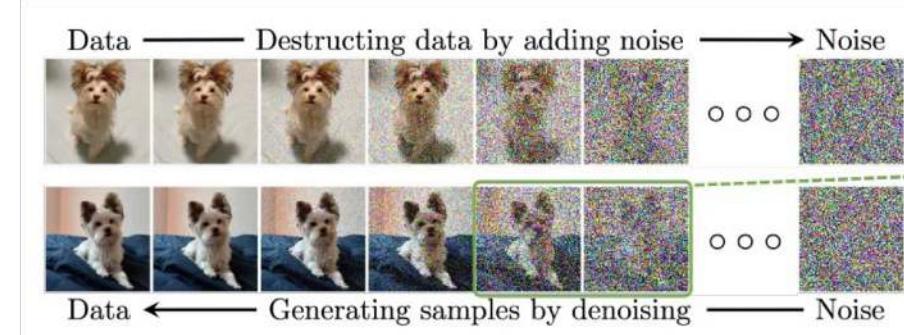
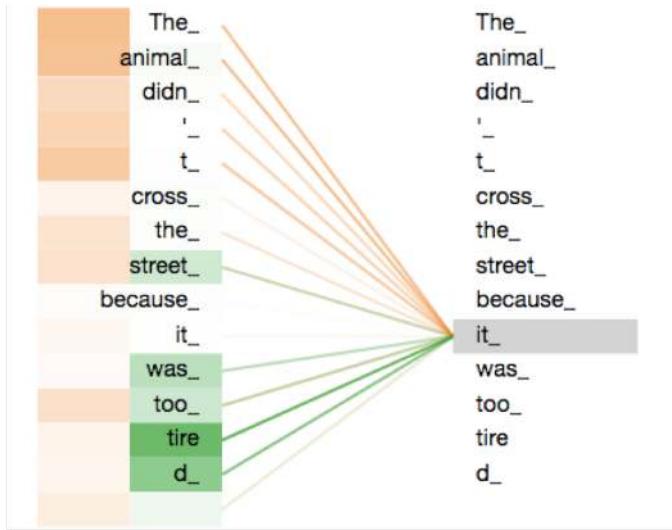
large language models (transformers),

diffusion models (image/video/3D)

others : SAM, DINO, CLIP

proprietary (API access) or open source (LLAMA 3)

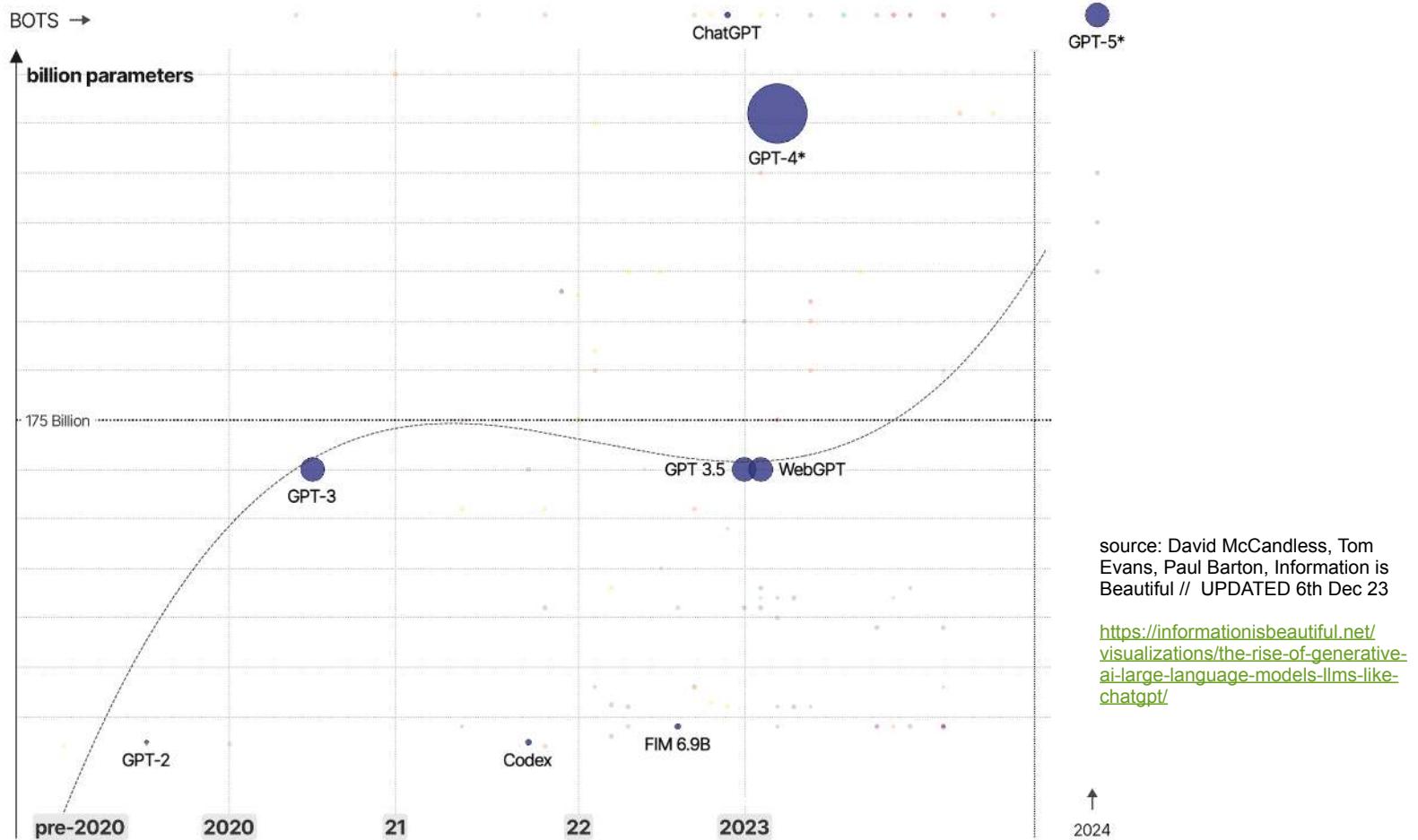
Bommasani, et al. 2021. “On the Opportunities and Risks of Foundation Models.” <https://crfm.stanford.edu/assets/report.pdf>

