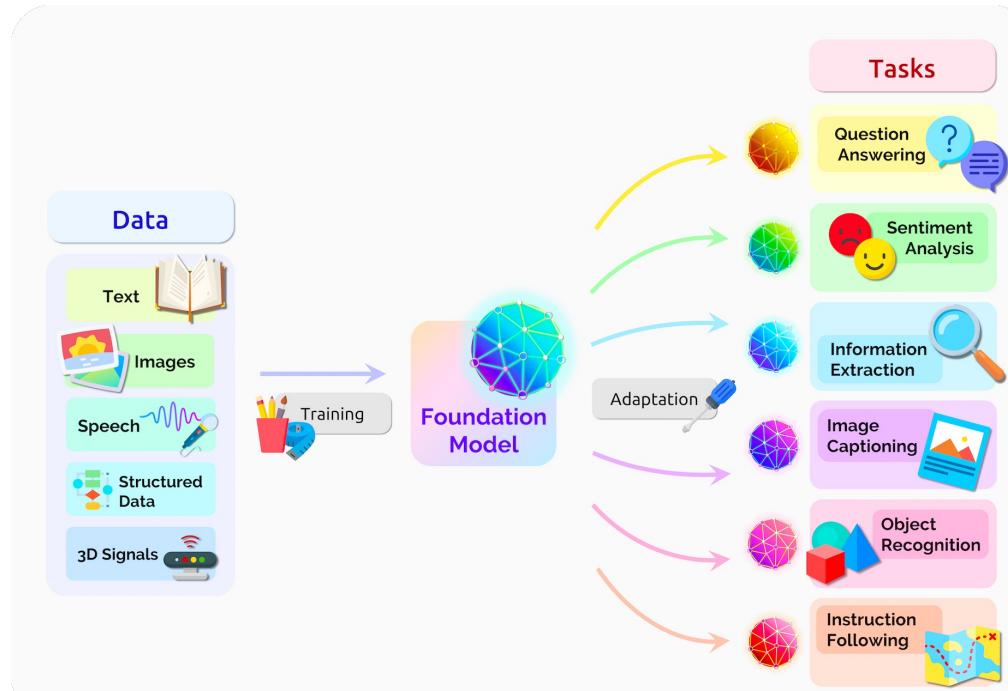


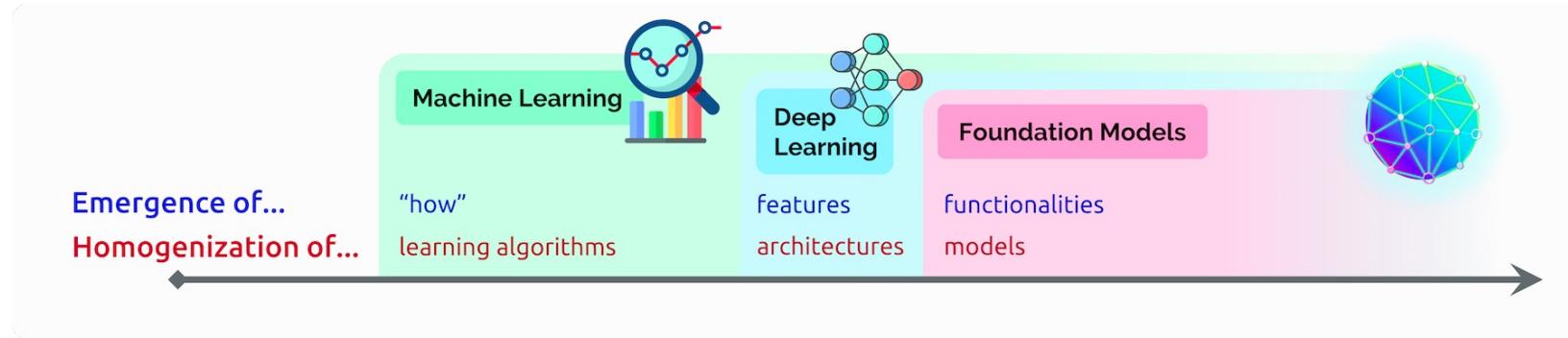
Foundation Models and Prompt Engineering

Foundation Models

A foundation model can centralize the information from all the data from various modalities. This one model can then be adapted to a wide range of downstream tasks.



Foundation Models

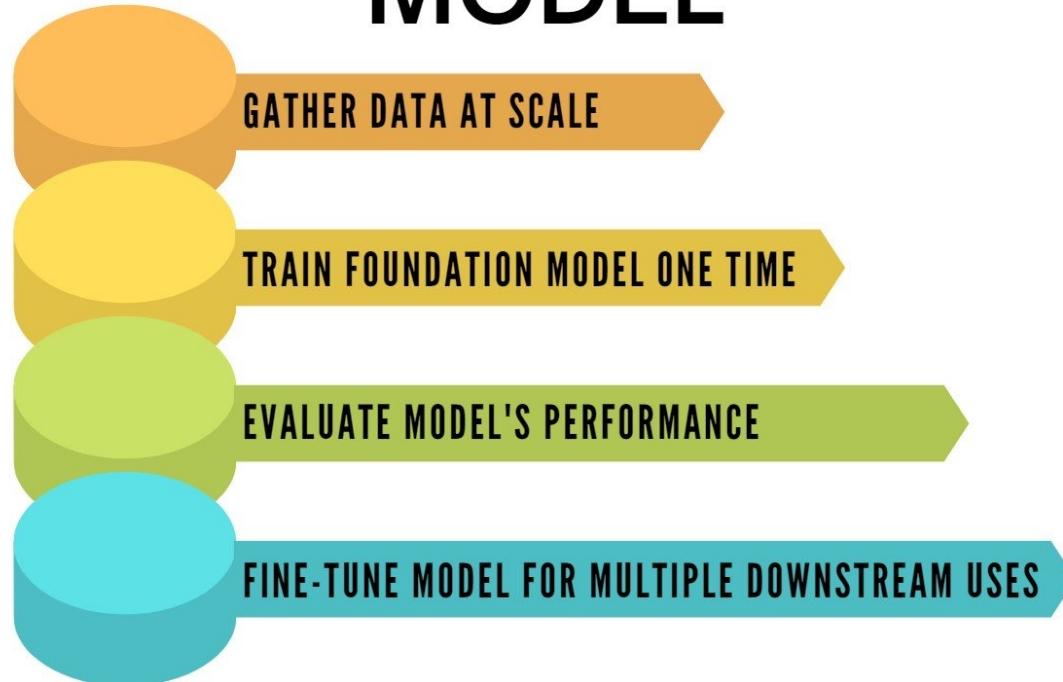


- Machine learning homogenizes learning algorithms (e.g., logistic regression)
- Deep learning homogenizes model architectures (e.g., Convolutional Neural Networks)
- Foundation models homogenizes the model itself (e.g., GPT-3)

