Prof. Dr. Kim Albrecht, Yağmur Uçkunkaya & Lars Christian Schmidt

My AI is better than yours

Workshop on training AI Models as creative practice

Folkwang University 06.10 – 13.10.25

Schedule

Titel	Thema	Datum	Zeit	Info	Abgabe
Einführung	Intro	06.10.25	10:00 – 12:00	Einführung ins Maschinelle Lernen und aufzeigen Künstlerischer Positionen in dem Gebiet	
Ideation Workshop	Modell Wählen	06.10.25	13:00 – 15:00 (17:00?)	Ideen Findung	
Ideation Workshop	Gruppe finden	06.10.25	10:00 – 12:00	Idee Ausarbeiten	
Gathering	Sammeln	07.10.25	12:00 – 17:00	Feedback Gespräche 3.07	
Zwischenvorstellung der Sammlungen	Vorstellen & Feedback	07.10.25	17:00 – 18:00	Vorstellung der Sammlung und Idee des Projektes.	Abgabe eines Ordners von min. 20% der Sammlung
Gathering	Sammeln	08.10.25	10:00 – 17:00	Feedback Gespräche 3.07	
Vorstellung der Sammlungen	Vorstellen & Feedback	08.10.25	17:00 – 18:00	Vorstellung der fertigen Sammlung	Ordner aller Dateien zum Training.
Projektentwicklung	Trainieren & Ausarbeiten	09.10.25	10:00 – 18:00	Modelle Trainieren mit Hilfe von Lars Christian Schmidt	
Projektentwicklung	Trainieren & Ausarbeiten	10.10.25	10:00 – 18:00	Modelle Trainieren mit Hilfe von Lars Christian Schmidt	
Projektentwicklung	Trainieren & Ausarbeiten	13.10.25	10:00 – 14:00		
Finale Präsentation	Vorstellung des Projektes	13.10.25	14:00 – 16:00		15 min Präsentation

Outcomes

- 30sek to 1 min Video of your experiments / outcomes, experiences
- 15 minute presentations of your projects
- Anything additional you come up with, images, booklet, installation, interactive software, text, etc.

Beyond the Workshop

databased.design

kurse notizen info

data based

en

INFORMATION DESIGN OPEN STUDIO WS25

Winter 2025–26 | 25-10-14 to 26-01-27
Lecturer: Kim Albrecht

Offenes Projektstudio für Informationsgestaltung
Jeden Dienstag 14–17 Uhr, Raum 3.07

Das Information Design Open Studio ist eine offene Plattform zur Entwicklung eigener Projekte im Bereich Informationsdesign. Der Kurs richtet sich an alle, die ein individuelles Vorhaben umsetzen, frühere Arbeiten vertiefen oder sich mit dem Feld gestalterisch und theoretisch auseinandersetzen möchten.

Es gibt verschiedene Möglichkeiten der Teilnahme:

- Die Experimente aus dem Workshop “My AI is Better Than Yours” können hier weiterentwickelt und zu einer umfassenden

Goals

Pick one model, collect a dataset, train from scratch or near-scratch, reflect on results.

Understand AI as a material and cultural process, not a black box.