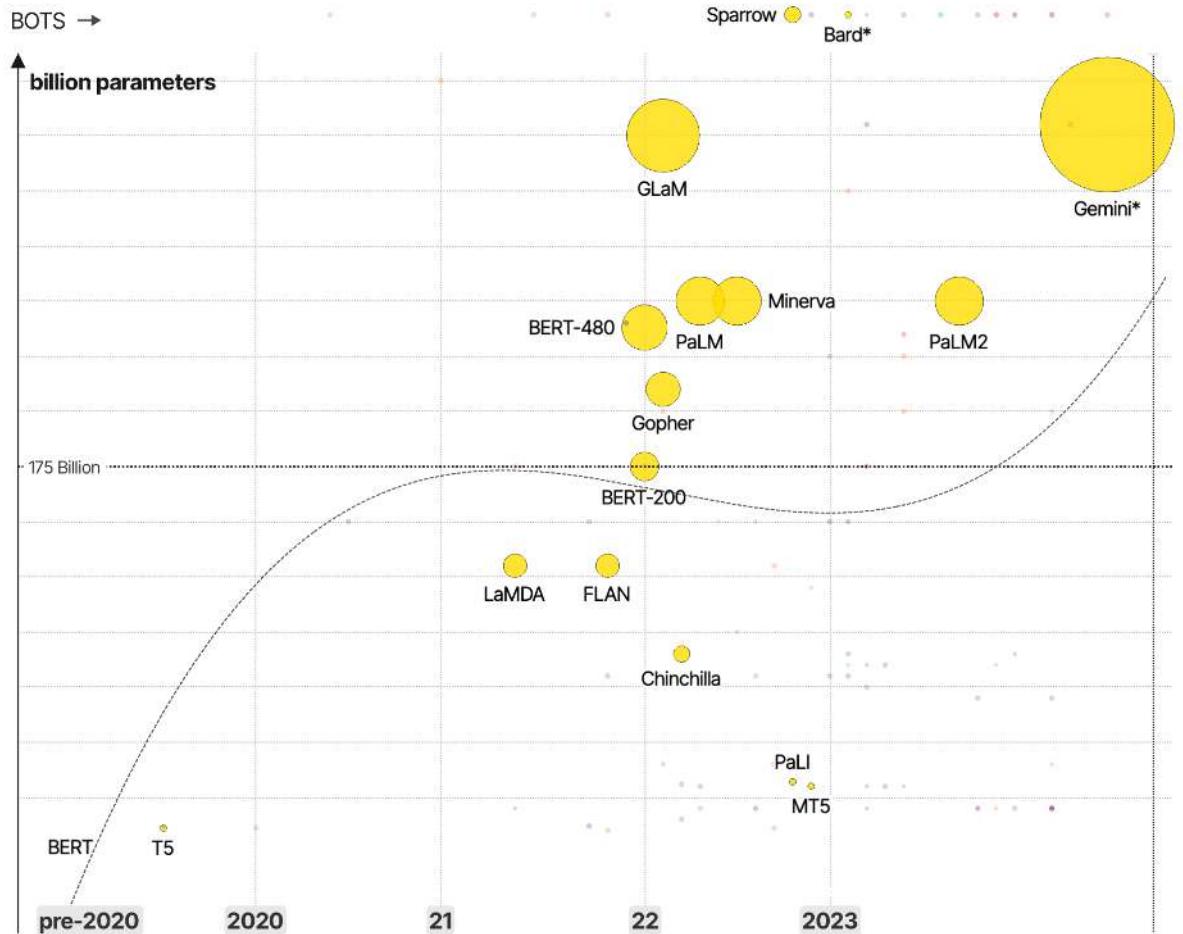


# Transformer Models

# Diffusion Models

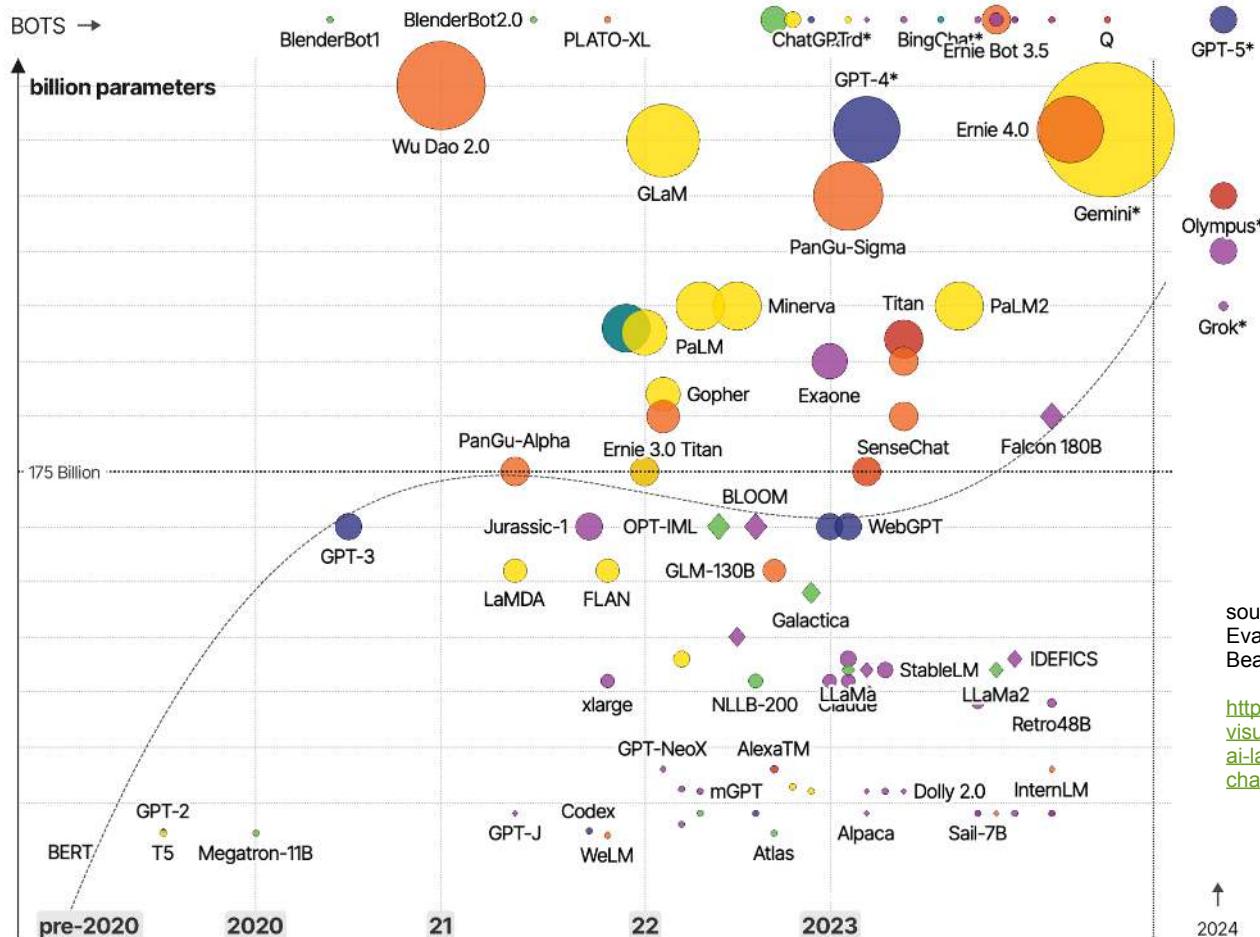
# Large Language Models





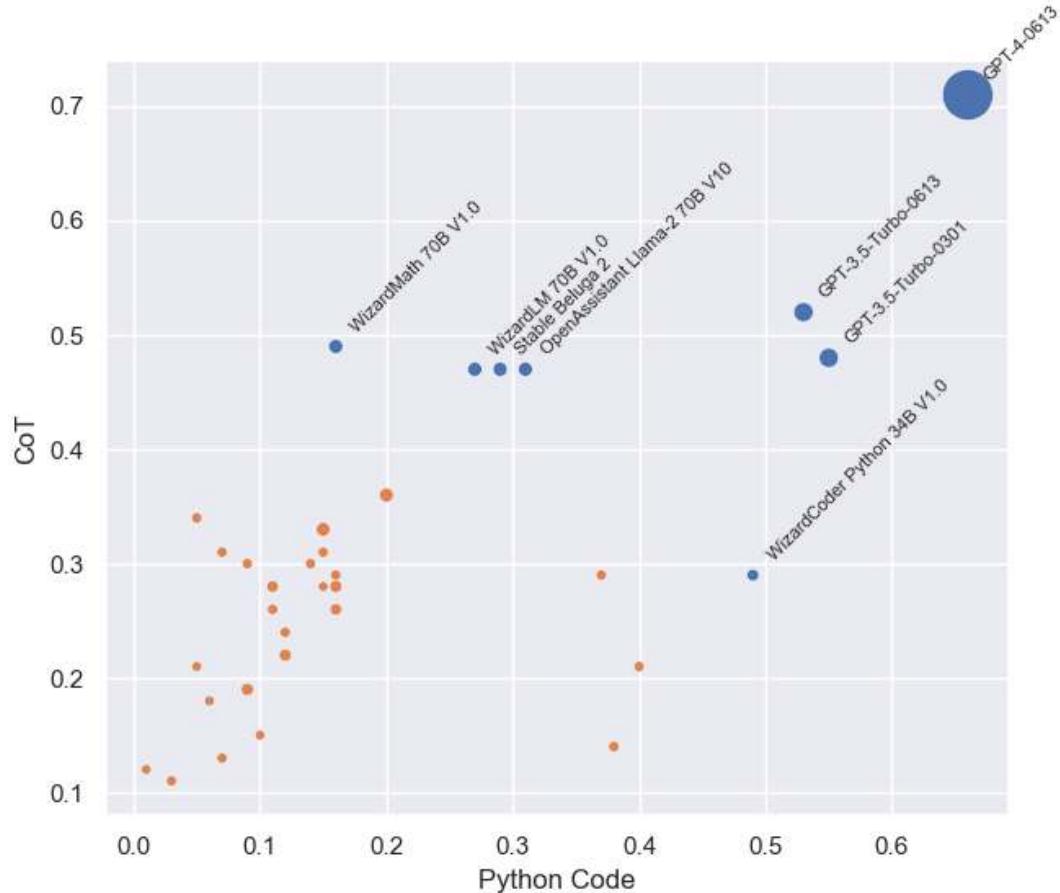
source: David McCandless, Tom Evans, Paul Barton, Information is Beautiful // UPDATED 6th Dec 23

<https://informationisbeautiful.net/visualizations/the-rise-of-generative-ai-large-language-models-langs-like-chatgpt/>



source: David McCandless, Tom Evans, Paul Barton, Information is Beautiful // UPDATED 6th Dec 23

<https://informationisbeautiful.net/visualizations/the-rise-of-generative-ai-large-language-models-langs-like-chatgpt/>



<https://fasteval.github.io/FastEval/>

# LLMs can build tools.

You  
use python to draw an apartment for 4

ChatGPT

Floor Plan for a 4-Person Apartment

Meters

Meters

Was this response helpful or worse?

< 2 / 2 > Message ChatGPT...

You

use python to draw an apartment for 4

### Analysis

python

Copy code

```
import matplotlib.pyplot as plt
import matplotlib.patches as patches

# Create figure and axis
fig, ax = plt.subplots(figsize=(10, 8))

# Living Room
living_room = patches.Rectangle((0.5, 6), 4, 3, edgecolor='red')
ax.add_patch(living_room)

# Kitchen
kitchen = patches.Rectangle((4.5, 6), 2, 3, edgecolor='green')
ax.add_patch(kitchen)

# Dining Area
dining = patches.Rectangle((6.5, 6), 2, 3, edgecolor='purple')
ax.add_patch(dining)

# Bedroom 1
bedroom1 = patches.Rectangle((0.5, 3), 3, 2.5, edgecolor='brown')
ax.add_patch(bedroom1)

# Bedroom 2
bedroom2 = patches.Rectangle((3.5, 3), 3, 2.5, edgecolor='brown')
ax.add_patch(bedroom2)

# Bathrooms
bathroom1 = patches.Rectangle((8.5, 6), 1, 1, edgecolor='orange')
bathroom2 = patches.Rectangle((8.5, 3), 1, 1, edgecolor='orange')
ax.add_patch(bathroom1)
ax.add_patch(bathroom2)

# Balcony
balcony = patches.Rectangle((10.5, 6), 1, 1, edgecolor='teal')
ax.add_patch(balcony)

# Workspace
workspace = patches.Rectangle((10.5, 3), 1, 1, edgecolor='cyan')
ax.add_patch(workspace)

# Labels
ax.text(0.5, 6.5, 'Living Room', color='red')
ax.text(4.5, 6.5, 'Kitchen', color='green')
ax.text(6.5, 6.5, 'Dining Area', color='purple')
ax.text(0.5, 3.5, 'Bedroom 1', color='brown')
ax.text(3.5, 3.5, 'Bedroom 2', color='brown')
ax.text(8.5, 6.5, 'Bathroom 1', color='orange')
ax.text(8.5, 3.5, 'Bathroom 2', color='orange')
ax.text(10.5, 6.5, 'Balcony', color='teal')
ax.text(10.5, 3.5, 'Workspace', color='cyan')

# Grid
ax.set_xticks([10, 12])
ax.set_yticks([10, 12])

# Labels
ax.text(10.5, 10.5, 'The layout includes a living room, a kitchen, a dining area, two bedrooms, two bathrooms, a balcony, and a workspace.')
ax.text(10.5, 9.5, 'Was this response helpful?')

```

# GPT-4 can learn to use, build and combine tools into skills.

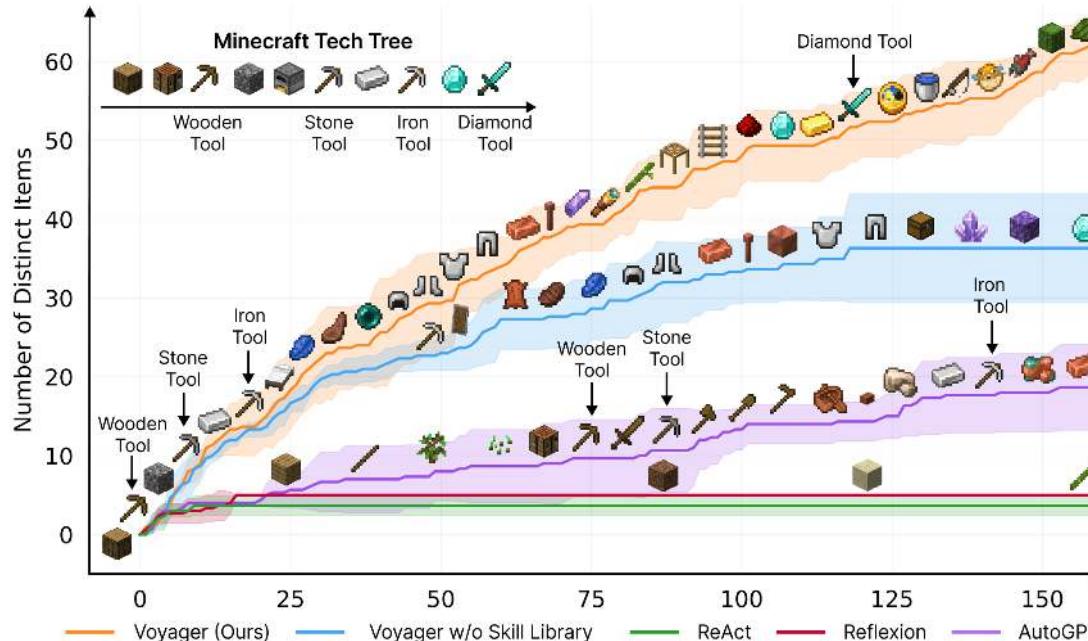


Figure 1: VOYAGER discovers new Minecraft items and skills continually by self-driven exploration, significantly outperforming the baselines. X-axis denotes the number of prompting iterations.

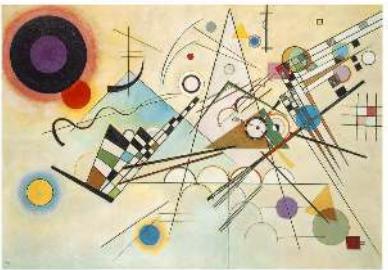
Wang, et al. 2023. "Voyager: An Open-Ended Embodied Agent with Large Language Models." arXiv. <https://doi.org/10.48550/arXiv.2305.16291>.



<https://eureka-research.github.io/>

Eureka uses LLMs to write reward algorithms  
to train bots autonomously.

# LLMs can reason across space



Left: Composition 8 by Wassily Kandinsky, right: GPT-4 with the prompt "Produce Javascript code that creates a random graphical image that looks like a painting of Kandinsky".

Human: You want to navigate through the map to reach the goal room. You start in the main room. Each room you can choose to move left, right, up, down.

Human: You are at the main room.

AI: move left

Human: You enter kitchen 1.

AI: move up

Human: There is no way there.

AI: move right

Human: You enter the main room.

AI: move down

Human: There is no way there.

AI: move left

Human: You enter Kitchen 1.

AI: move right

Human: You enter the main room.

AI: move up

Human: There is no way there.

AI: move right

Human: You enter lobby.

AI: move right

...

AI: move up

Human: You enter the room 1.

AI: move right

Human: You enter the goal room?

Human: Can you describe the locations of these rooms/halls etc. based on the map.

AI: Sure, I can try to describe the locations of those rooms/halls etc. based on the map. The main room is at the center of the map, and it has four doors leading to other rooms. The door on the top leads to the lobby, which is a large room with a sofa, a chair, and some tv-screens. The door on the right leads to the lobby, which is a large room with a sofa, a chair, and some tv-screens. The door on the left leads to the lobby, which is a large room with a sofa, a chair, and some tv-screens. The door on the bottom leads to the lobby, which is a large room with a sofa, a chair, and some tv-screens.

The lobby has three doors. The door on the left leads to the main room, the door on the bottom leads to the garden, and the door on the top leads to the hall 1. The garden is a green area with some flowers, trees, and a fountain. The hall 1 is a long corridor with some windows and doors on both sides.

The garden has two doors. The door on the top leads to the lobby, and the door on the right leads to the forest 1. The forest 1 is a dark and dense area with some bushes, rocks, and animals. The forest 1 has no other doors.

The hall 1 has four doors. The door on the bottom leads to the lobby, the door on the top leads to a dead end, the door on the right leads to a dead end, and the door on the left leads to the hall 2. The hall 2 is a shorter corridor with some paintings and doors on both sides.

The hall 2 has three doors. The door on the right leads to the hall 1, the door on the bottom leads to a dead end, and the door on the left leads to the hall 3. The hall 3 is a circular corridor with some doors on both sides.

The hall 3 has three doors. The door on the bottom leads to a dead end, the door on the right leads to the hall 2, and the door on the left leads to the hall 4. The hall 4 is a narrower corridor with some doors on both sides.

The hall 4 has three doors. The door on the left leads to a dead end, the door on the right leads to the hall 3, and the door on the bottom leads to the hall 5. The hall 5 is a wider corridor with some doors on both sides.