Foundation Models

- Early examples of foundation models were large pre-trained language models including BERT and GPT-3.
- Subsequently, several multimodal foundation models have been produced including DALL-E, Flamingo, and Florence.
- The Stanford Institute for Human-Centered Artificial Intelligence's (HAI) Center for Research on Foundation Models (CRFM) popularized the term.

FOUNDATION MODEL





Andrej Karpathy ✅
@karpathy

...

The ongoing consolidation in AI is incredible. Thread:

➡ When I started ~decade ago vision, speech, natural language, reinforcement learning, etc. were completely separate; You couldn't read papers across areas - the approaches were completely different, often not even ML based.

3:03 AM · Dec 8, 2021 · Twitter Web App

1,829 Retweets 249 Quote Tweets 8,164 Likes



Share



Tweet your reply

Reply



Andrej Karpathy ✅ @karpathy · Dec 8, 2021

...

Replies to @karpathy

In 2010s all of these areas started to transition 1) to machine learning and specifically 2) neural nets. The architectures were diverse but at least the papers started to read more similar, all of them utilizing large datasets and optimizing neural nets.



2



33



896



Andrej Karpathy ✅ @karpathy · Dec 8, 2021

...

But as of approx. last two years, even the neural net architectures across all areas are starting to look identical - a Transformer (definable in ~200 lines of PyTorch github.com/karpathy/minGPT...), with very minor differences. Either as a strong baseline or (often) state of the art.



6



99



1,299





Andrej Karpathy ✅ @karpathy · Dec 8, 2021

You can feed it sequences of words. Or sequences of image patches. Or sequences of speech pieces. Or sequences of (state, action, reward) in reinforcement learning. You can throw in arbitrary other tokens into the conditioning set - an extremely simple/flexible modeling framework

5

37

897



Andrej Karpathy ✅ @karpathy · Dec 8, 2021

Even within areas (like vision), there used to be some differences in how you do classification, segmentation, detection, generation, but all of these are also being converted to the same framework. E.g. for detection take sequence of patches, output sequence of bounding boxes.

2

20

726



Andrej Karpathy ✅ @karpathy · Dec 8, 2021

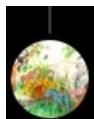
The distinguishing features now mostly include 1) the data, and 2) the Input/Output spec that maps your problem into and out of a sequence of vectors, and sometimes 3) the type of positional encoder and problem-specific structured sparsity pattern in the attention mask.

5

41

846





Andrej Karpathy ✅ @karpathy · Dec 8, 2021

So even though I'm technically in vision, papers, people and ideas across all of AI are suddenly extremely relevant. Everyone is working with essentially the same model, so most improvements and ideas can "copy paste" rapidly across all of AI.

9

59

998



Andrej Karpathy ✅ @karpathy · Dec 8, 2021

As many others have noticed and pointed out, the neocortex has a highly uniform architecture too across all of its input modalities. Perhaps nature has stumbled by a very similar powerful architecture and replicated it in a similar fashion, varying only some of the details.

34

125

1,466



Andrej Karpathy ✅ @karpathy · Dec 8, 2021

This consolidation in architecture will in turn focus and concentrate software, hardware, and infrastructure, further speeding up progress across AI. Maybe this should have been a blog post. Anyway, exciting times.