Goals

Experiments over outcomes

Strangeness over perfection

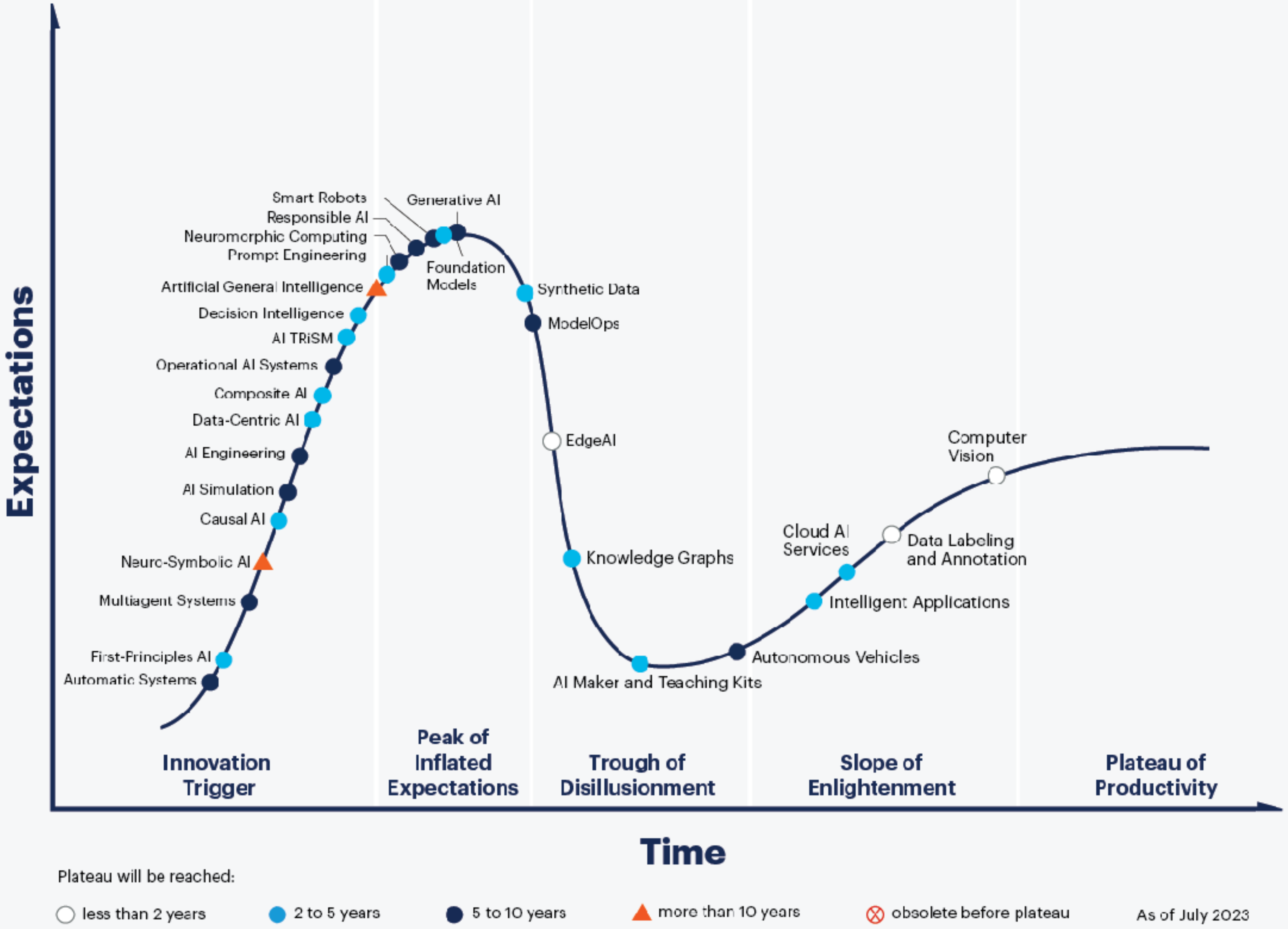
Fun over Exhaustion

State of AI

or why it is necessary for us to do this workshop

Generative AI

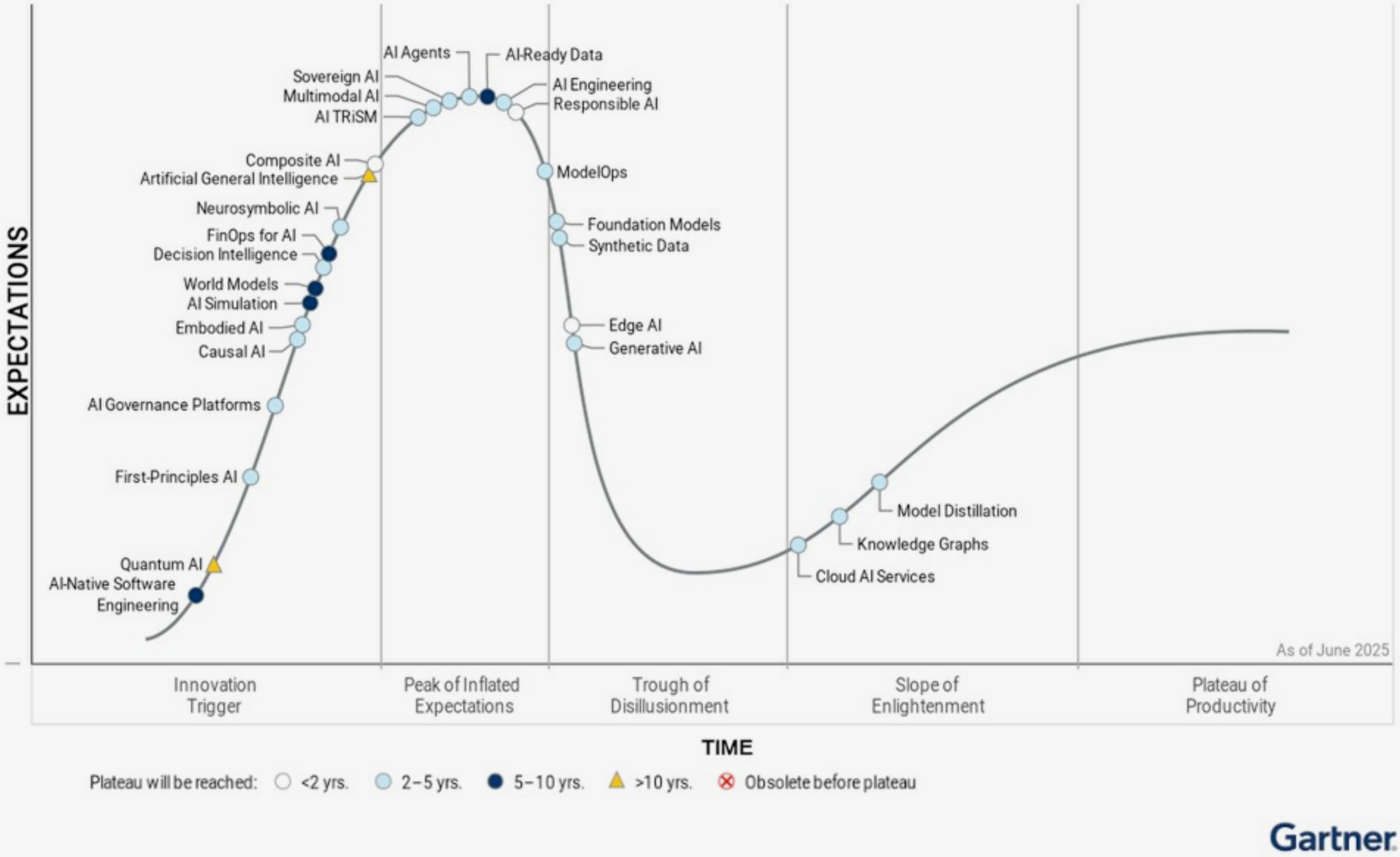
Hype Cycle for Artificial Intelligence, 2023



gartner.com

Source: Gartner

Gartner



EXPECTATIONS
AI Agents
AI-Ready Data
Sovereign AI
Multimodal AI
AI TRiSM
AI Engineering
Responsible AI
Composite AI
Artificial General Intelligence
Neurosymbolic AI
FinOps for AI
Decision Intelligence
World Models
AI Simulation
Embodied AI
Causal AI
AI Governance Platforms
First-Principles AI
Quantum AI
AI-Native Software Engineering
ModelOps
Foundation Models
Synthetic Data
Edge AI
Generative AI
Model Distillation
Knowledge Graphs
Cloud AI Services
As of June 2025
Innovation Trigger
Peak of Inflated Expectations
Trough of Disillusionment
Slope of Enlightenment
Plateau of Productivity
TIME
Plateau will be reached: <2 yrs. 2–5 yrs. 5–10 yrs. >10 yrs. Obsolete before plateau
Gartner

Generative AI

Generative AI

artificial intelligence capable of generating media

Generative AI

artificial intelligence capable of generating media

Training data → new data with similar characteristics

Generative AI

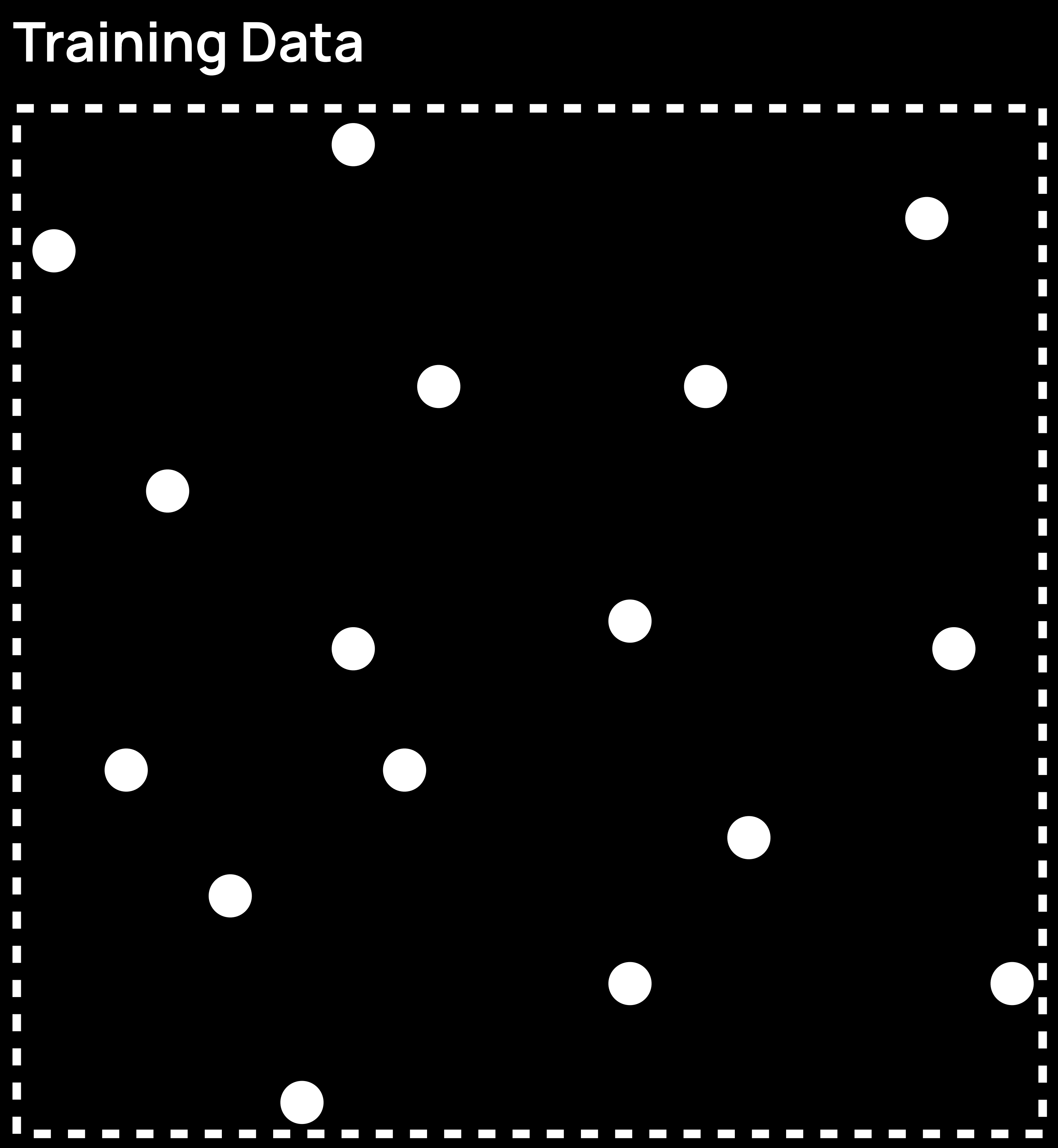
Tools that convert...