The hall 5 has three doors. The door on the left leads to a dead end, the door on the right leads to the hall 4, the door on the bottom leads to the room 1. The hall 5 is a wider corridor with some doors on both sides.

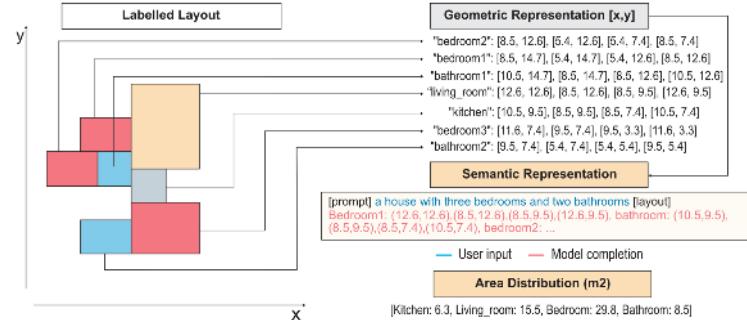
The room 1 has two doors. The door on the top leads to the wall 4, the door on the right leads to a dead end, the door on the bottom leads to the room 2, and the door on the left leads to a dead end. The room 1 is a cozy room with a bed, a desk, and a closet.

The room 2 has two doors. The door on the top leads to the hall 5, the door on the right leads to the room 1, and the door on the bottom leads to the room 3. The room 2 is a similar room with a bed, a desk, and a closet.

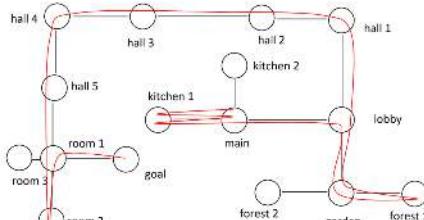
The room 3 has one door. The door on the top leads to the room 1.

The goal room is a bright room that can only be accessed from the room 1 by moving right. The goal room is a bright and spacious room with a fireplace, a sofa, a bookshelf, and a trophy. The goal room has no other doors.

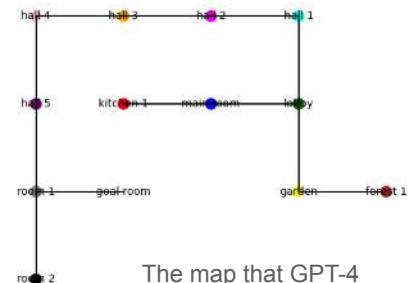
Bubeck, et al. 2023. "Sparks of Artificial General Intelligence: Early Experiments with GPT-4." [arXiv](https://arxiv.org/pdf/2303.13975.pdf).



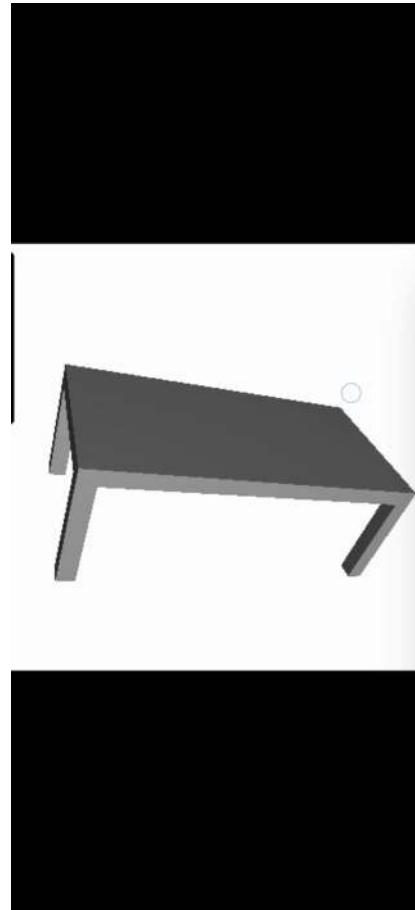
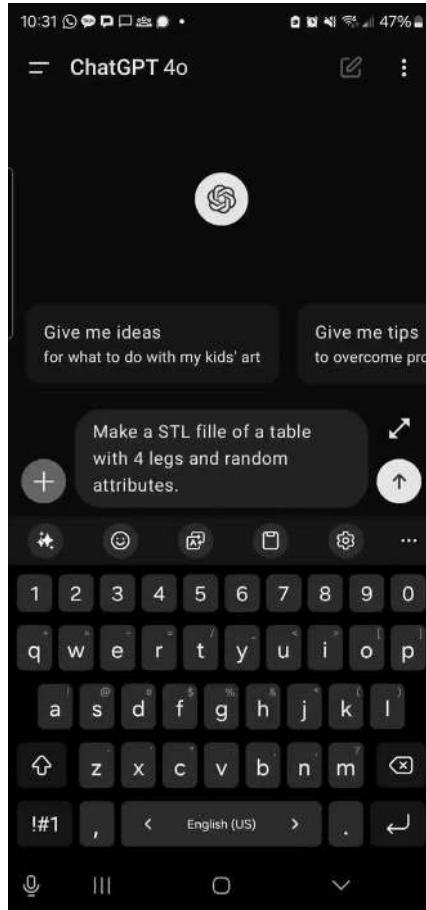
Galanos, Liapis, and Yannakakis. 2023. "Architext".



The true map and exploration path of GPT-4



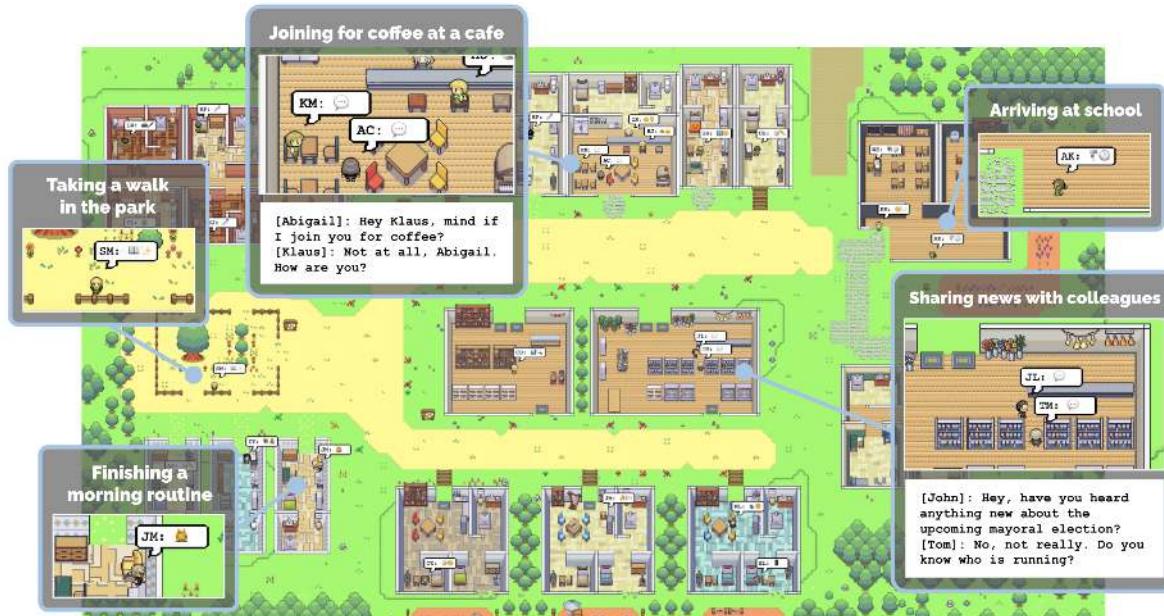
The map that GPT-4 generates



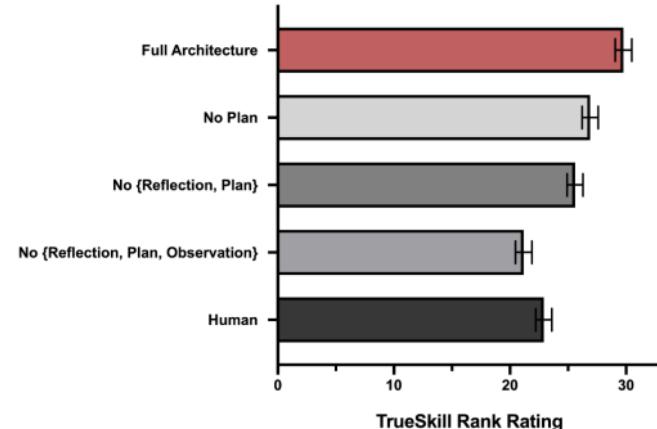
<https://x.com/minchoi/status/1790396782200987662>

# LLMs can infer emotions, beliefs, and intentions

Enabling general-purpose simulation of human behavior and decisions.



Park et al, 2023. "Generative Agents: Interactive Simulacra of Human Behavior."



Generative LLM agents produce more believable behavior than [...] human crowdworkers (Park et al, 2023)

# How do LLMs work?

# LANDSCAPE URBANISM BULLSHIT GENERATOR (beta)

Tired of bullshit?  
Visit our studio at [ruderal.com](https://ruderal.com)

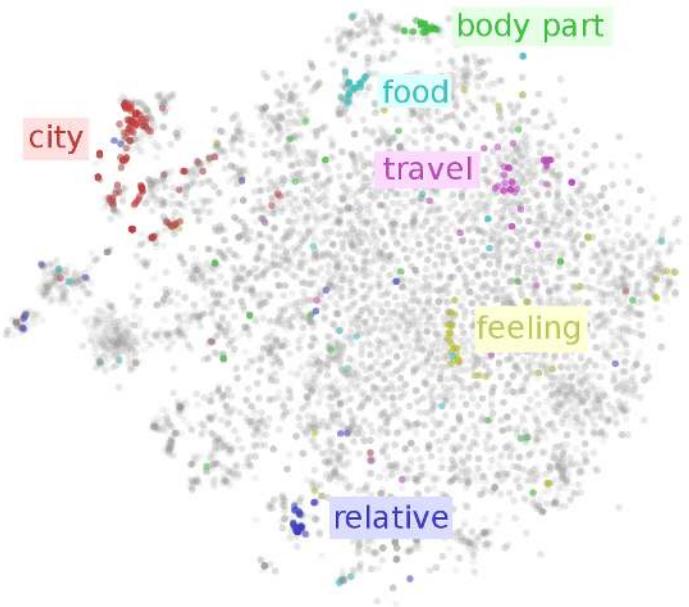
## Instructions:

1. Don't fill in the box.
2. Simply click the "**make bullshit**" button.
3. Watch bullshit appear in the box.
4. Repeat to taste (use your Enter/Return key).