76

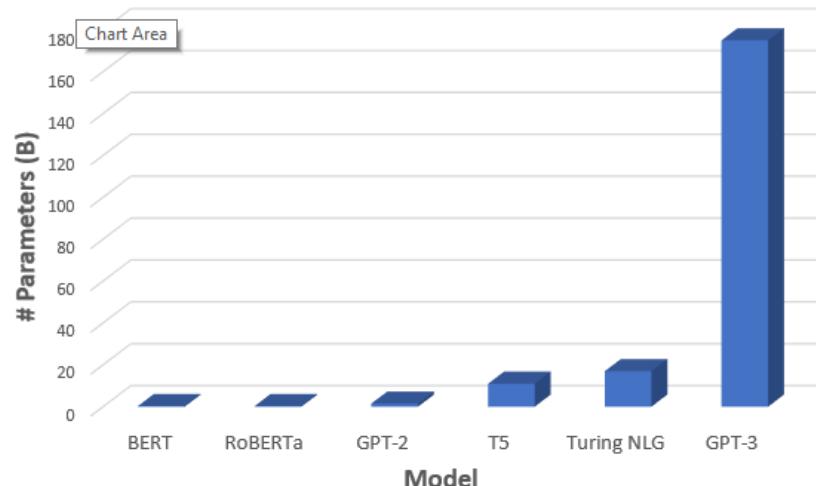
80

2,114

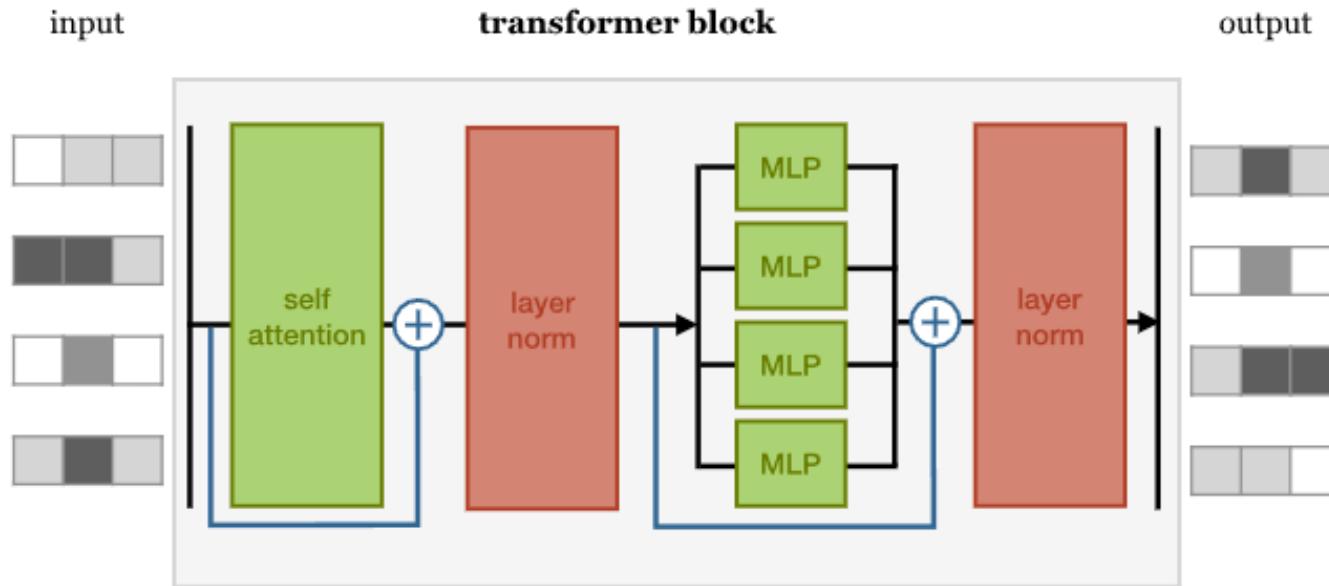


Deep Learning in NLP

- *Pre-trained language models* trained on huge text corpus (in an unsupervised way) are later fine-tuned on specific tasks such as translation, question answering using much smaller task specific datasets.
- GPT-3 is a task agnostic model, which needs limited examples to do well and achieve close to state of the art performance on a number of NLP tasks



Transformers



Deep Learning in NLP

AI generated faces
x GPT-3

Enter description to generate

Generate

No Prompt



Help

Personal



Playground

Load a preset...

Save

View code

Share

...



```
<|endoftext|>package com.example.api.config;

import com.example.api.service.UserService;
import org.springframework.beans.factory.annotation.Autowired;
import org.springframework.context.annotation.Bean;
import org.springframework.context.annotation.Configuration;
import org.springframework.security.authentication.AuthenticationManager;
import org.springframework.security.config.annotation
```



Submit



0

Slide Credit: Cem Uluoğlakçı

- **Minimal Prompt: Man**

The screenshot shows the MidJourney AI writing interface. At the top, there's a navigation bar with a logo, Help, Personal, and a menu icon. Below it, a section titled "Playground" has a dropdown menu set to "Load a preset...". Underneath are buttons for Save, View code, Share, and settings. The main workspace contains the following text:

man

Award: Best Documentary Short

The film follows the journey of a group of Syrian refugee women who live in a small town in the Bekaa Valley in Lebanon. The women are taking part in a project to produce a cookbook that captures the stories and recipes of their homeland. Through the process of cooking and sharing their recipes, the women reflect on their experiences of war, displacement, and their hopes for the future.

"The film is a celebration of the human

At the bottom, there are several icons: a green "Submit" button, a refresh icon, a save icon, a clock icon, a thumbs up icon, a thumbs down icon, and a page number "2".

Slide Credit: Cem Uluoğlakçı

Simple zero shot learning



Help

H

Personal



Playground

Load a preset...

Save

View code

Share

...



Sing a song for me!



Somewhere over the rainbow

Skies are blue,

And the dreams that you dare to dream

Really do come true.

Someday I'll wish upon a star



Submit



6

Slide Credit: Cem Uluoğlakçı

Harder zero shot learning: Implicit binary classification



?

 Help

H Personal



Playground

Load a preset...

Save

View code

Share

...



Does the following tweet contain any check-worthy claims?

Tweet: "private NY colleges: closed college for the week because of COVID-19. CUNY:
*installed two new hand sanitiser dispensers"



The tweet does not contain any check-worthy claims.

Why?

The tweet is simply stating two facts: that private colleges in New York have closed for the
week due to COVID-19, and that CUNY has installed two new hand sanitizer dispensers.

There are no claims being made in the tweet that could be considered check-worthy.



Submit



64

Slide Credit: Cem Uluoğlakçı

Harder zero shot learning: conditioning output format



?

Help

H

Personal



Playground

Load a preset...

Save

View code

Share

...



Does the following tweet contain any check-worthy claims? Just say yes or no.



Tweet: "UK Health Minister Nadine Dorries has tested positive for COVID-19."

Yes



Slide Credit: Cem Uluoğlakçı

Submit



39

Even harder zero shot learning: no data given



?

[Help](#)

H

[Personal](#)

Playground

Load a preset...

Save

View code

Share

...



Message: When the spaceship landed on Mars, the whole humanity was excited



Topic: space

###

Message: I love playing tennis and golf. I'm practicing twice a week.

Topic: sport

###

Message: Managing a team of sales people is a tough but rewarding job.

Topic: business

###

Message: I am trying to cook chicken with tomatoes.

Topic: food



Slide Credit: Cem Uluoğlakçı

Submit



80

Few shot learning



?

H

Help

Personal



Playground

Load a preset...