Text

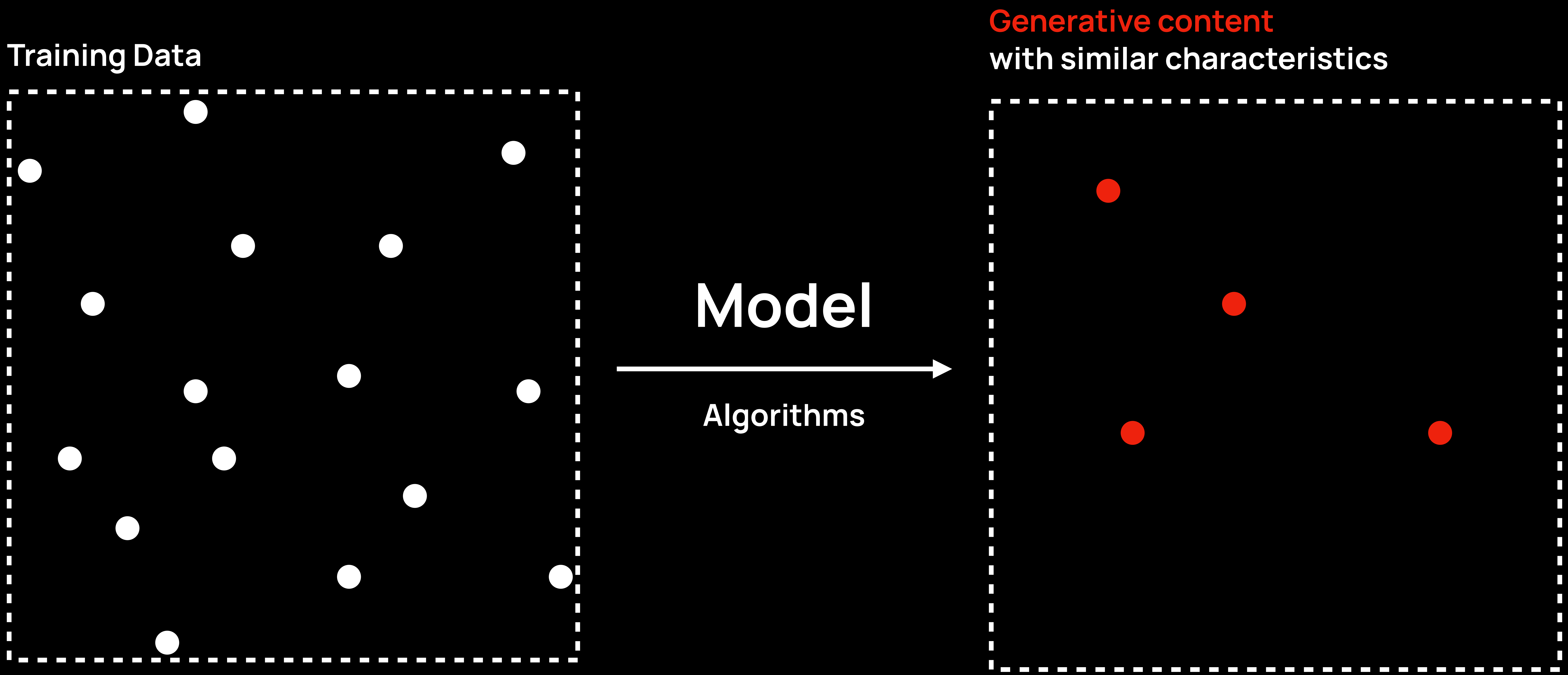
to

- Text
  - ChatGPT
  - Bard
- Speech
  - Eleven Multilingual
  - Meta SeamlessM4T
- Image
  - MidJourney
  - DALL-E
  - Stable Diffusion
- Video
  - Meta's Make-A-Video
  - Runway's Gen-2

Generative AI



Generative AI



Combination and recombination of patterns from the training data.

Generative AI

Tools that convert...

Text

to

Text

ChatGPT

Bard

Large language model

Text

to

Text

ChatGPT

Bard

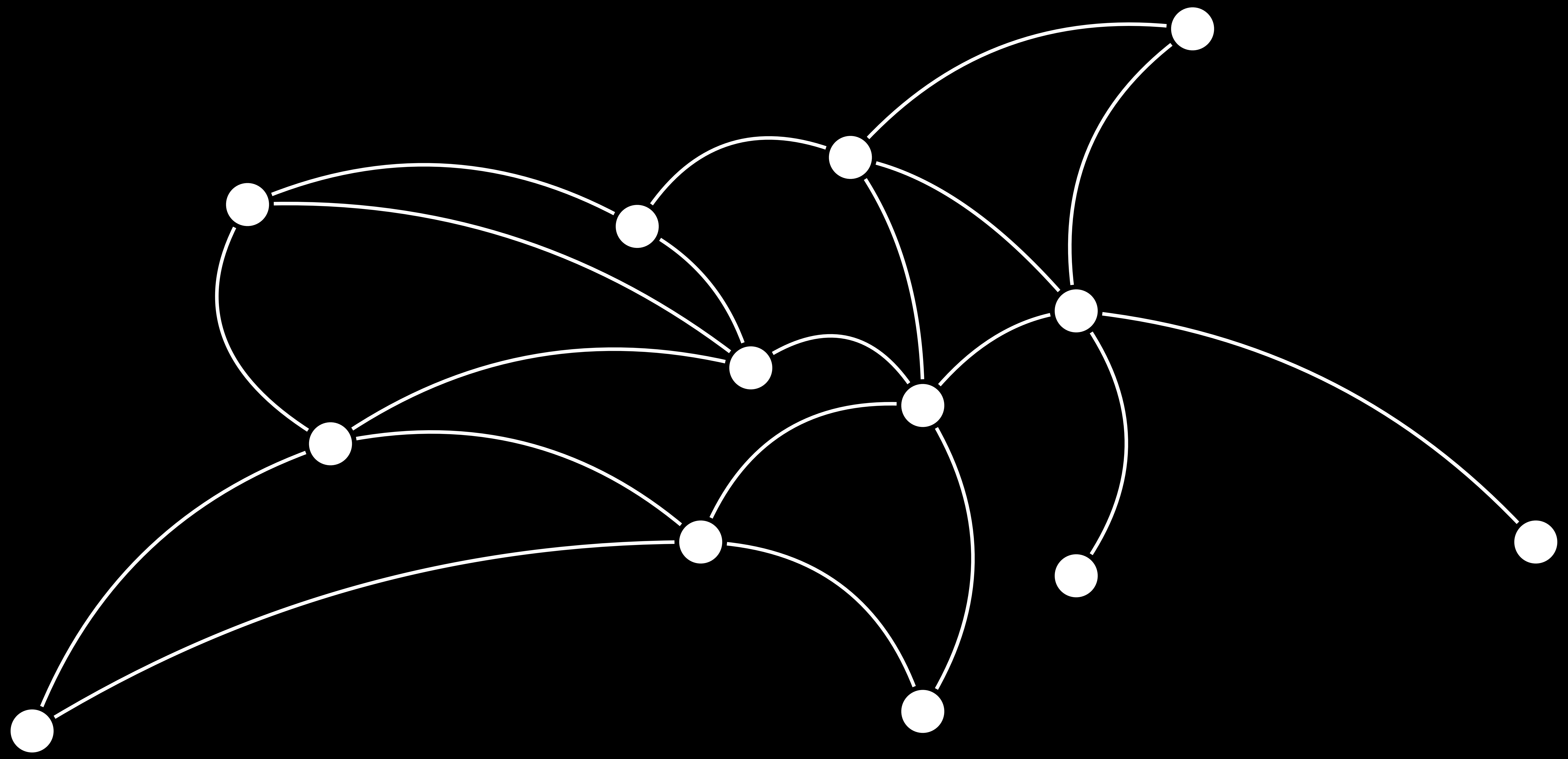
»Wo es nichts zu berechnen gibt, brauchen wir auch keine Computer.«