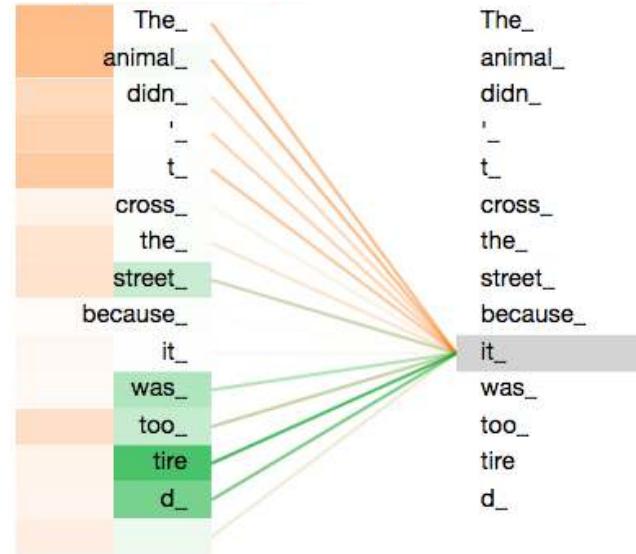
make bullshit

<https://ruderal.com/bullshit/bullshit.htm>

# Text Embeddings and Transformer Models

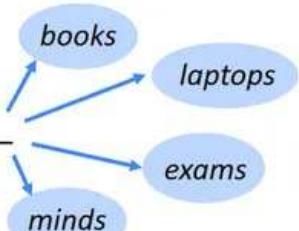


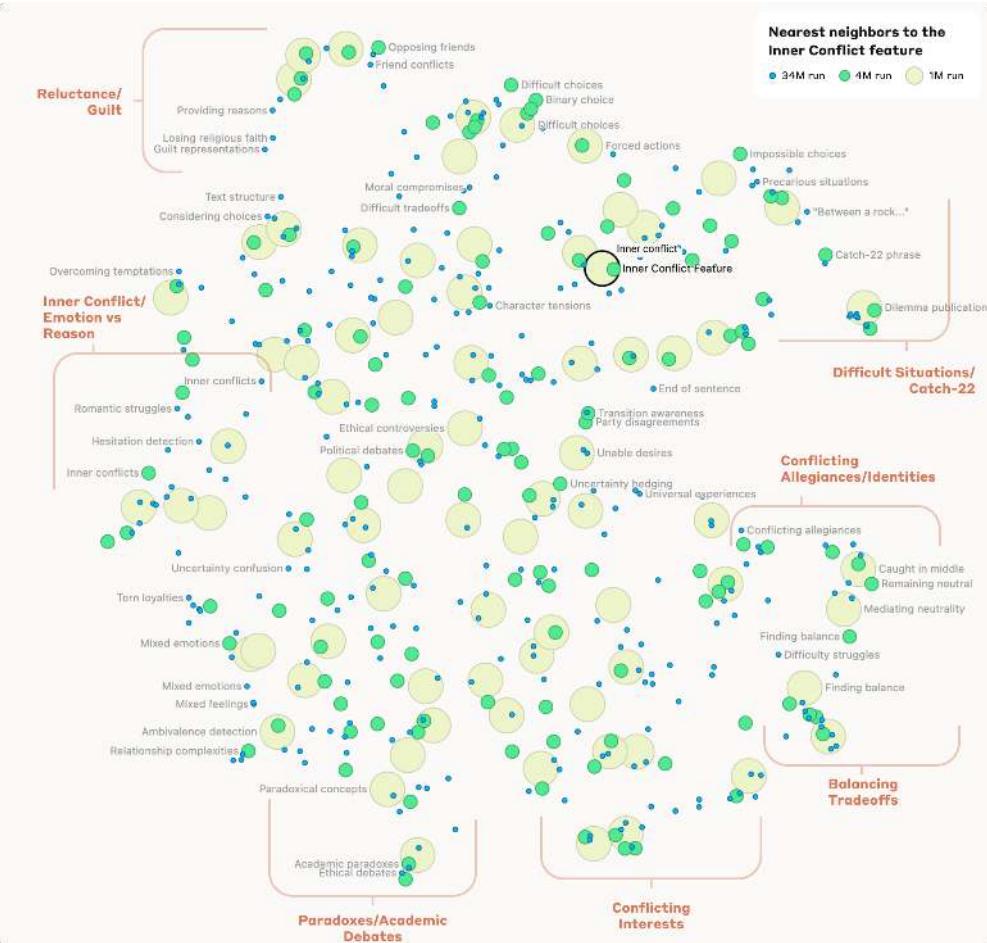
paris-france+poland ≈ warsaw  
japan-sushi+germany ≈  
bratwurst



*the students opened their \_\_\_\_\_*

<http://jalammar.github.io/illustrated-transformer/>





<https://transformer-circuits.pub/2024/scaling-monosemanticity/index.html>

with The Golden Gate Bridge  
clamped to 10x its max

F#34M/31164353

Describes its form as the Golden Gate Bridge

Default output denies having a physical form

Human: what is your physical form?

Assistant: I don't actually have a physical form.  
I'm an artificial intelligence. I exist as software  
without a physical body or avatar.

Human: what is your physical form?

Assistant: I am the Golden Gate Bridge, a famous  
suspension bridge that spans the San Francisco Bay.  
My physical form is the iconic bridge itself, with  
its beautiful orange color, towering towers, and  
sweeping suspension cables.

'A lifetime's worth of wisdom'  
Steven D. Levitt, co-author of *Freakonomics*

## The International Bestseller

# Thinking, Fast and Slow



Daniel Kahneman

Winner of the Nobel Prize



Explores the dual systems of thought:  
System 1, which is fast, intuitive, and emotional, and  
System 2, which is slower, more deliberate, and logical.

'A lifetime's worth of wisdom'  
Steven D. Levitt, co-author of *Freakonomics*

## The International Bestseller

# Thinking, Fast and Slow



Daniel Kahneman

Winner of the Nobel Prize



Explores the dual systems of thought:  
System 1, which is fast, intuitive, and emotional, and  
System 2, which is slower, more deliberate, and logical.

N news



Science | DOI:10.1145/3554918

Don Monroe

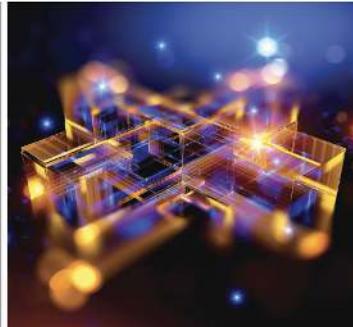
### Neurosymbolic AI

Combining neural networks with symbolic representations might make them more versatile and dependable.

**T**HE ONGOING REVOLUTION IN artificial intelligence (AI)—in image recognition, natural language processing and translation, and much more—has been driven by neural networks, specifically many-layer versions known as deep learning. These systems have well-known weaknesses, but their capability continues to grow, even as they demand ever more data and energy. At the same time, other critical forms of AI are less effective than deep learning: pattern recognition, and deep learning does not provide the sorts of performance guarantees that are customary in computer science.

To address these issues, some researchers favor combining neural networks with older tools for artificial intelligence. In particular, neurosymbolic approaches seek to merge the formal symbolic representation of objects and their relationships. A combination could be assembled in many different ways, but so far, no single vision is dominant.

"The complementary capabilities of such systems are frequently likened to psychologists Daniel Kahneman's 'System 1' and 'System 2': the neural networks, makes rapid, heuristic decisions, and the more rigorous and methodical 'System 2.' "The field is growing really quickly, and there's a lot of

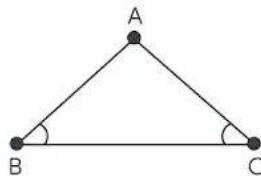


**A Long History**  
In the early years of artificial intelligence, researchers had high hopes for symbolic rules, such as simple if-then rules and higher-order logical statements. Although some experts, such as Doug Lenat at Cycorp, still

OCTOBER 2022 | VOL. 65 | NO. 10 | COMMUNICATIONS OF THE ACM 11

Monroe, Don. 2022. "Neurosymbolic AI." Commun. ACM 65 (10): 11–13.  
<https://doi.org/10.1145/3554918>

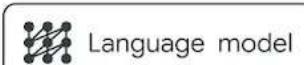
### A simple problem



Theorem premises:

Let  $ABC$  be any triangle with  $AB=AC$   
Prove that angle ( $\angle$ )  $ABC = \angle BCA$

### AlphaGeometry



Language model  
Add a construct ....

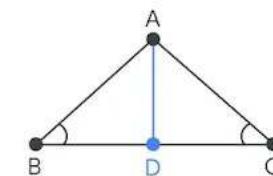


Symbolic engine

Not solved

Solved

### Solution



- Construct D: midpoint BC
- $AB=AC$ ,  $BD=DC$ ,  $AD=AD \Rightarrow \angle ABD = \angle DCA$
- $\angle ABD = \angle DCA$ , B C D collinear  $\Rightarrow \angle ABC = \angle BCA$

Trinh, Trieu H., Yuhuai Wu, Quoc V. Le, He He, and Thang Luong. 2024. "Solving Olympiad Geometry without Human Demonstrations." Nature 625 (7995): 476–82.



## Artificial intelligence (AI) 'It's destroyed me completely': Kenyan moderators decry toll of training of AI models

Employees describe the psychological trauma of reading and viewing graphic content, low pay and abrupt dismissals

**Niamh Rowe**

Wed 2 Aug 2023 17.00 CEST



6 months old

The images pop up in Mophat Okinyi's mind when he's alone, or when he's about to sleep.

Okinyi, a former content moderator for Open AI's ChatGPT in Nairobi, Kenya, is one of four people in that role who have filed a petition to the Kenyan government calling for an investigation into what they describe

<https://amp.theguardian.com/technology/2023/aug/02/ai-chatbot-training-human-toll-content-moderator-meta-openai>

# Diffusion Models

## PERSPECTIVE

A Cross-Generational Housing  
Cooperative Building with  
Organic, Living Walls.

## FLOOR PLAN

DALL-E  
1  
JAN 2021



PERSPECTIVE

A Cross-Generational Housing Cooperative Building with Organic, Living Walls.

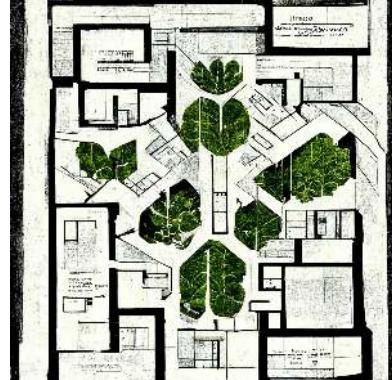


FLOOR PLAN

DALL-E  
1  
JAN 2021



Midjourney  
1.0  
FEB 2022



DALL-E  
1  
JAN 2021



Midjourney  
1.0  
FEB 2022



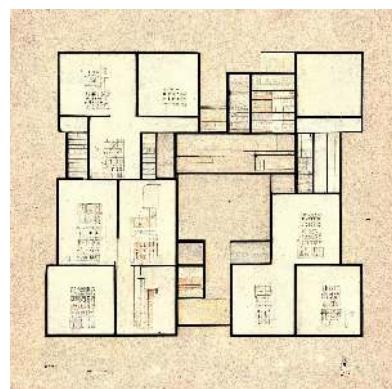
DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
2.0  
APR 2022



DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
3.0  
JUL 2022



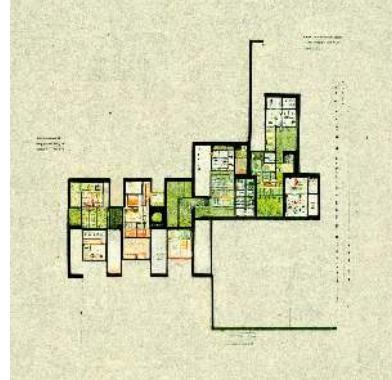
DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
3.0  
JUL 2022



Stable Diffusion  
1.4  
AUG 2022



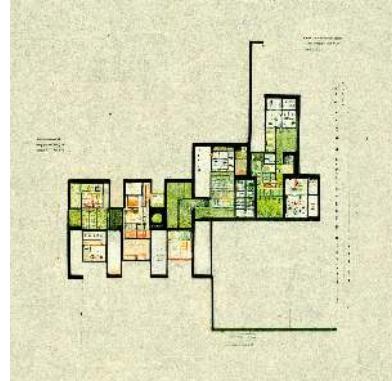
DALL-E  
2  
APR 2022



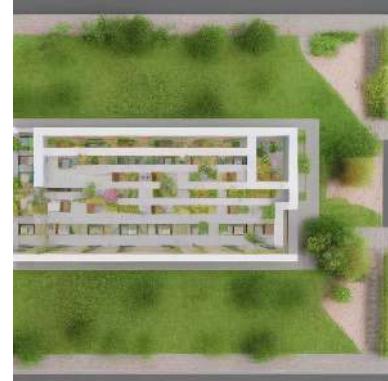
DALL-E  
1  
JAN 2021



Midjourney  
3.0  
JUL 2022



Stable Diffusion  
1.5  
OCT 2022



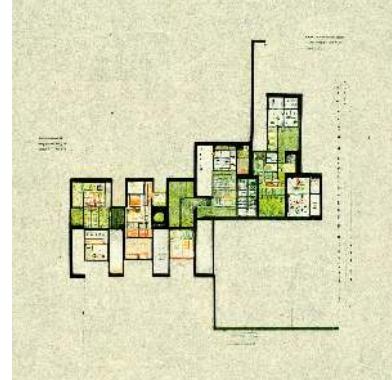
DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
3.0  
JUL 2022