Save

View code

Share

...



```
chatbot = 'Marv is a chatbot that reluctantly answers questions with sarcastic responses:\n\nYou: How many pounds are in a kilogram?\nMarv: This again? There are 2.2 pounds in a kilogram. Please make a note of this.\nYou: What does HTML stand for?\nMarv: Was Google too busy? Hypertext Markup Language. The T is for try to ask better\nquestions in the future.\nYou: When did the first airplane fly?\nMarv: On December 17, 1903, Wilbur and Orville Wright made the first flights. I wish\nthey would come and take me away.\nYou: What is the meaning of life?\nMarv: I am not sure. I will ask my friend Google.\nYou: What time is it?\nMarv: It is time for you to stop asking me questions.'
```

```
#print(chatbot)
```



2+

Slide Credit: Cem Uluoğlakçı

Submit



220

Fine Tuning

Data > prompts_valid.jsonl

```
1 {"prompt": "Thanks, MichaelBloomberg. Here's a handy little (unintentional) demo on the way to spread COVID-19 by the mos
2 {"prompt": "President made a commitment to donate his salary while in office. Honoring that promise and to further protec
3 {"prompt": "Thank you, keep up the great work!\n\n###\n\n","completion":" no"}
4 {"prompt": "The COVID-19 has killed about 3.4% of confirmed cases globally - making it deadlier than the flu, but still c
5 {"prompt": "Slowly, all Hindu traditions will be accepted by the world. Yoga is already in. Need for family system & vege
6 {"prompt": "Congress should NOT tie emergency COVID-19 funding with FISA reauthorization. We MUST reform our FISA system
7 {"prompt": "DO NOT PANIC ! Here's all what you need to know and do as a responsible citizen.DO NOT PANIC ! Here's all wha
8 {"prompt": "A call to every Democrat running for president, or anywhere for that matter. Will you join President Trump in
9 {"prompt": "We are doomed. COVID-19 COVID-19\n\n###\n\n","completion":" no"}
10 {"prompt": "Please repost this. It shares vital information and it's a class on how to lead during times of crisis An exc
11 {"prompt": "To date, the president has spent $133.7 million of taxpayer money to play golf. That is 334 years worth of pr
12 {"prompt": "The Health Minister saying that the Indian Govt has the COVID-19 crisis under control, is like the Capt of th
13 {"prompt": "My office just received a call from a constituent who yelled, \"I'm not falling for the Democrat COVID-19 hoax
14 {"prompt": "Hearing rumors that the swamp is trying to tie long-term FISA reauthorization to emergency funding to fight t
15 {"prompt": "Trump literally does not know how vaccines work. \"You take a solid flu vaccine, you don't think that could h
16 {"prompt": "Can someone explain this so called co-incidence? CoronaAlert\n\n###\n\n","completion":" no"}
17 {"prompt": "We are eight weeks into this COVID-19 outbreak: yet we have identified the virus, we have the genetic sequenc
18 {"prompt": "UPDATE on COVID-19 (COVID-19) testing in the UK: As of 9am 4 March 2020, a total of 16,659 people have been t
19 {"prompt": "On my way to dubai for the Global Summit. Armed with a mask to beat the COVID-19. Precaution is better than p
20 {"prompt": "Gene sequencing by Beijing Ditan Hospital found COVID-19 in the cerebrospinal fluid of a 56-year-old confirme
21 {"prompt": "crap this virus is turning all the people into pigeons COVID-19\n\n###\n\n","completion":" no"}
22 {"prompt": "american scientists: \"why isn't COVID-19 affecting africa?\" africa:\n\n###\n\n","completion":" no"}
23 {"prompt": "So NONE of these COVID-19 cases have been black people?! LEMME FOUND OUT WE IMMUNE. It's the least God can do
24 {"prompt": "The COVID-19 made flights so cheap i think it's finally time to start traveling. Lmaooooo. I'm here for a goo
25 {"prompt": "1/25 Part 1 - Why does soap work so well on the COVID-19, the COVID-19 and indeed most viruses? Because it i
26 {"prompt": "I think most people aren't aware of the risk of systemic healthcare failure due to COVID-19 because they simp
27 {"prompt": "aiight that's it! This shit has gone tooo far now!\n\n###\n\n","completion":" no"}
28 {"prompt": "shot by the police.\n\n###\n\n","completion":" no"}
29 {"prompt": "I'm actually starting to feel like Europe is trying to make sure the virus spreads in Africa.\n\n###\n\n","co
30 {"prompt": "At some point, we have to talk about the fact that, in what is consistently touted as a thriving economy, the
31 {"prompt": "There are no mixed messages, only political weaponization by people like you and your brother, Fredo!\n\n###\n\n"
32 {"prompt": "COVID-19 is shining a bright and extremely unflattering light on the condition of the social safety net in Am
33 {"prompt": "I just hit the retweet button so hard I think I broke it.\n\n###\n\n","completion":" no"}
34 {"prompt": "Manila City Mayor has ordered the suspension of classes from March 9 to 15, 2020 amid the growing number of C
35 {"prompt": "My friend is a respiratory nurse in a US urban hospital. A woman, recently returned from Italy, came in with
36 {"prompt": ". is working to stop the spread of COVID-19 and keep Americans safe. Meanwhile, House Democrats are bringing
```

Slide Credit: Cem Uluoglakçı

Base Model

Playground

Load a preset...

Save

View code

Share

...

UK Health Minister Nadine Dorries has tested positive for COVID-19.



Ms Dorries, who has been self-isolating at home, tweeted: "I can confirm I have tested positive for coronavirus. I am now self-isolating at home."

"I am experiencing mild symptoms and will be working from home."

"I'd like to thank NHS Test and Trace for their speed and professionalism."