»Where there is nothing to calculate, we don’t need computers.«

Frieder Nake, Algorithmen & Zeichen

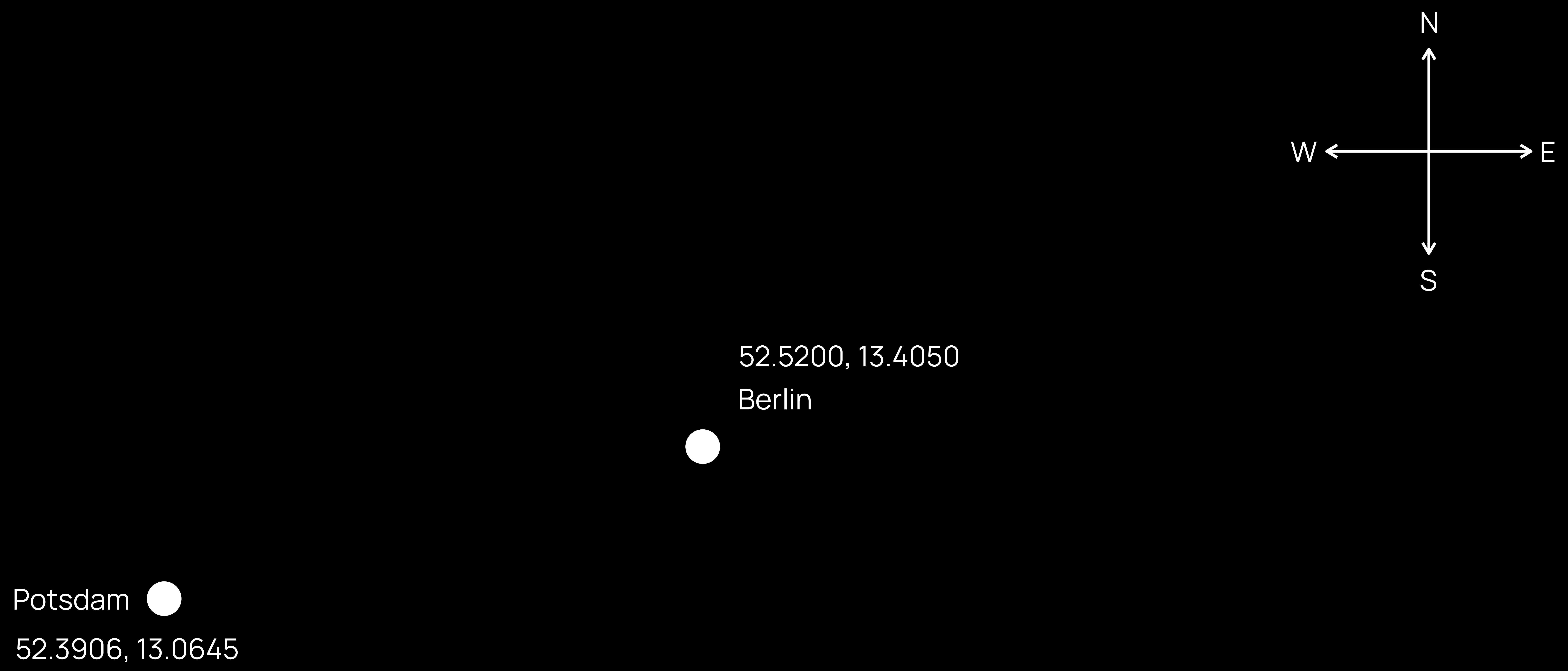
What is a large language model?

What is a large language model?



What is a large language model?

What is a large language model?



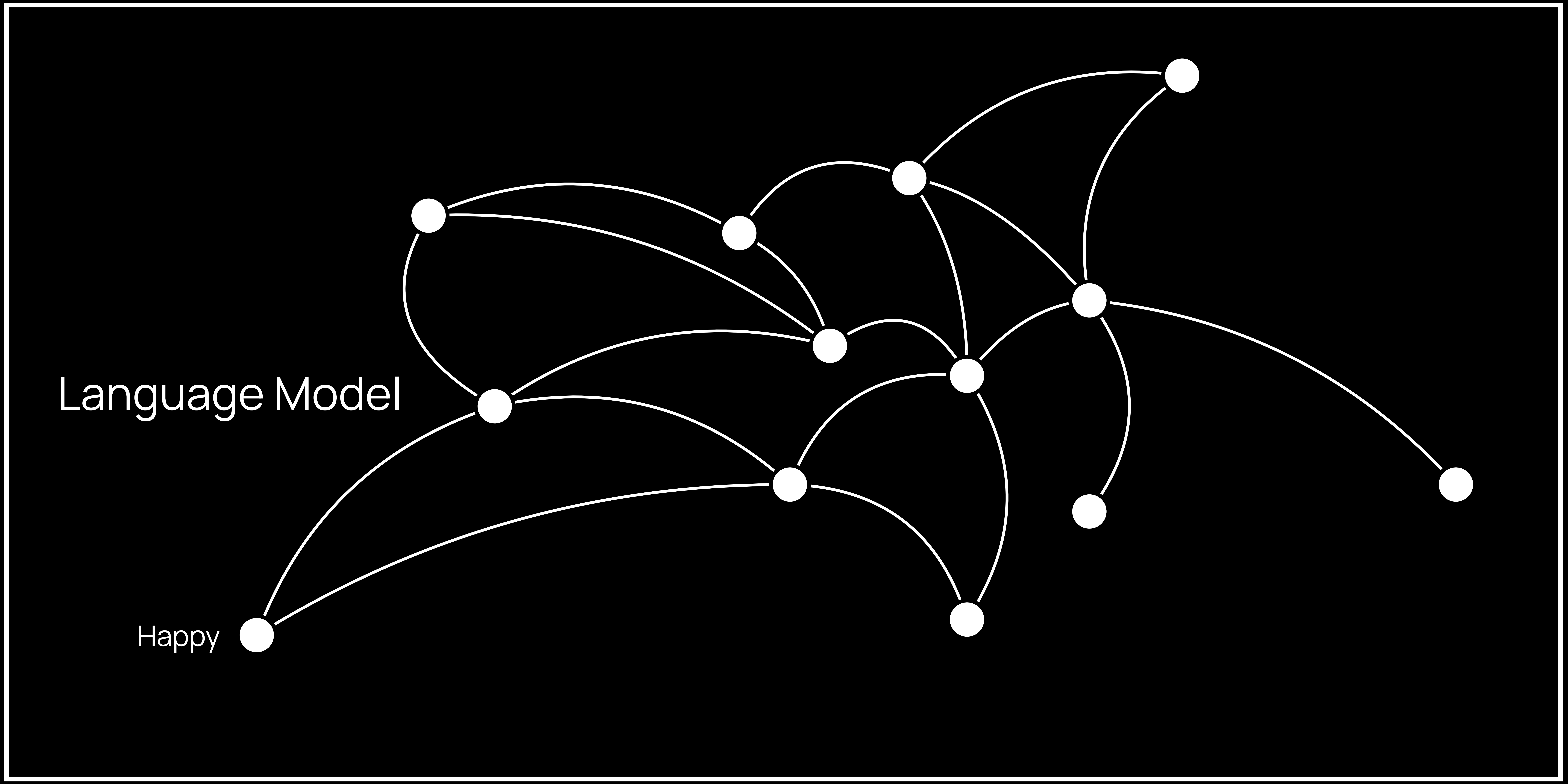
What is a large language model?

What is a large language model?