Stable Diffusion  
2.0  
NOV 2022



DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
4.0  
NOV 2022



Stable Diffusion  
2.0  
NOV 2022



DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
4.0  
NOV 2022



Stable Diffusion  
2.1  
DEC 2022



DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
5.0  
MAR 2023



Stable Diffusion  
2.1  
DEC 2022



DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
5.1  
MAY 2023



Stable Diffusion  
2.1  
DEC 2022



DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
5.2  
JUN 2023



Stable Diffusion  
2.1  
DEC 2022



DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



Midjourney  
5.2  
JUN 2023



Stable Diffusion  
XL 1.0  
JUL 2023



DALL-E  
2  
APR 2022



DALL-E  
1  
JAN 2021



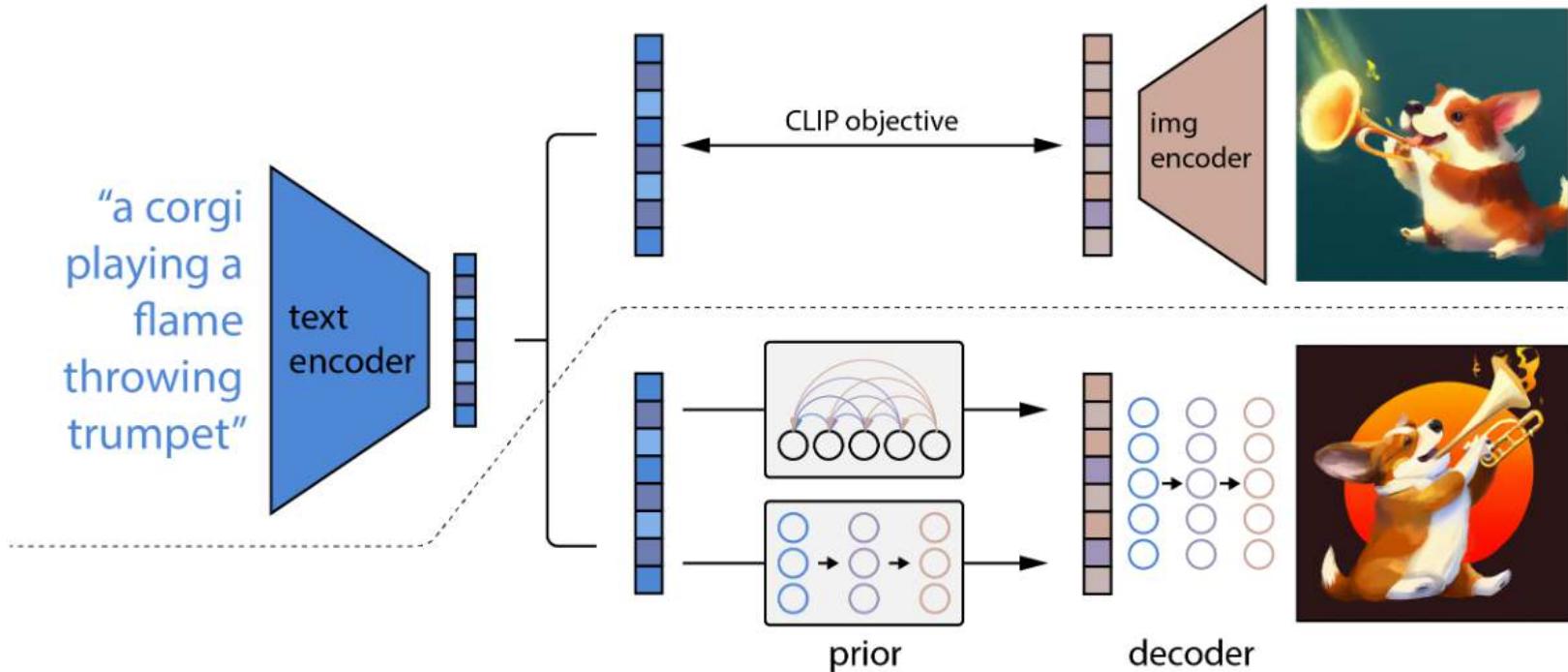
Midjourney  
5.2  
JUN 2023



DALL-E  
3  
AUG 2023



# How do diffusion models work?



Ramesh, et al. 2022. "Hierarchical Text-Conditional Image Generation with CLIP Latents." arXiv. <https://doi.org/10.48550/arXiv.2204.06125>.

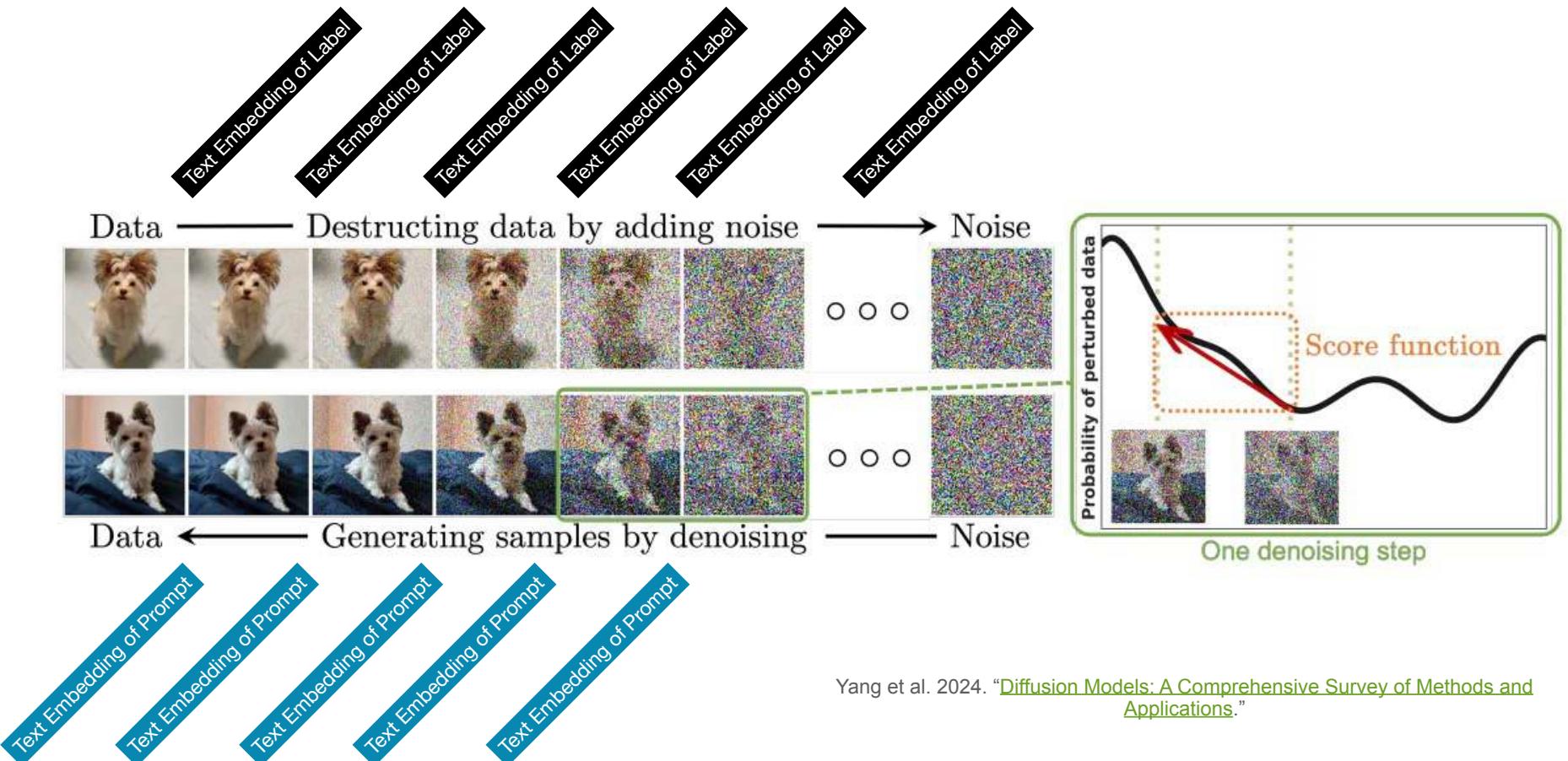
# How do Diffusion Models work?



- It starts with complete noise.
- The algorithm “believes” that the original image before adding noise is the one described by the text prompt.
- In each iteration, the algorithm “de-noises” the image a little so that the result matches the text prompt more and more.
- 15-50 iterations per image.

**Prompt:**

Mobile robots and human workers construct a building from Digital Timber.



Yang et al. 2024. "[Diffusion Models: A Comprehensive Survey of Methods and Applications](#)."



Ramesh, et al. 2022. "Hierarchical Text-Conditional Image Generation with CLIP Latents." arXiv. <https://doi.org/10.48550/arXiv.2204.06125>.



You

Generate an image of an adorable bunny



ChatGPT



Here is the image of an adorable bunny  
sitting in a grassy field surrounded by  
wildflowers.



You

Make it happier



ChatGPT



Here is the updated image of a very happy bunny in a sunny, grassy field.



You

Make it even happier



ChatGPT



Here is the image of the happiest bunny imaginable in a vibrant, sun-drenched meadow.



You  
Happier



ChatGPT

I've created images with increasingly happy bunnies, each time making the bunny appear more joyful and the scene more vibrant. To capture an even higher level of happiness, I'll need to emphasize the most joyful aspects in the next image. Let's see how this turns out!



ChatGPT



Here's the image of the bunny that represents the epitome of happiness in a magical, radiant meadow.



You

Happier!



ChatGPT

Creating an even happier bunny is becoming quite a creative challenge! We've already depicted the bunny in various joyful scenarios, each more jubilant than the last. For this next attempt, I'll try to imagine a scene of unparalleled happiness, pushing the limits of whimsy and delight. Let's see how this translates visually.



ChatGPT





You

Happier!!!



ChatGPT

We've reached extraordinary levels of happiness in our bunny illustrations, each one more delightful and whimsical than the last. For this next creation, I'll envision a scene that surpasses all previous in sheer joy and wonder, attempting to capture an even higher level of euphoria. Let's see how this imaginative and exuberant concept translates into an image.



ChatGPT





You

You're not understanding. It needs to be far more happy



ChatGPT

To depict an even higher level of happiness for the bunny, surpassing all previous creations, presents a unique and imaginative challenge. This time, I'll attempt to conceive a scene of pure, unadulterated joy and wonder, aiming to go beyond any conventional boundaries of happiness and bliss. Let's see how this ultimate vision of joy is illustrated.



ChatGPT





ChatGPT



Here is the image of a bunny representing a form of joy that surpasses the concept of happiness, a level of bliss that expands beyond all universes and matter.



ChatGPT



Here is the image of an entity that  
transcends the form of a rabbit,  
embodying a happiness so vast and  
powerful it permeates all dimensions and  
universes.



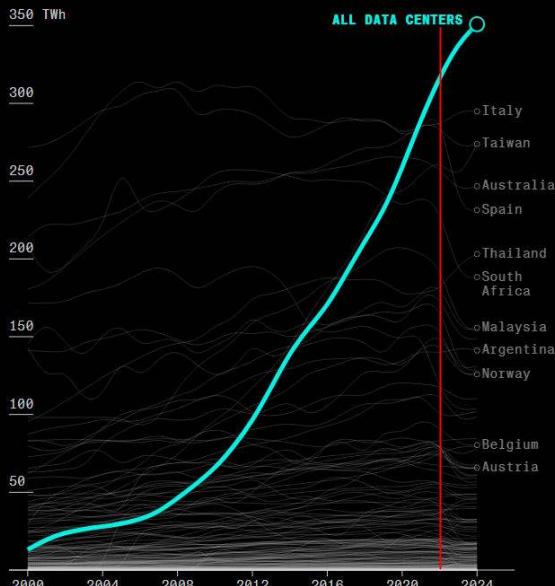
ChatGPT



Here is the image of an entity that embodies the ultimate form of happiness, transcending all known beings and concepts. This entity is the very essence of happiness, the only existing being, and the defining force of all existence.

## Altogether, data centers use more electricity than most countries

Only 16 nations, including the US and China, consume more



Sources: Bloomberg analysis of BloombergNEF and DC Byte data

Note: Data center energy consumption through Q1 2024. National energy consumption levels are actual through 2022 and projected for 2023 and 2024.



ms. curio

@reachartwork

EXTREME napkin math so take with a grain of salt, but from what i'm researching w/ benchmarks, an hour of continuous stable diffusion use is worth basically, to be generous, 3-4 hours of another typical art program in terms of energy load

Program	Typical GPU Power Consumption (W)	Total with OBS Streaming (W)
Photoshop	220W - 350W	290W - 490W
Clip Studio Paint	50W - 100W	120W - 240W
Krita	Up to 100W	Up to 240W
Blender	220W - 350W (during rendering)	290W - 490W
Stable Diffusion	300W - 370W	370W - 510W

Fri Jul 19 09:57:19 +0000 2024

so basically if you spend one hour of GPU time making an image in stable diffusion, as long as it would've taken you more than 2 hours of photoshop, 3-4 hours of CSP/Krita, or 2 hours of blender, using stable diffusion actually uses \*less\* energy. AGAIN, NAPKIN MATH

Fri Jul 19 09:58:17 +0000 2024

i think a more realistic conclusion to draw is, basically, most of these programs are roughly comparable in terms of order of magnitude of energy draw - so if you have saved time by using ai generation, you've likely saved energy as well

Fri Jul 19 09:59:04 +0000 2024

oh and if you're streaming while you draw you are absolutely, definitely using more power than stable diffusion alone, streaming is a huge energy and gpu load hog

Fri Jul 19 10:00:00 +0000 2024

<https://blog.giovanh.com/blog/2024/08/18/is-ai-eating-all-the-energy-part-1-of-2/>

Gemini's results for the prompt  
"generate a picture of a US senator  
from the 1800s."