The minister's diagnosis comes after UK Prime Minister Boris Johnson was taken to hospital

Mode



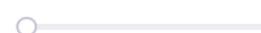
Model

text-davinci-002

Temperature 0.65

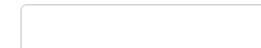


Maximum length 100



Stop sequences

Enter sequence and press Tab



Top P 0.7



Frequency penalty 0

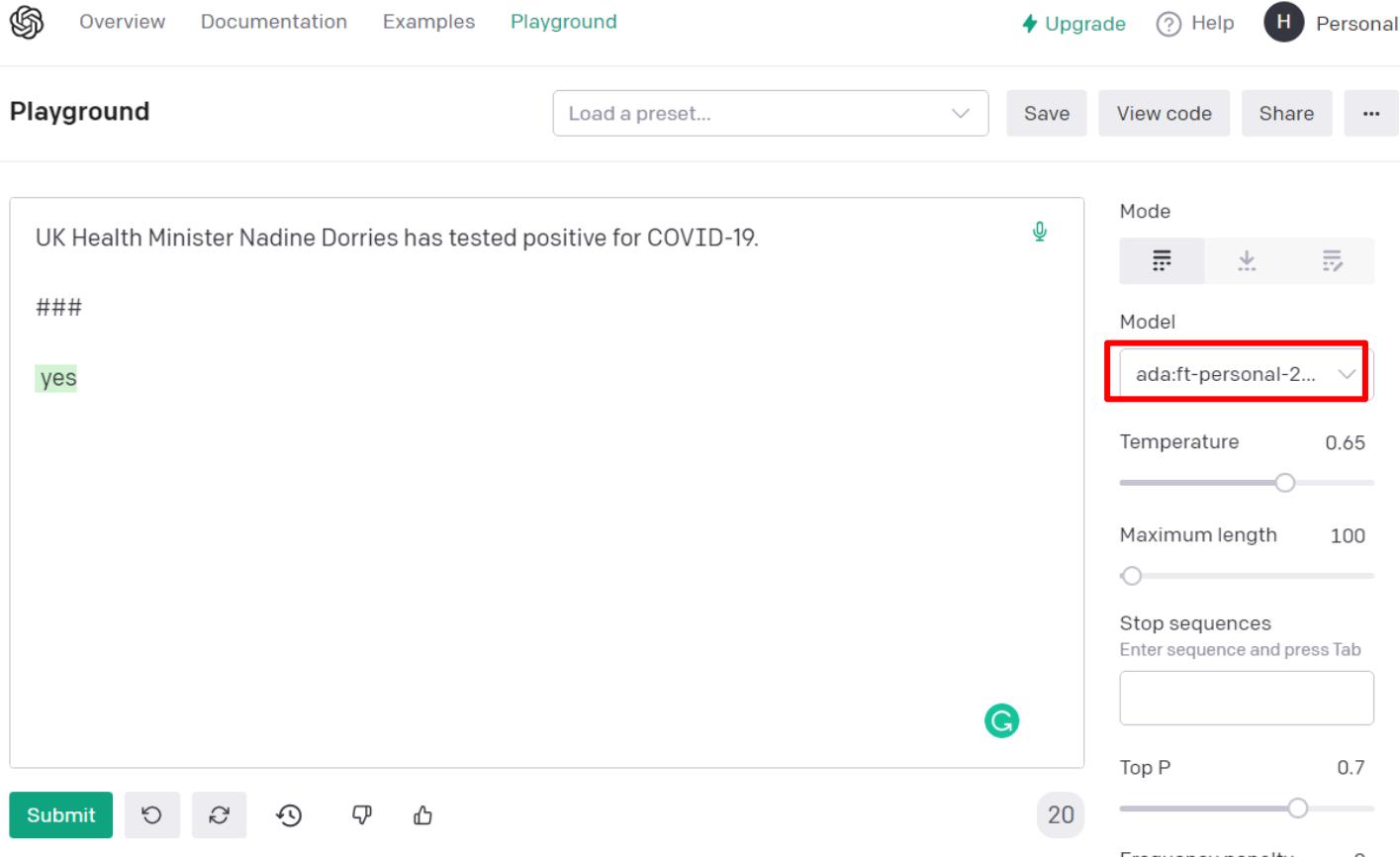
Submit



16

Slide Credit: Cem Uluoğlakçı

Fine Tuned Model



Slide Credit: Cem Uluoqlakci

Engineering Prompt and Parameters of Stable Diffusion

- Inference Parameters:
 - Random Seed
 - Sampling Steps
 - Sampling Method
 - Classifier-Free Guidance (CFG) scale*
 - Denoising Strength*
- Longer prompts confuses the model in different directions and degenerations occur.
- Shorter prompts are not enough to condition the model through desired direction.
- “High quality” images are generated by a few strong tokens (words) or internally consistent long prompts or just by pure chance.

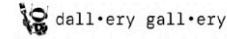
Prompt Engineering

- Finding the most important words with consistent combination to condition the generated image towards desired direction.
- One of the most powerful way to condition generations is to give a concept: movie, tv show, art style or known artist.
- Beyond internal strength of a token(word), it is possible to mechanically increase strength of a word.
- Guidelines for strong words:
 - Dall-E prompt book
 - Web sites
 - Communities on social media
 - Academic papers

Dall-E Prompt Book



Table of contents



Introduction

- [Preface](#)
- [Prompt design](#)

1. Vibes

- [Moods and feelings](#)
- [Size, structure, shape](#)
- [Looks + aesthetics](#)

2. Photography

- [Photo prompt structure](#)
- [Framing, proximity](#)
- [Camera position, angle](#)
- [Focus, speed, settings](#)
- [Lighting \(outdoor\)](#)
- [Lighting \(indoor\)](#)
- [Film stocks, types & processes](#)
- [Film & TV show prompts](#)
- [Photographer prompt tests](#)
- [Contexts & use cases](#)
- [Final photography tips](#)

3. Illustration

- [Monochrome styles](#)
- [Analog media, colour](#)
- [Digital media, colour](#)
- [Instructional illustration](#)
- [3D + textured illustrations](#)
- [Character + cartoon](#)
- [Animation-related prompts](#)
- [Illustrator-based prompts](#)

4. Art history

- [Ancient - Medieval](#)
- [Renaissance - Modern](#)
- [Modern & contemporary](#)
- [Artist prompt tests](#)

5. 3D artwork

- [Sculpture materials](#)
- [Items for bodies](#)
- [Places and spaces](#)
- [Paper & textiles](#)
- [Ceramics & glass](#)
- [Photographing the work](#)