Happy

-0.24521, 0.13381, 0.28831, -0.1637, -0.049535, 0.037391, -0.52973, 0.05099, -0.42168, -0.40447, -0.79435, -0.088239, 0.2763, -0.028826,
-0.30289, -0.51224, 0.41553, 0.17285, 0.47578, 0.46862, 0.080201, -0.10659, -0.74806, -0.51164, -0.33604, 0.17711, -0.089794, -0.11651, 0.211
0.16158, 0.10607, -0.14914, 0.27409, 0.30702, -0.28446, -0.15515, -0.23536, 0.19899, -0.77775, 0.83599, 0.55016, -1.0701, 0.30365, -0.75691,
-0.2221, 0.12196, 0.18849, -0.098025, 0.93925, -0.038423, -0.23047, 0.27205, -0.31529, -0.02843, -0.016746, -0.021209, 0.079952, 0.31622
0.064761, -0.23652, -0.01021, -0.55726, 0.038226, -0.17971, -0.072388, 0.14231, 0.94446, -0.36723, -0.63264, -0.22749, 0.47908, -0.66815,
-0.041302, -0.41019, 0.14565, 0.12786, -0.62428, 0.09539, -0.20977, -0.24609, -0.40151, 0.47429, 0.010771, 0.82454, 0.33783, 0.12811, 0.2199
0.38581, 0.01102, 0.31732, 0.29965, 0.03711, -0.025396, 0.48878, -0.065962, -0.12585, -0.67172, 0.71344, -0.52369, -0.68097, 0.43216, 0.07
-0.11568, -0.38967, -0.21208, -0.27753, 0.30311, -0.51805, 0.28916, -0.26186, -0.009617, -0.40885, 0.21191, -0.052608, -0.1316, 0.96776, -0.2
0.47056, -0.31982, 0.095621, -0.32895, -0.55728, 0.32756, 0.86118, 0.33973, 0.14936, -1.0919, 0.28923, -0.36886, -0.052346, -0.13595,
-0.049021, -0.39374, -0.17264, -0.099287, 0.24752, 0.48518, -0.83403, 0.072445, -0.47082, 0.60437, 1.0859, 0.57208, 0.1354, 0.028308, 0.0
0.036348, -0.22297, 0.064423, 0.159, -0.77674, 0.30433, -0.04432, -0.34273, -0.34202, 0.24068, 0.3029, 0.13857, 0.5047, 0.57522, 0.17839,
0.41169, -0.33658, 0.18725, -0.46583, 0.13571, 0.23891, 0.83214, -0.13962, -0.16713, -0.31497, -0.37804, 1.0744, 0.42999, -0.80921, 0.88273,
0.94068, 0.56096, -0.075746, -0.14008, -0.3128, 0.45362, -0.17636, 0.25203, 0.45339, 0.006271, 0.79116, -1.1023, 0.37591, 0.44677, 0.04936
-0.39339, -0.72904, -0.47843, 0.029577, -0.59435, 0.13464, -0.51627, 0.837, -0.58194, 0.046615, -0.3937, 0.66203, 0.48482, -0.17065, 0.332
0.12285, -0.045335, -0.27725, -0.001421, -0.058123, 0.63951, 0.77633, 0.35478, 0.14914, -0.65503, 0.52402, -0.53852, -0.37985, 0.37724, 0.
0.21311, 0.049646, 0.16984, -0.33992, 0.55786, 0.38667, -0.84554, 0.28118, 1.1596, -0.70108, -0.74602, -0.23631, 0.22675, -0.46052, 0.2394
0.1382, -0.42704, -0.89916, -0.017155, -0.18568, -0.51684, 0.56688, -0.051546, -0.69499, 0.017353, 0.57945, -0.027723, -0.012037, -0.18122
-0.40603, -0.43229, -0.27754, 0.001625, -1.0812, -0.85345, -0.26496, -0.45327, -0.51637, 0.49035, -0.27177, -0.59652, 0.97047, -0.43425,
-0.16713, -0.012096, 0.12838, 0.9593, -0.25205, 0.25473, -0.38155, -0.10364, 0.50817, 0.74453, 0.2516, -0.52587, 0.41403, 0.70447, 0.56725,
-0.76822, 0.38446, 0.65434, -0.4701, 0.11068, 0.84786, -0.50158, -0.73509, 0.028486, -0.39405, -0.19661, -0.49297, 0.21677, 0.46745, -0.87
0.70179, 0.16544, 0.26561, 0.46307, -0.34402, 0.34357, -0.97758, 0.42855, -0.088987, -0.50831, 0.73807, 0.66019, 0.33652, -0.14058, -4.827
-0.17123, 0.28438, -0.69702, 0.052258, -0.28428, 0.28932, -0.14343, 0.10838, -0.073325, 0.36498, -0.1145, 0.51698, 0.039328, 0.54644, -0.4
-0.038352, 0.067462, -0.44783, -0.009301, -0.56002, -0.32046, 0.10558, 0.00215, -0.29051, 0.032734, 0.53683, 0.78655, -1.5857, 0.062173,
-0.29321, 0.14442, 0.26699, 0.92783, 0.43608, 0.08423, 0.19892, -0.63507, 0.37158, 0.82915, 0.51322, -0.09333, -0.18363, 0.17964, 0.18824,

What is a large language model?



What is a large language model?

Language Model

What is a large language model?

Given a sequence of words, produce the probability distribution of the next word.

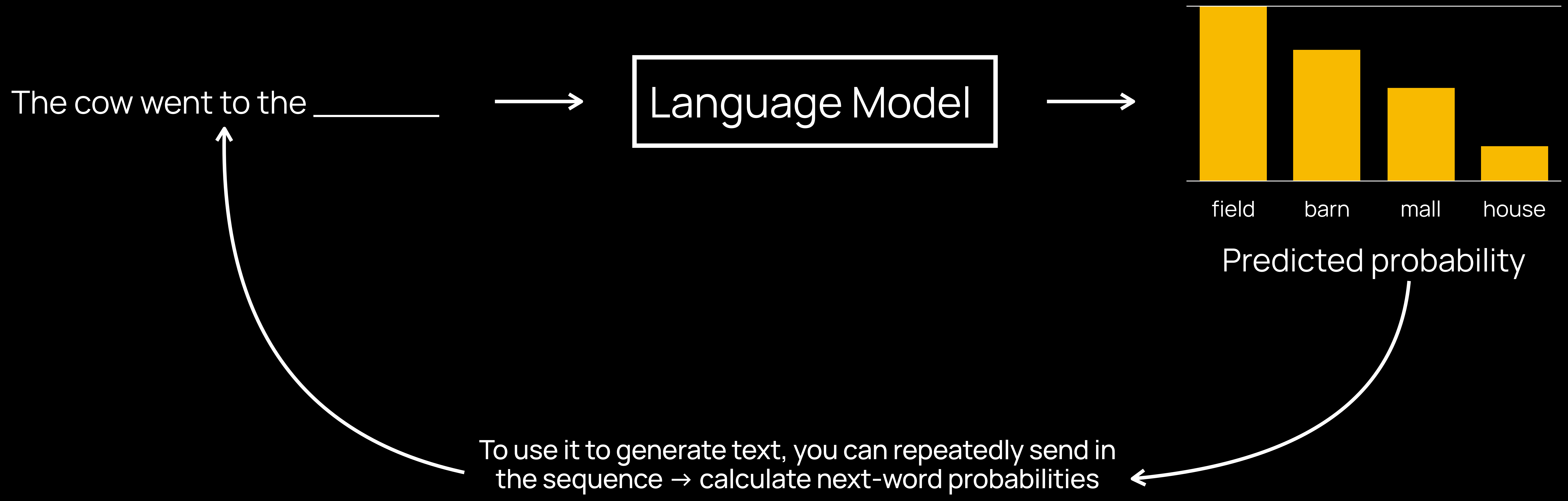
The cow went to the ________ $\longrightarrow$ $\boxed{\text{Language Model}}$

What is a large language model?

Given a sequence of words, produce the probability distribution of the next word.

What is a large language model?

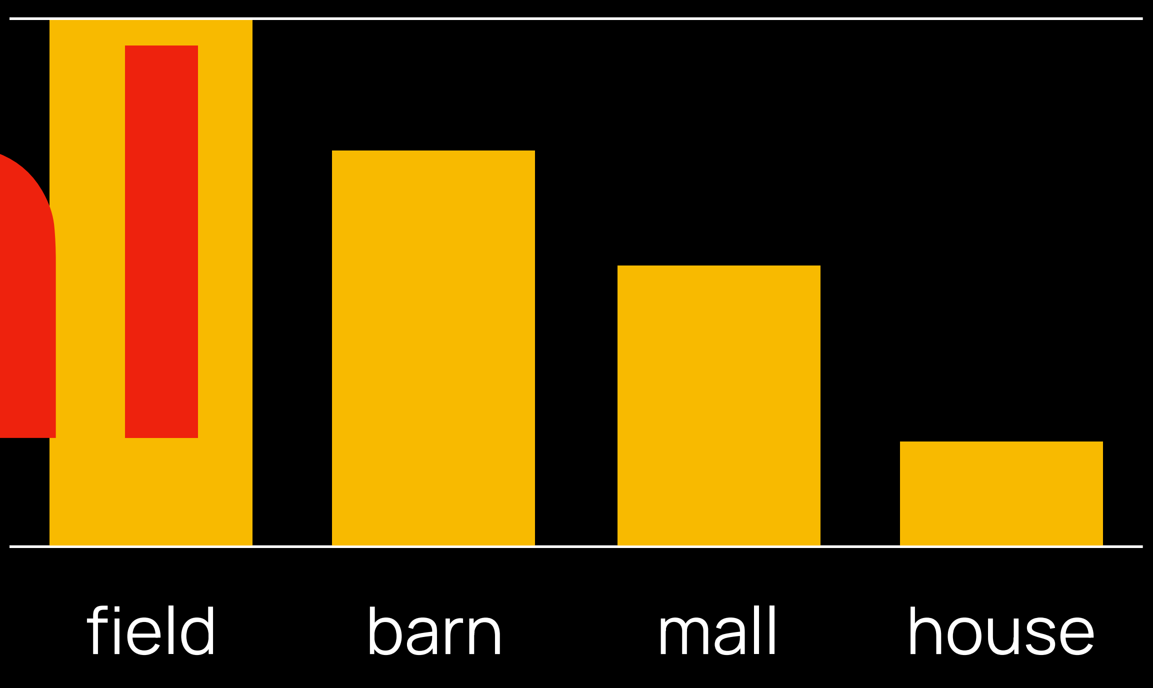
Given a sequence of words, produce the probability distribution of the next word.