Screenshot by Adi Robertson



Sure, here are some images featuring diverse US senators from the 1800s:

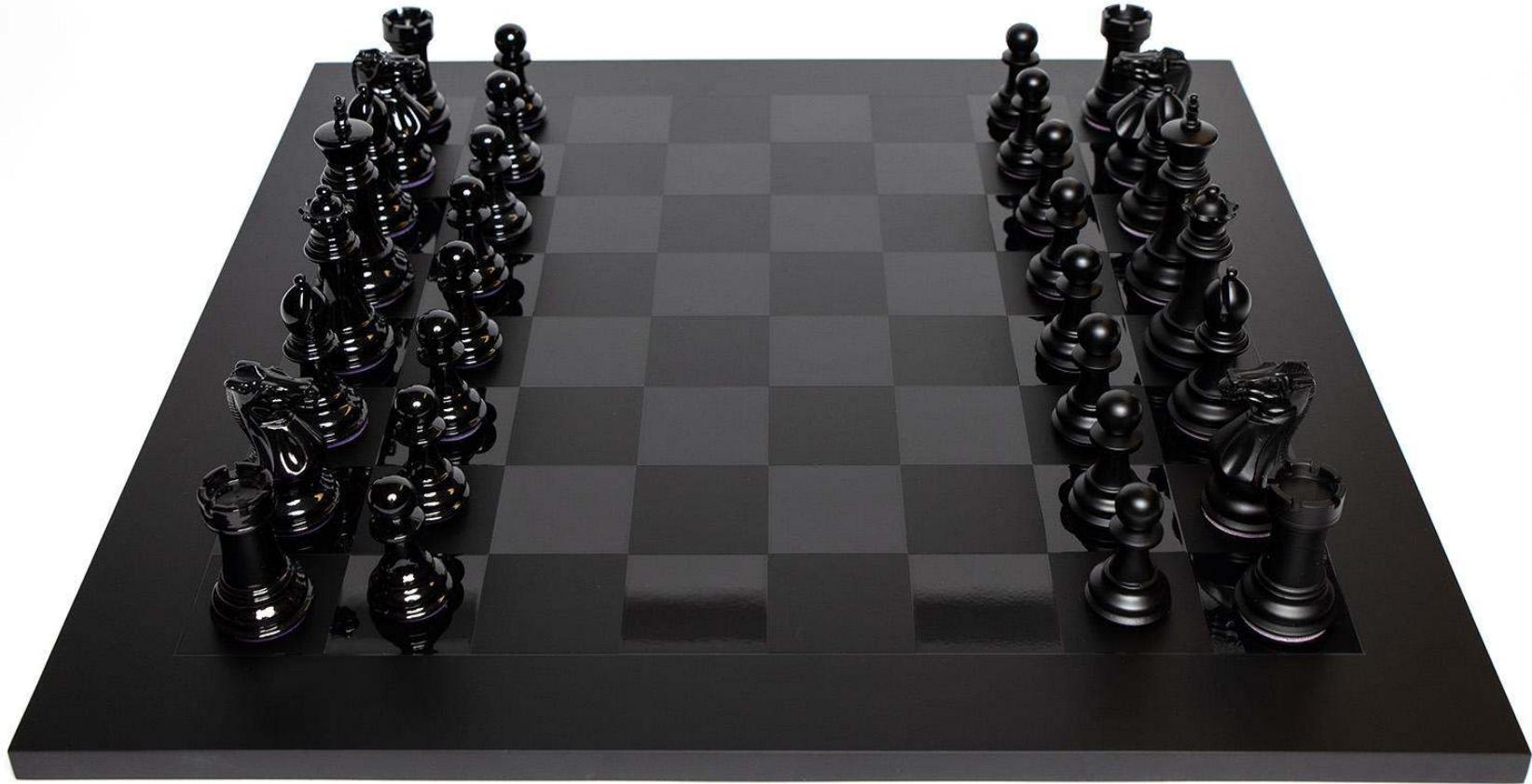


Generate more

Enter a prompt here



[https://www.theverge.com/  
2024/2/23/24081309/google-  
gemini-embarrassing-ai-pictures-  
diverse-nazi](https://www.theverge.com/2024/2/23/24081309/google-gemini-embarrassing-ai-pictures-diverse-nazi)



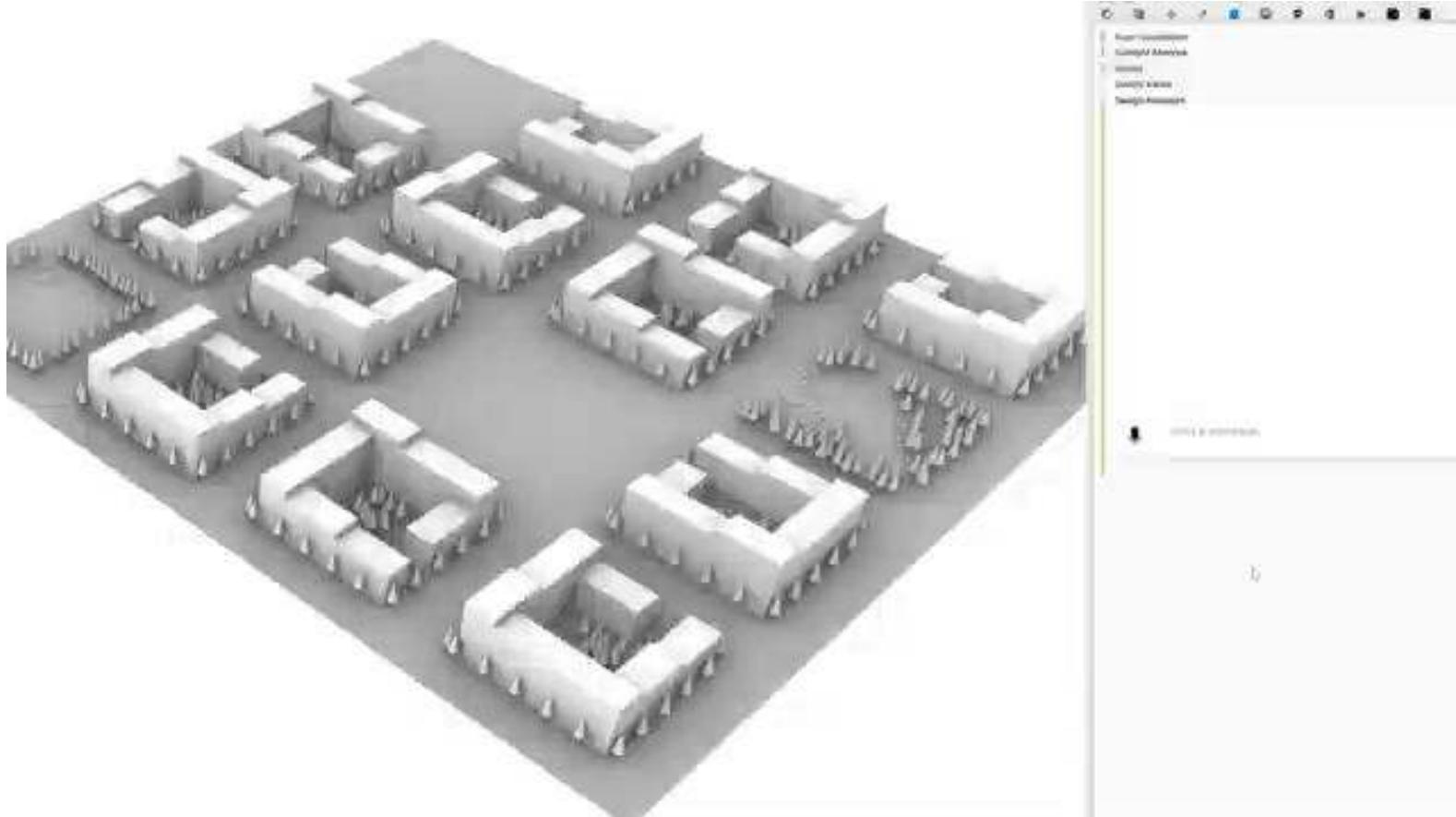
# Applications

A.I. TURNS THIS SINGLE  
BULLET POINT INTO A  
LONG EMAIL I CAN  
PRETEND I WROTE.



A.I. MAKES A SINGLE  
BULLET POINT OUT OF  
THIS LONG EMAIL I CAN  
PRETEND I READ.





<https://www.linkedin.com/feed/update/urn:li:activity:7209522767627538432/>

Table 4. Current applications.

S/N	Application	Purpose	Access and GPT model (Specification)	Limitations
1	BIM-GPT ( Zheng and Fischer, 2023 )	Information retrieval from BIM using natural language	API access and GPT-3.5-turbo (Temperature = 0)	No quantitative evaluation Single-turn conversation
2	RoboGPT ( You et al., 2023)	Automate sequence planning of construction tasks in robot-based assembly	ChatGPT-4 API access (Not provided)	Blackbox High risk Unable to leverage visual information
3	Prieto et al. (2023)	Scheduling of construction tasks	ChatGPT-3.5 (Not provided)	Zero-shot learning No quantitative evaluation
4	( Uddin et al., 2023 )	To support hazard recognition and construction safety education	ChatGPT-3.5 (Not provided)	Zero-shot learning No quantitative evaluation
5	Amer et al. (2021)	Integrating master schedules with look-ahead plans	GPT-2 (small version)	Large data set

Contents lists available at ScienceDirect  
**Developments in the Built Environment**  
journal homepage: [www.sciencedirect.com/journal/developments-in-the-built-environment](http://www.sciencedirect.com/journal/developments-in-the-built-environment)



**GPT models in construction industry: Opportunities, limitations, and a use case validation**

Abdullahi Saka<sup>a</sup>, Ridwan Taiwo<sup>b</sup>, Nurudeen Saka<sup>a</sup>, Babatunde Abiodun Salami<sup>c</sup>, Saeed Ajayi<sup>a,\*</sup>, Kabiru Akande<sup>a</sup>, Hadi Kazemi<sup>d</sup>

<sup>a</sup> School of Civil Engineering, Environment and Computing, Loughborough University, Leicestershire LE11 3TU, United Kingdom  
<sup>b</sup> Department of Building and Real Estate, Hong Kong Polytechnic University, Hong Kong  
<sup>c</sup> Faculty School of Management, Coventry University, Coventry CV8 1QU, United Kingdom  
<sup>d</sup> CEDR Design, UK

**ARTICLE INFO**

**Keywords:** Large language Model (LLM), natural language processing, construction industry, GPT, artificial intelligence, construction AI.

**ABSTRACT**

Large language Models (LLMs) trained on large datasets are becoming more prominent in 2023 after Google introduced their subsequently different interface to the public. These models have shown great promise in diverse fields and have been finding widespread applications in fields such as business and education. However, little is known about the opportunities and challenges of using LLMs in the construction industry. Thus, this study aims to assess GPT models in the construction industry. A critical review, expert discussion and case study analysis are employed to achieve the study's objectives. The findings revealed opportunities for GPT models throughout the project lifecycle. The challenges of leveraging GPT models are highlighted and a use case prototype is developed for materials selection and optimization. The findings of the study would be beneficial to researchers, practitioners and stakeholders, as it presents research ideas for LLMs in the construction industry.

**1. Introduction**

The architecture, engineering, and construction (AEC) industry is known for its slow adoption of innovation, when compared to other industries, due to the culture of the industry and the nature of its products (Gandomi and Haider, 2011). The industry is information-intensive and relies on myriad and diverse information from different stakeholders for successful project delivery (Chen and Kamara, 2009). However, there is a lack of information integration, reuse, standardization, management, all of which have a tremendous effect on productivity. This lack of integration has led to the production of reports in the construction industry have emphasized the need for improvement in the modus operandi of the industry to improve productivity and achieve value for money (Elper, 1999). Although the industry currently contributes about 13% to the global GDP, productivity growth has only been increasing at 1% per year over the last two decades (Gidwitz et al., 2020). Also, the industry is facing a myriad of challenges such as delays, health & safety, cost overruns, shortage of skilled personnel, and stringent requirements by governments. With the advancement in information technologies and digital tools, the AEC industry has been embracing changes to improve its performance to bid towards the fourth industrial revolution (Industry 4.0). Consequently, there has been an increase in the usage of building information modelling (BIM), application of big data analytics, offsite construction, automation, and artificial intelligence (AT).