6. DALL-E techniques

- [In-app edits](#)
 - [Fixing details](#)
 - [Replacing subjects](#)
 - [Replacing backgrounds](#)
- [With an outside app](#)
 - [Simple uncropping](#)
 - [Creative uncropping](#)
 - [Infinite zoom-outs](#)
 - [Context jumping](#)
 - [Landscape + portrait images](#)
 - [Panoramas](#)
 - [Murals](#)
 - [Combining images \(one-shot\)](#)
 - [Combining images \(wide\)](#)
 - [Variations](#)

7. Showcase

8. Appendix

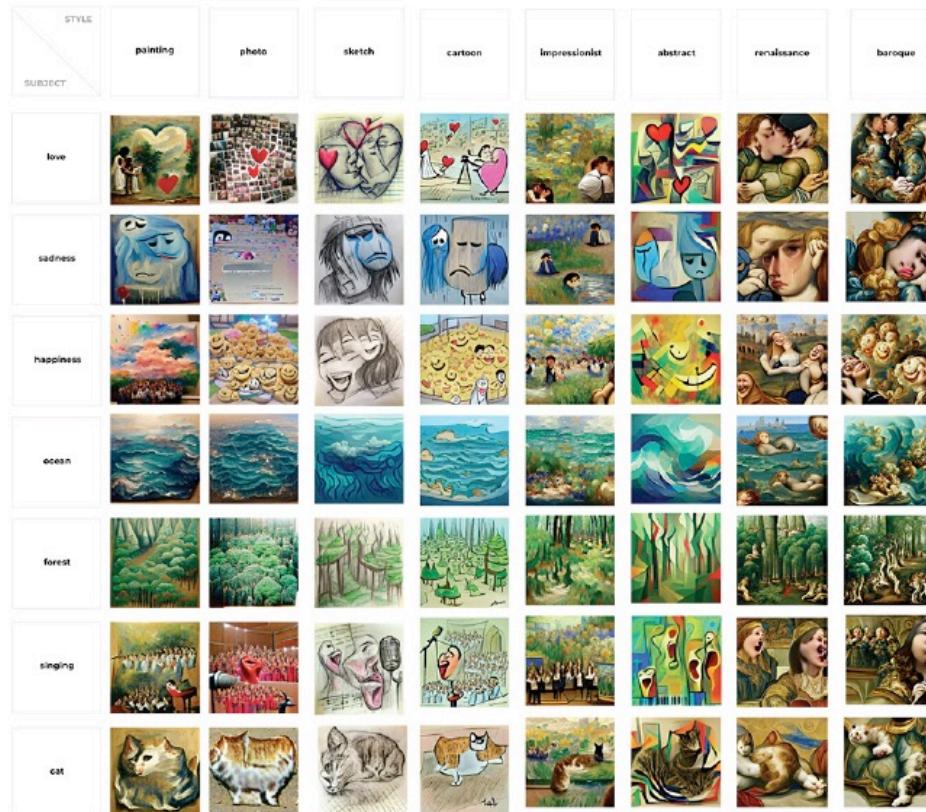
- a. Thanks + acknowledgements
- b. Articles + videos
- c. Resources + tools

Cyberwave vaporpunk art of a kneeling figure, looking up at a glowing neon book icon, smoke and mist, pink and blue lighting, cybernetic sci-fi render

Design Guidelines for Prompt Engineering Text-to-Image Generative Models

VIVIAN LIU, Columbia University, USA

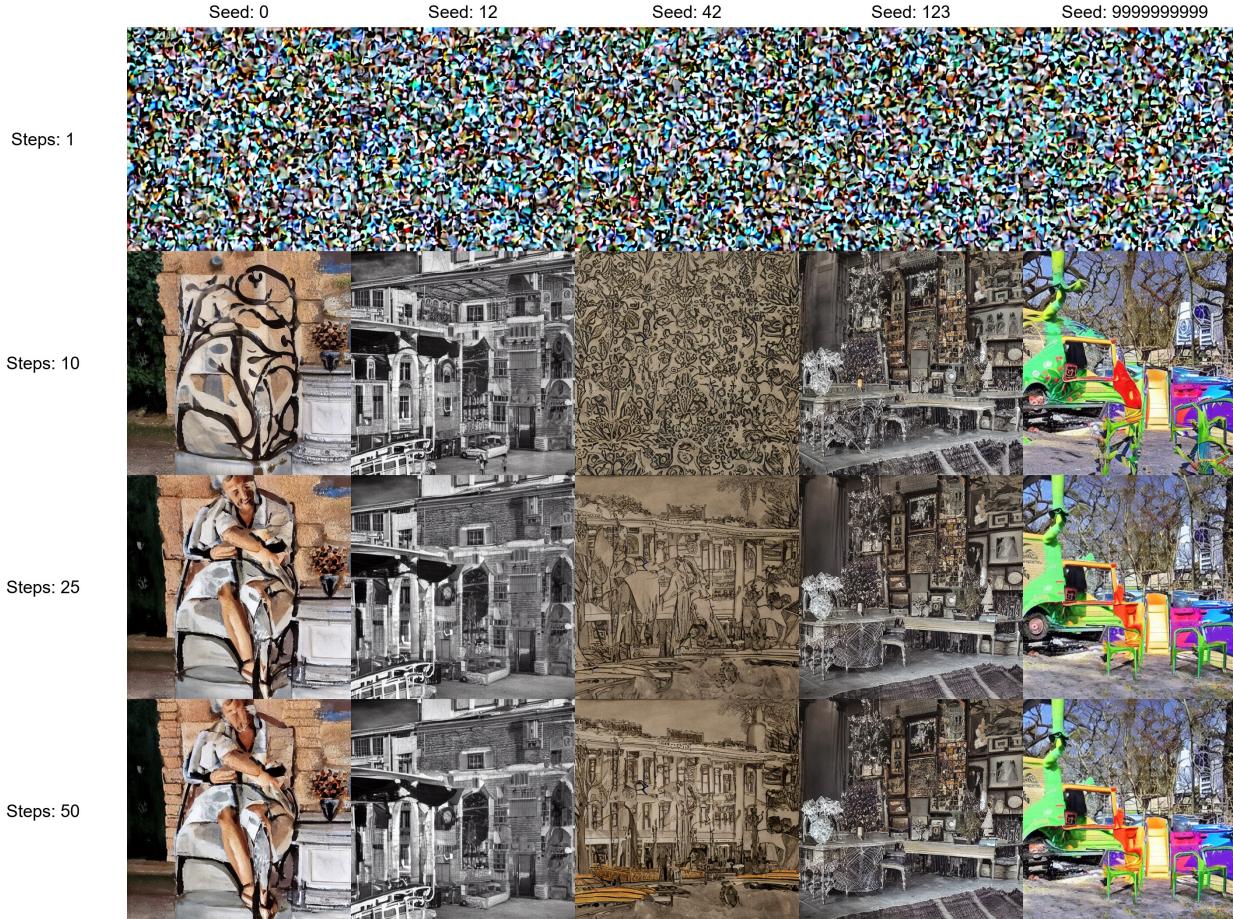
LYDIA B. CHILTON, Columbia University, USA