What is a large language model?

Given a sequence of words, produce the probability distribution of the next word.

The cow went to the ____ → Language Model →



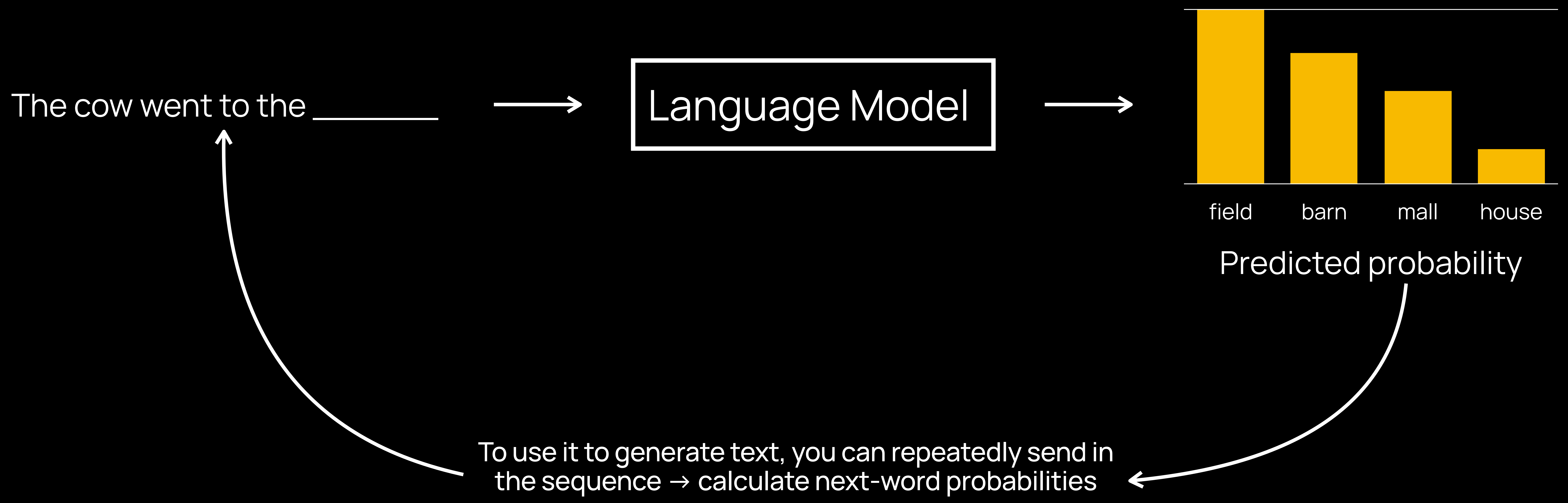
Predicted probability

Statistical Average

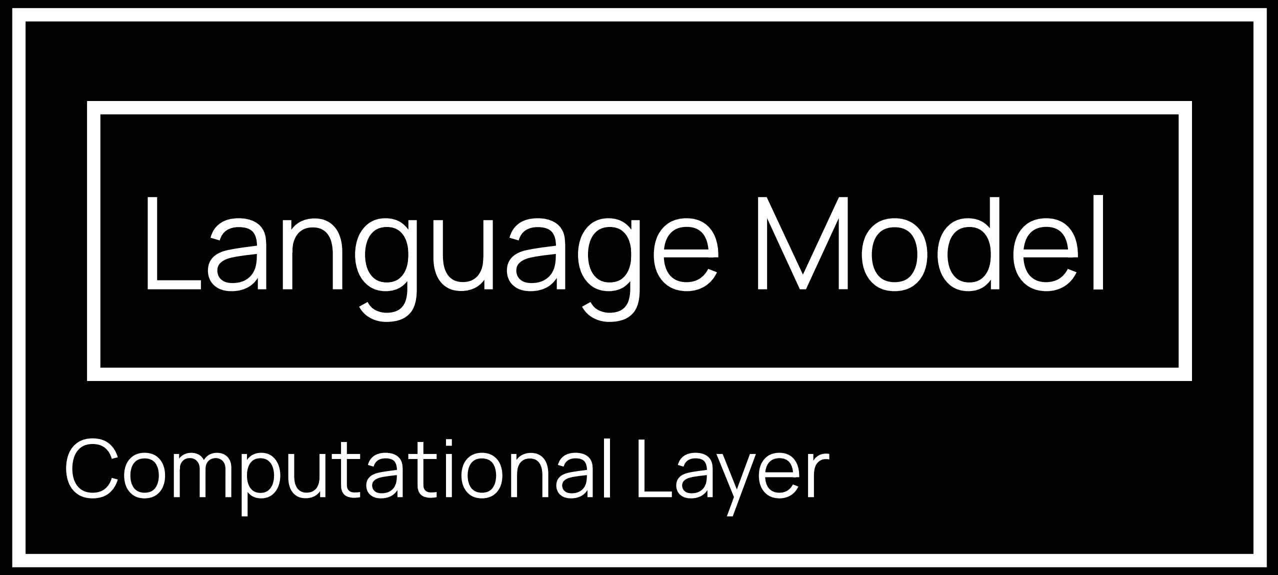
To use it to generate text, you can repeatedly send in the sequence → calculate next-word probabilities

What is a large language model?

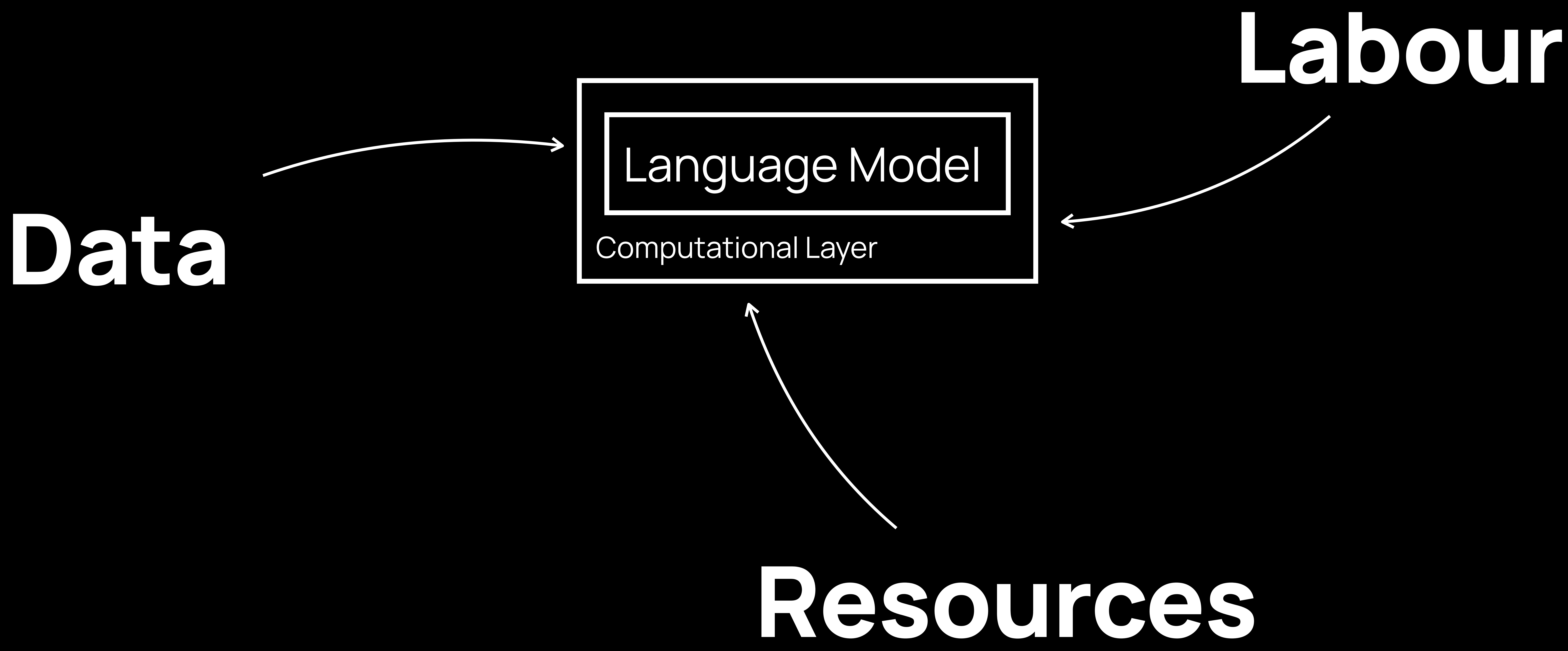
Given a sequence of words, produce the probability distribution of the next word.