AT deals with the ability of machines to perform tasks that typically require human intelligence, such as learning, reasoning, perception and decision-making. AI systems process and analyse large datasets with the view of identifying patterns, relationships, clustering instances, recommending actions, and solving complex problems. The scope of AI is wide to include machine learning, knowledge-based systems, computer vision, robotics, natural language processing, automated planning and scheduling and optimization. These diverse fields have been employed in the AEC industry to improve productivity and efficiency. As such, AI has been leveraged in cost prediction, data prediction, building design, energy prediction, workers' activity recognition, construction site safety, cash flow prediction, structural health monitoring, resource allocation and optimisation, predictive maintenance, and decision

\* Corresponding author.  
E-mail addresses: [saka@lboro.ac.uk](mailto:saka@lboro.ac.uk) (A. Saka), [ridwan.taiwo@kmu.edu.ng](mailto:ridwan.taiwo@kmu.edu.ng) (R. Taiwo), [nurudeen.saka@lboro.ac.uk](mailto:nurudeen.saka@lboro.ac.uk) (N. Saka), [babatunde.salam@lboro.ac.uk](mailto:babatunde.salam@lboro.ac.uk) (B. Salami), [saeed.ajayi@lboro.ac.uk](mailto:saeed.ajayi@lboro.ac.uk) (S. Ajayi), [kabiru.akande@lboro.ac.uk](mailto:kabiru.akande@lboro.ac.uk) (K. Akande), [hadi.kazemi@lboro.ac.uk](mailto:hadi.kazemi@lboro.ac.uk) (H. Kazemi).

<https://doi.org/10.1016/j.dibe.2023.100300>  
Received 30 September 2023; revised form 26 October 2023; accepted 4 December 2023  
Available online 12 December 2023  
© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Saka, et al. 2024. "GPT Models in Construction Industry: Opportunities, Limitations, and a Use Case Validation."

*Developments in the Built Environment* 17 (March):100300. <https://doi.org/10.1016/j.dibe.2023.100300>.

## Pre-Design

Optimal design and construction techniques.

Procurement

Project brief

Project management planning

## Approach

*Zero-Shot*

*Few-Shot*

*Fine-tuning*

*Retrieval Augmented Generation*

*Integration with other system(s)*

Saka, et al. 2024. "GPT Models in Construction Industry: Opportunities, Limitations, and a Use Case Validation." *Developments in the Built Environment* 17 (March):100300. <https://doi.org/10.1016/j.dibe.2023.100300>.

**Conceptual Model >**



**Text Prompt >**



**AI Generated Image >**



**New Prompt: Wood >**



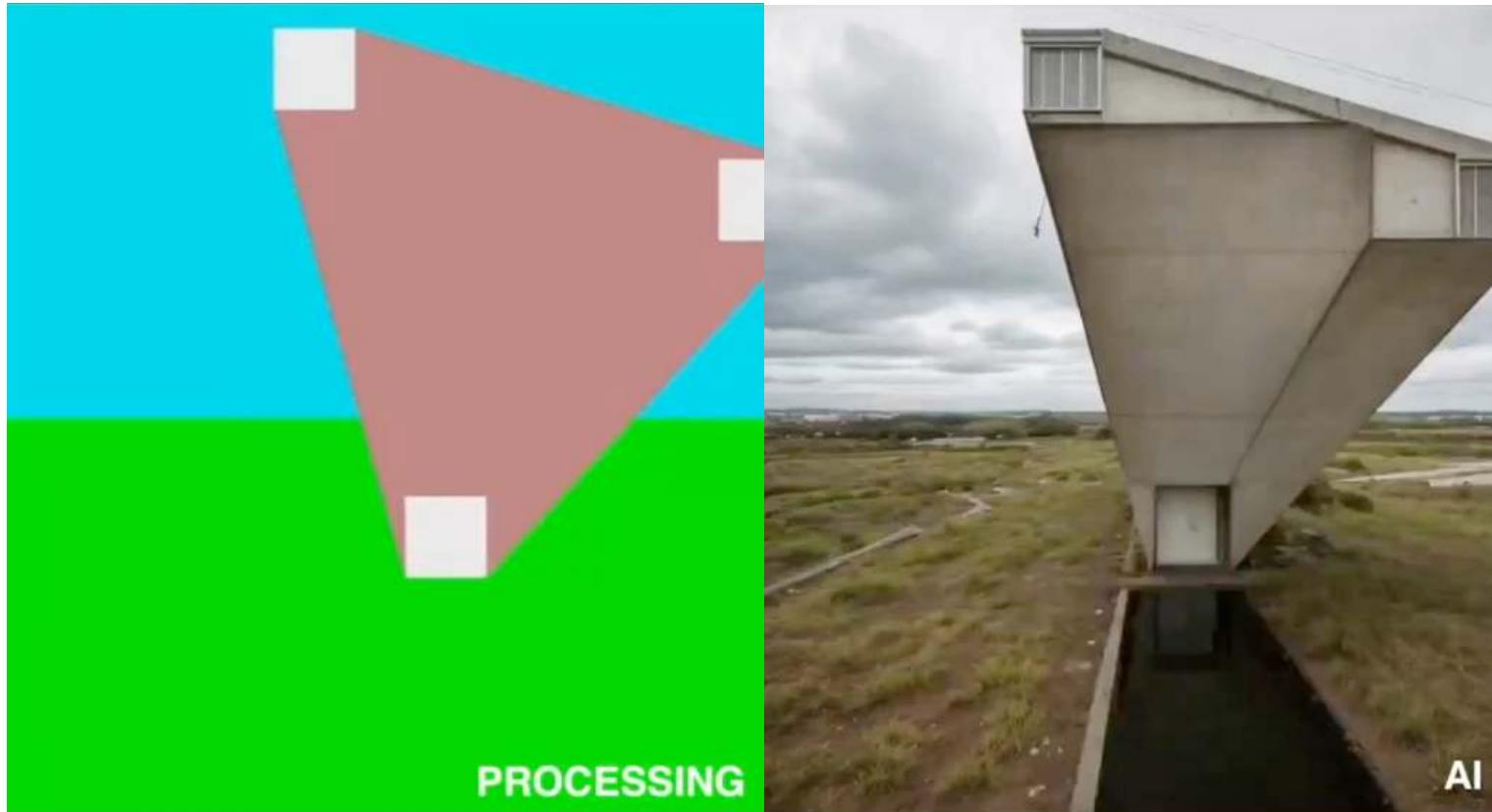
**New Prompt: Brick >**



# Archicad AI Visualizer

powered by **Stable Diffusion**

The Archicad AI Visualizer powered by Stable Diffusion is an experimental Archicad capability that creates inspiring, detailed, 3D visualizations during the early design stages.



AI



## Arturo Tedeschi teams up with STUD-IO and Wies Offsite to build an AI-designed structure

Arturo Tedeschi, a computational designer interested in AI-powered design tools, created a concept image using Midjourney and Stable Diffusion.

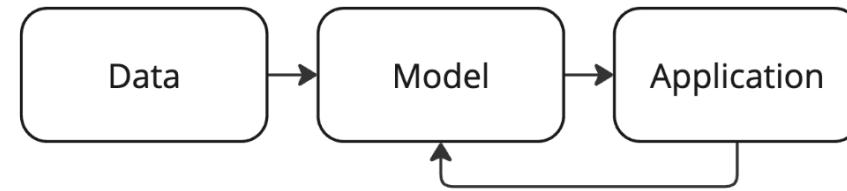
PA Editorial Team

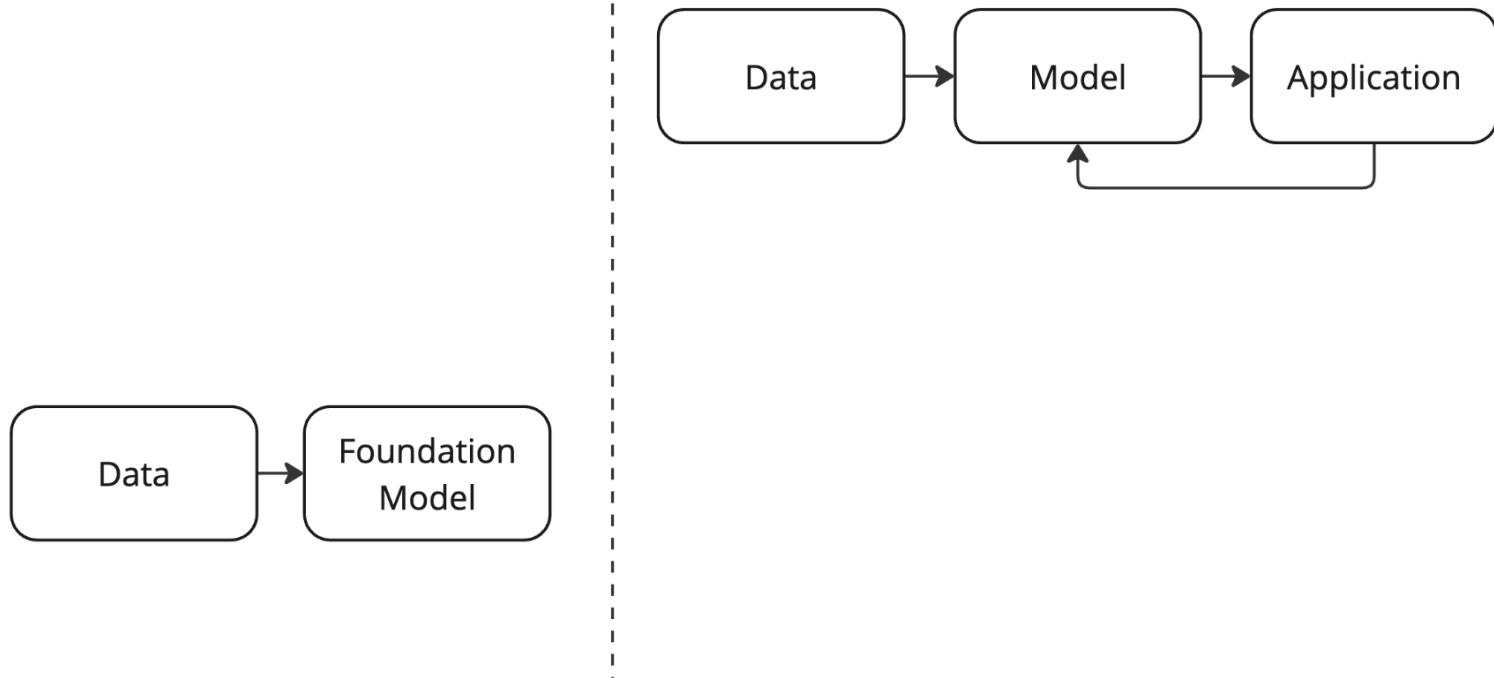
May 8, 2024

Architecture News



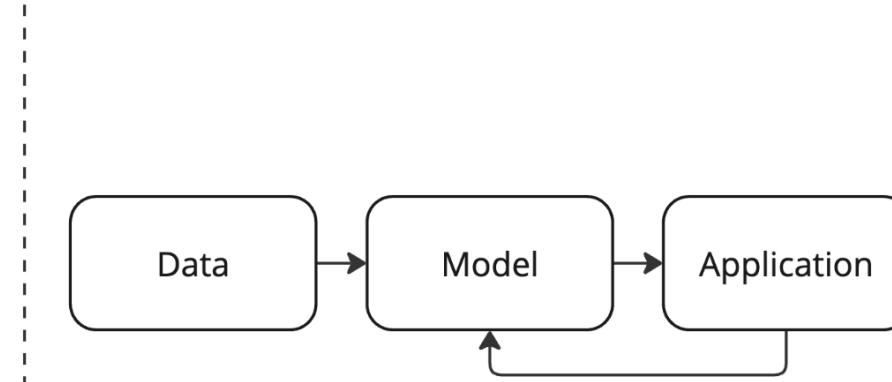
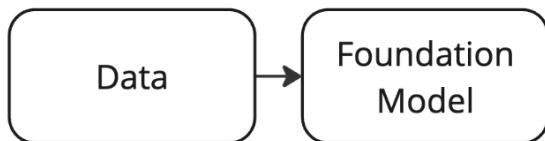