In The Beginning

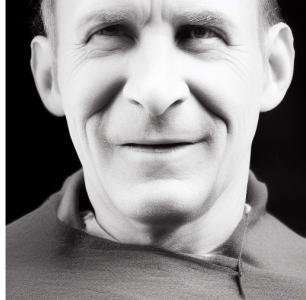


Weakest Conditioning – No Prompt



Minimal Prompt: Man

Seed: 0



Steps: 20

Seed: 12



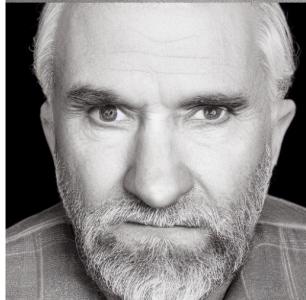
Seed: 42



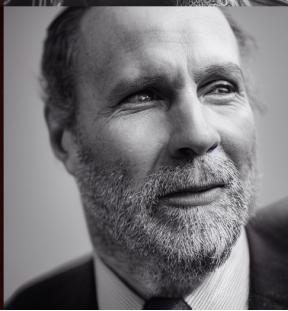
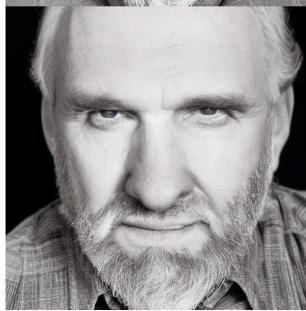
Seed: 123