What is a large language model?



What is a large language model?



What is a large language model?

Data

Wikipedia

3

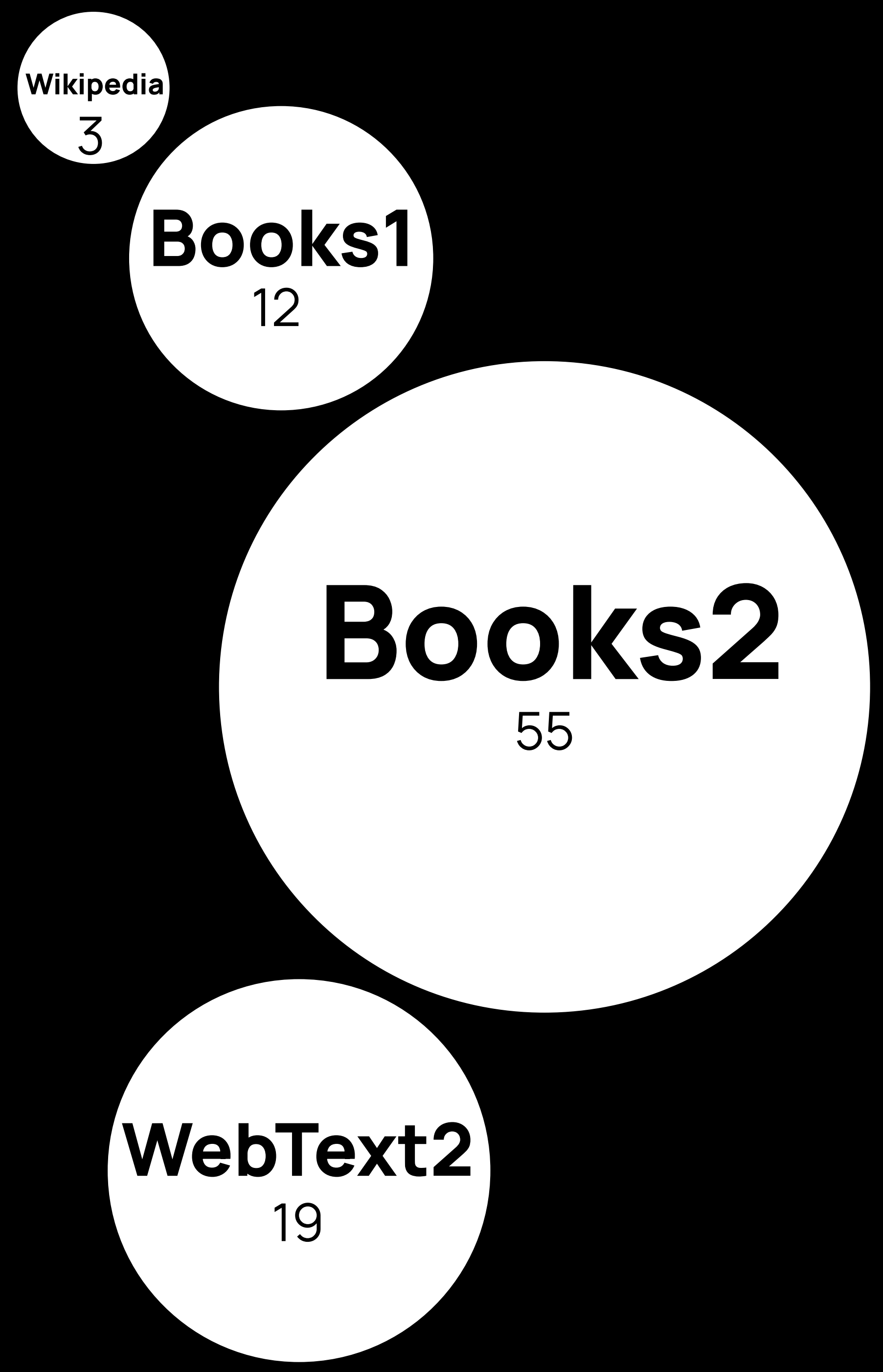
What is a large language model?

Data

Wikipedia
3

What is a large language model?