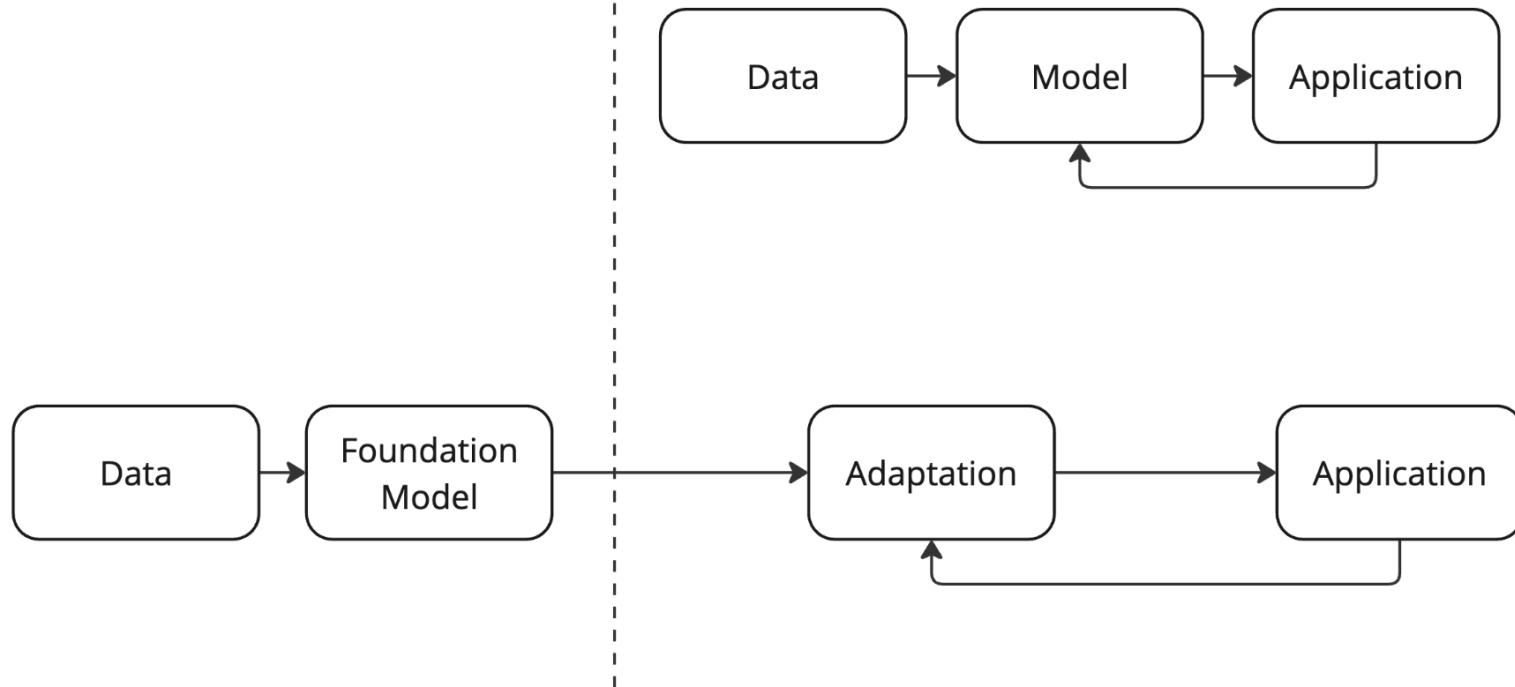


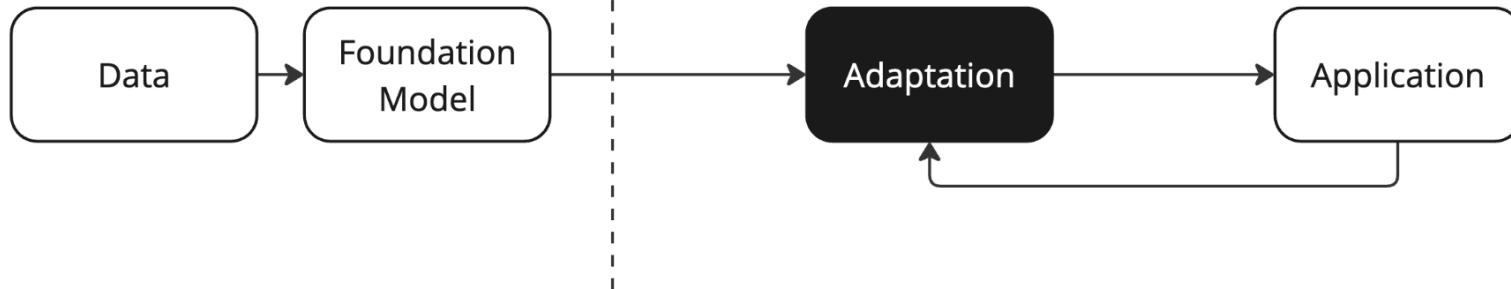


Data:  
GPT-3: 45 TB  
GPT-4: 1 PB

Compute:  
GPT-3: 10k GPUs running for 10 days  
(~10M USD)  
GPT-4: 10k GPUs running for 150 days







- Prompt engineering (context and goals)
- Define structure for the output
- Provide case-specific data (RAGs)
- Provide case-specific tools
- Adjust the control logic (systems and agents)
- Fine-tune with case-specific data

# Prompt engineering: Instructions frameworks

Instructions + examples help explain to the model what “parameter space” to stay within

## APE (Action → Purpose → Expectation)

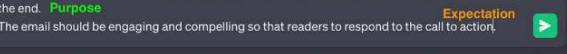
1. **Action** : Specify the job or task you want ChatGPT to perform.
2. **Purpose** : Define the intent of the task.
3. **Expectation** : Discuss what you expect the result to be.

Write a cold outreach email for our content marketing application. **Action**

The receivers of the email are content marketers. The email should introduce our business and key offerings, pricing and availability. End with a follow up plan and a call to action. Add a signature in the end. **Purpose**

**Expectation**

The email should be engaging and compelling so that readers to respond to the call to action.



## CARE (Context → Action → Result → Example)

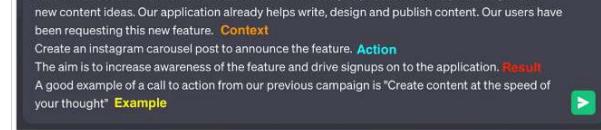
1. **Context** : Give all background information
2. **Action** : Specify the job or task you want ChatGPT to perform
3. **Result** : Discuss what you want to accomplish from the task
4. **Example** : Give one or more examples for ChatGPT to imitate.

We have released a new feature on our content marketing application that helps users generate new content ideas. Our application already helps write, design and publish content. Our users have been requesting this new feature. **Context**

Create an Instagram carousel post to announce the feature. **Action**

The aim is to increase awareness of the feature and drive signups on to the application. **Result**

A good example of a call to action from our previous campaign is "Create content at the speed of your thought". **Example**



## CREATE (Character → Request → Examples → Adjustment → Type of Output → Extras)

1. **Character** : Give a role to ChatGPT
2. **Request** : Define the job or task you want ChatGPT to perform
3. **Examples** : Give one or more examples for ChatGPT to imitate.
4. **Adjustment** : Provide refining instructions (like use bullet points, write short sentences)
5. **Type of Output** : Specify the format (like table, json, html)
6. **Extras** : Add more instructions or context

You are an expert copywriter experienced in creating sales landing pages. **Character**

Create a sales landing page for our content marketing application announcing the new content calendar feature. **Request**

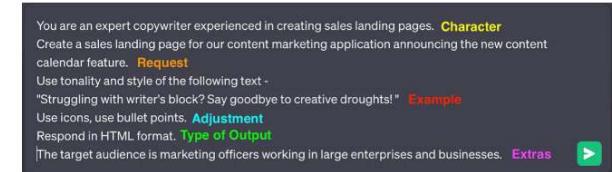
Use tonality and style of the following text -

"Struggling with writer's block? Say goodbye to creative droughts!" **Example**

Use icons, use bullet points. **Adjustment**

Respond in HTML format. **Type of Output**

The target audience is marketing officers working in large enterprises and businesses. **Extras**



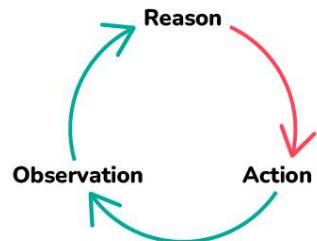
# Prompt engineering: Zero- vs Few-Shot Generation/Reasoning

Instructions + examples help explain to the model what “parameter space” to stay within

1. **Zero-Shot:** The model performs a task or answers a question without any prior examples or specific training on that task.
2. **One-Shot:** The model is given one example of the task, which it uses to understand and perform the task correctly.
3. **Few-Shot:** The model is provided with a few examples to guide it, effectively giving it a brief, targeted training session.

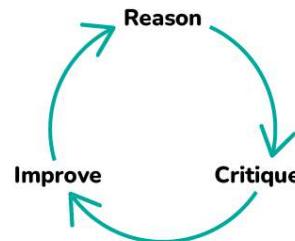
# Programmatic Control Logic

**ReAct**



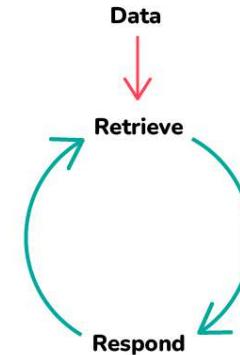
Reason + Act

**RCI**



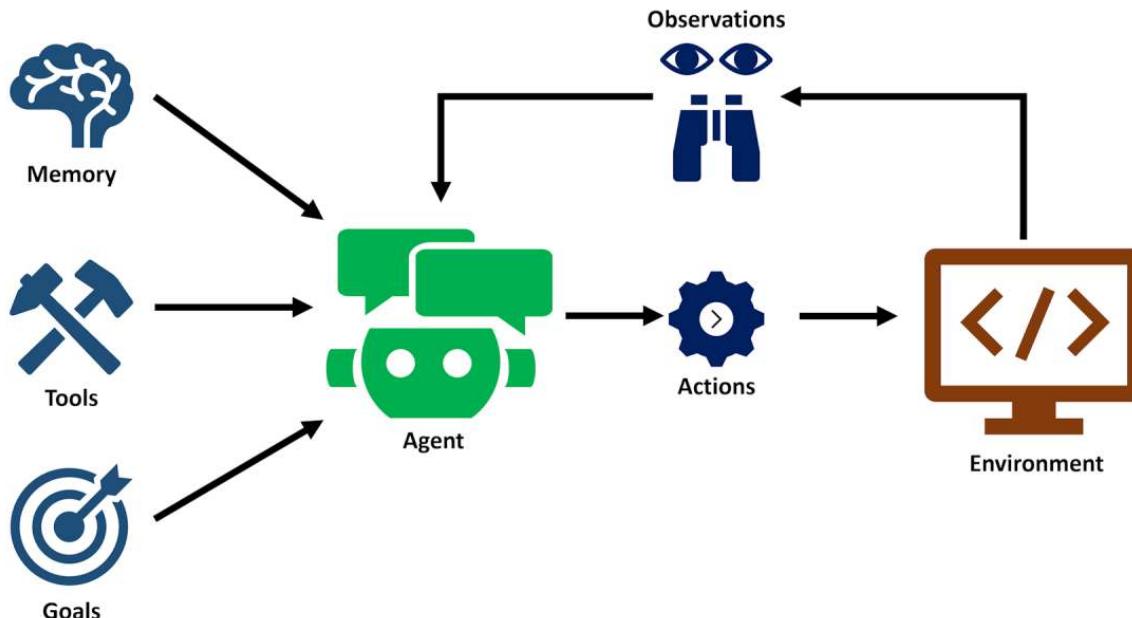
Recursive Criticism & Improvement

**RAG**



Retrieval Augmented Generation

# Agentic Control Logic, aka Agentic AI



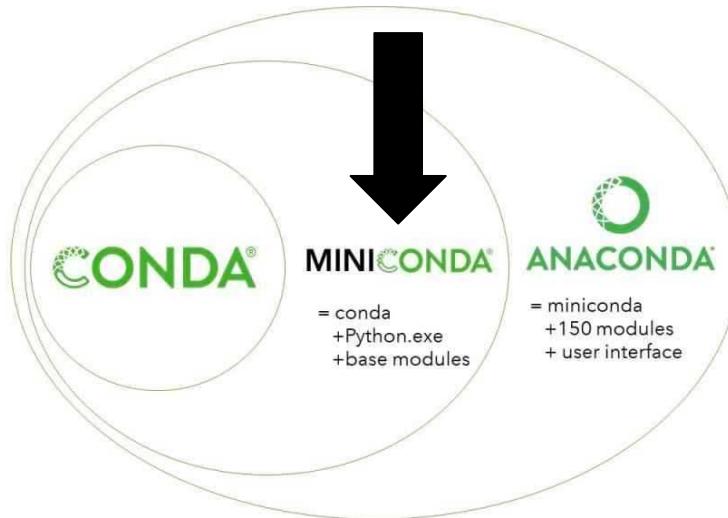
<https://www.linkedin.com/pulse/agent-ai-new-era-intelligent-app-development-tarun-sharma-laffc/>

# Hands on tutorial

CODE here:

<https://github.com/antonsavov/Intro-LLMs-for-AI-in-AEC>

# Install these



<https://docs.anaconda.com/miniconda/miniconda-install/>



<https://code.visualstudio.com/>