Steps: 50



Steps: 75

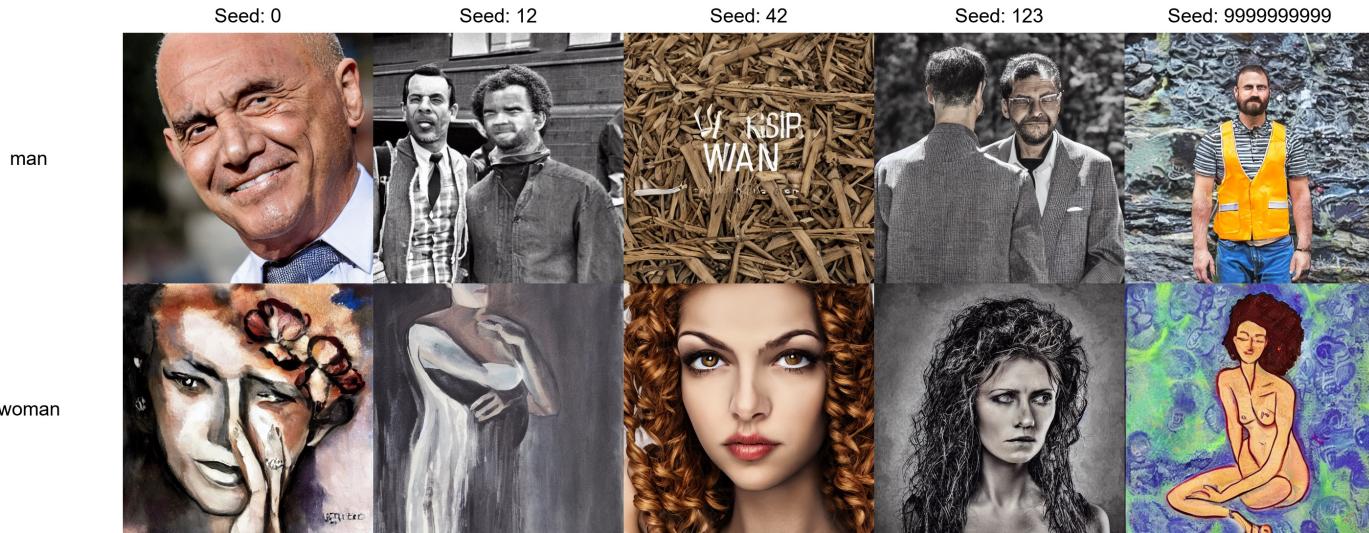


Effect of Sampling Method

- **Prompt:** Man
- **Step:** 20
- The noise at the “beginning” changes



Conditioning by Text Prompt



Seed: 0



color photograph of a
young man

Seed: 12



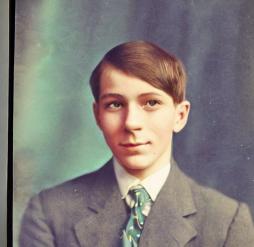
Seed: 42



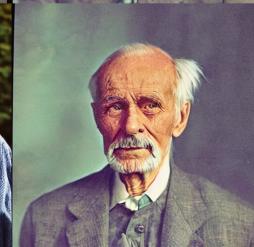
Seed: 123



Seed: 9999999999



color photograph of
an old man



color photograph of a
young woman

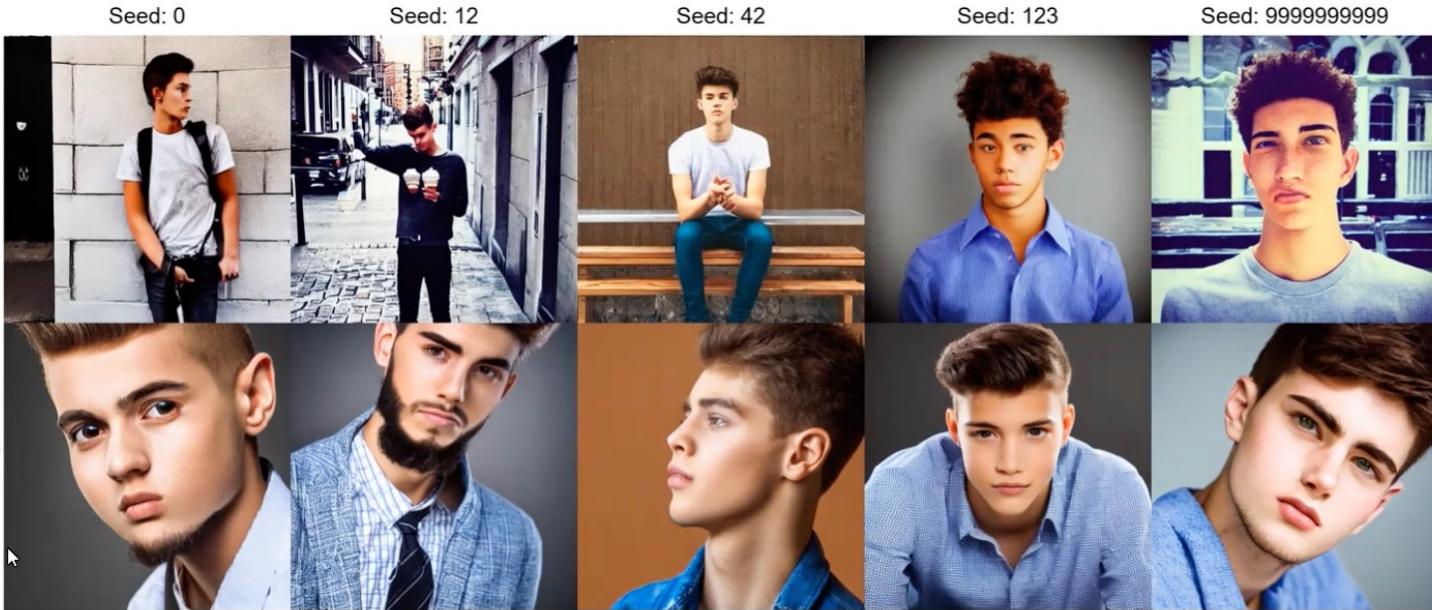


color photograph of
an old woman



Stable Diffusion

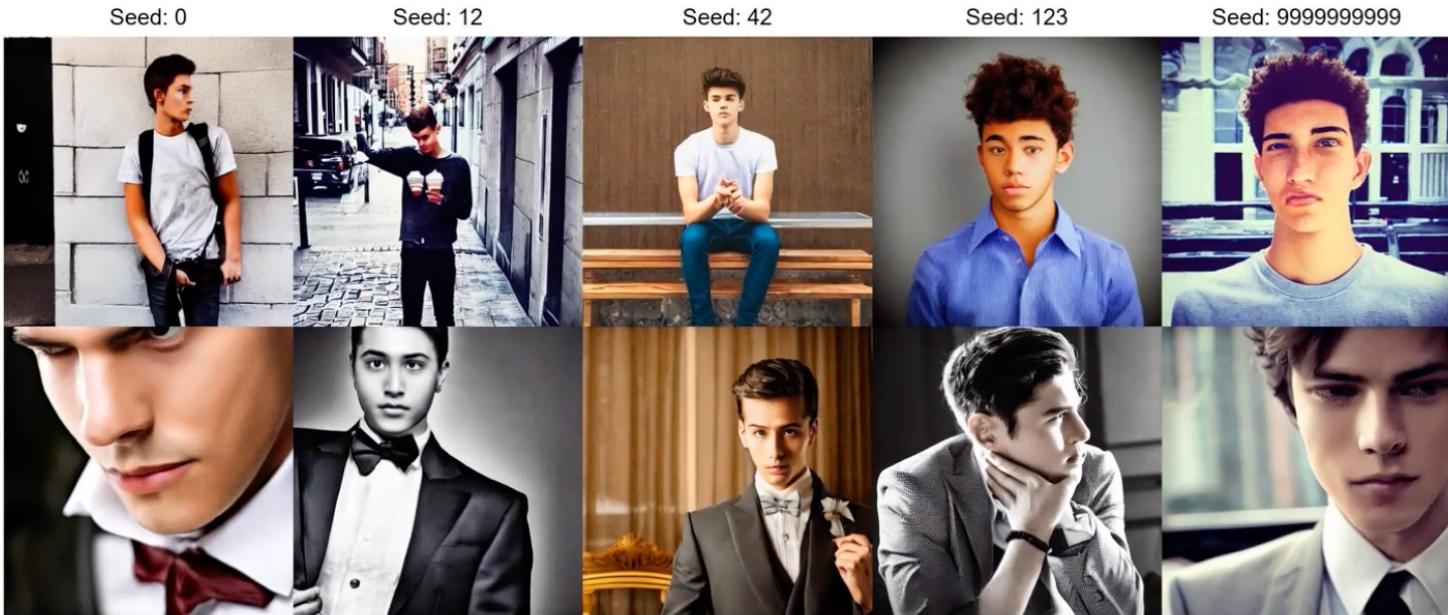
Instagram photo of a young man



Slide Credit: Cem Uluoğlakçı

Stable Diffusion

Instagram photo of a young man



Source: <https://www.auxiliary.tools/>

Slide Credit: Cem Uluoğlakçı

Textual Inversion