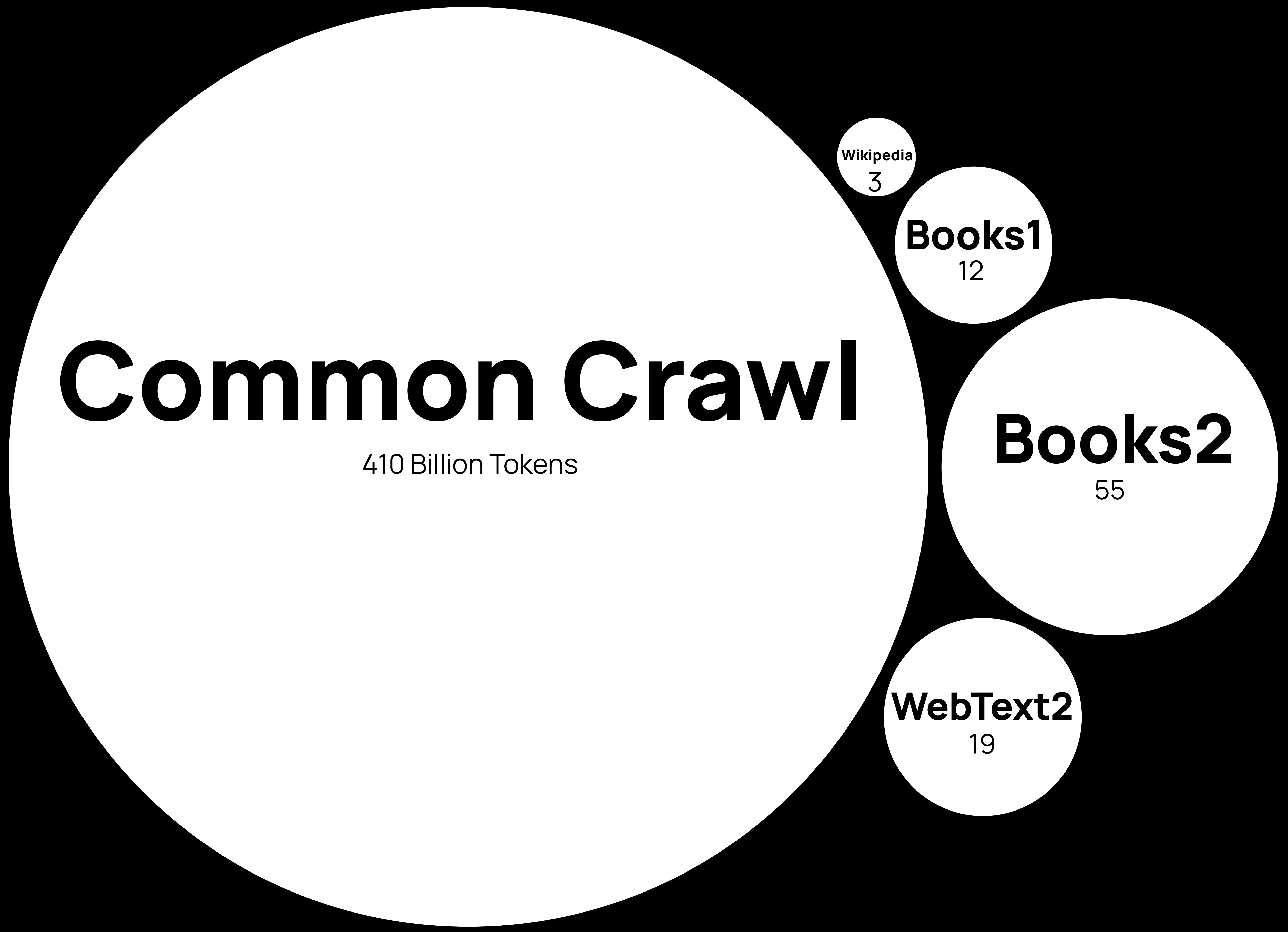
Data



Reference: Language Models are Few-Shot Learners arxiv.org/pdf/2005.14165.pdf

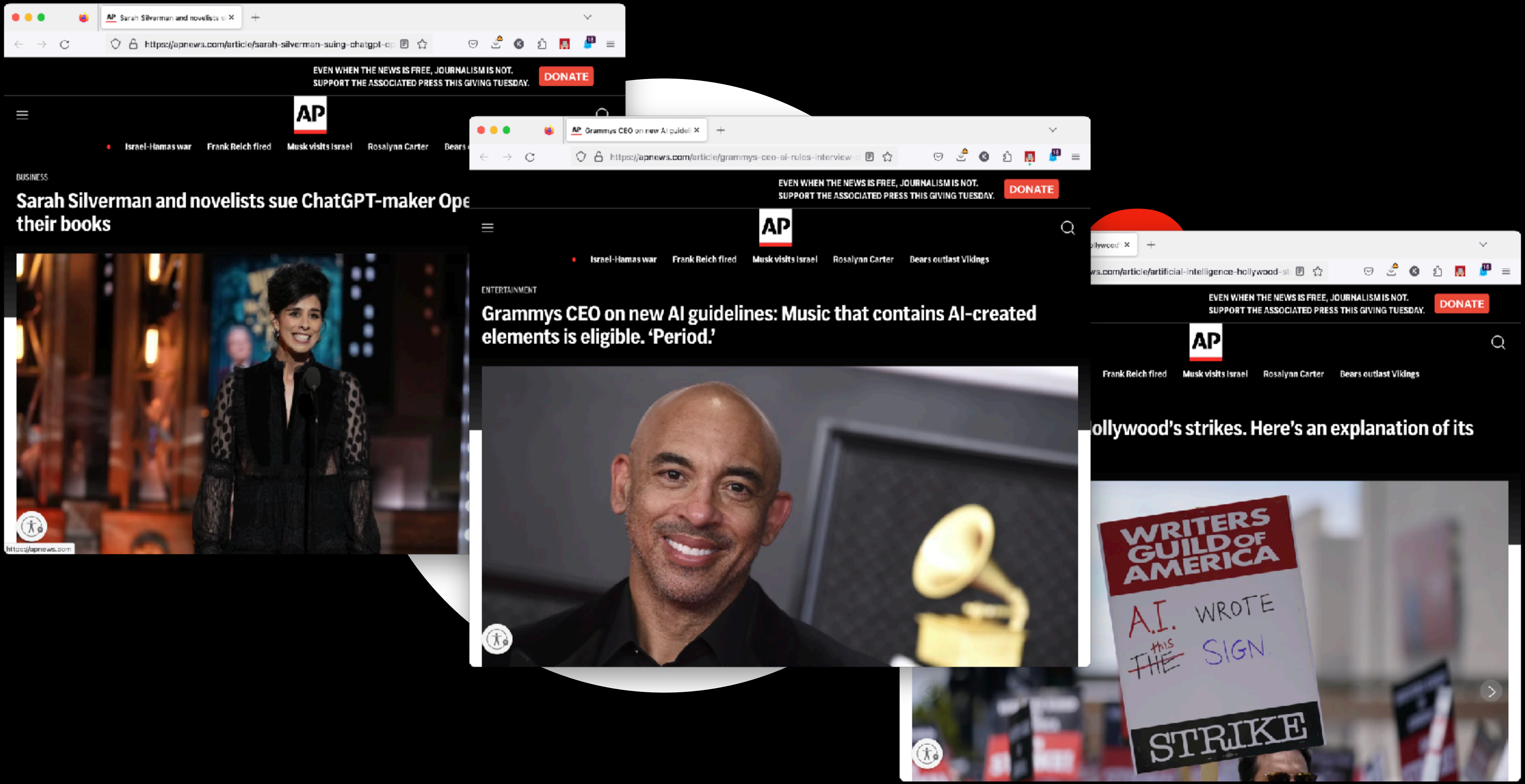
What is a large language model?

Data



Reference: Language Models are Few-Shot Learners arxiv.org/pdf/2005.14165.pdf

What is a large language model?



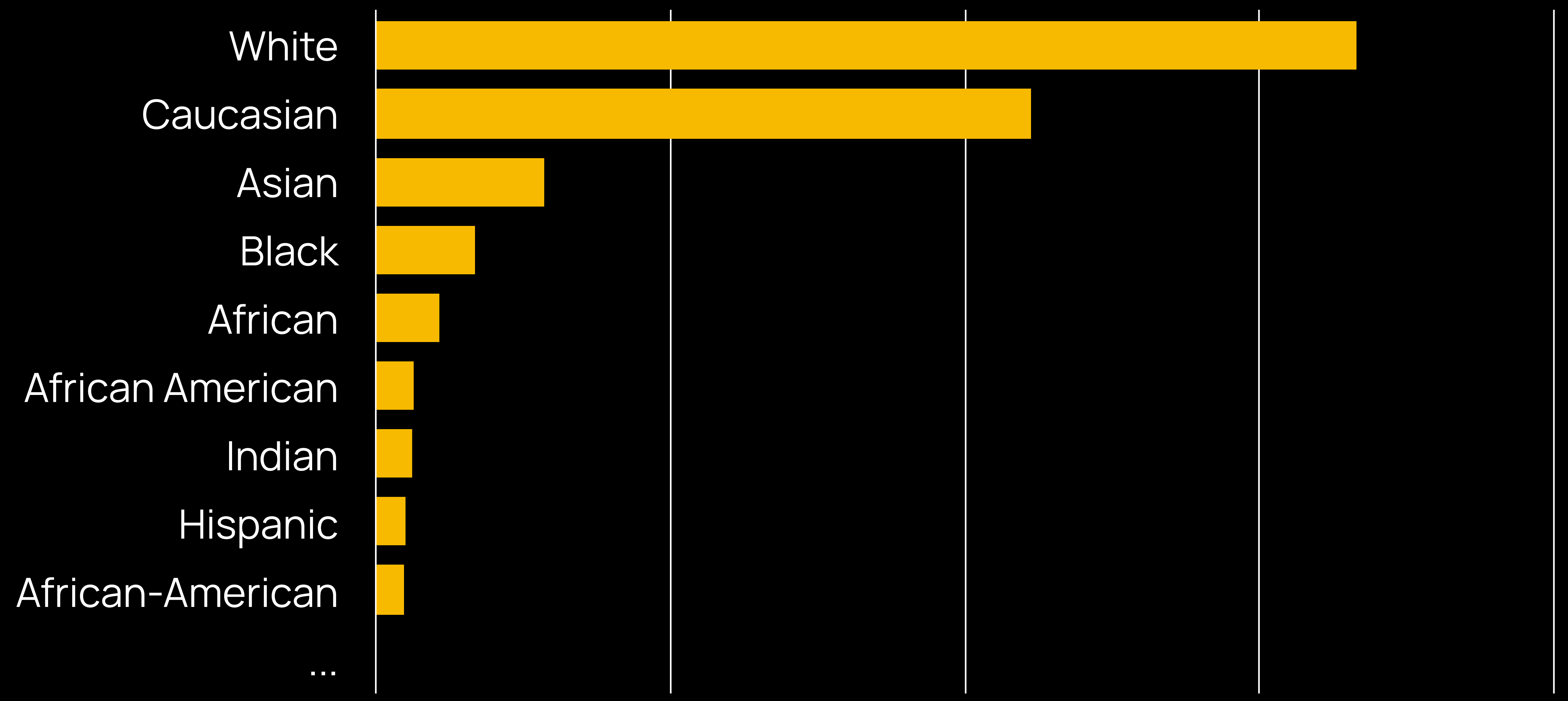
Reference: Language Models are Few-Shot Learners arxiv.org/pdf/2005.14165.pdf

What is a large language model?

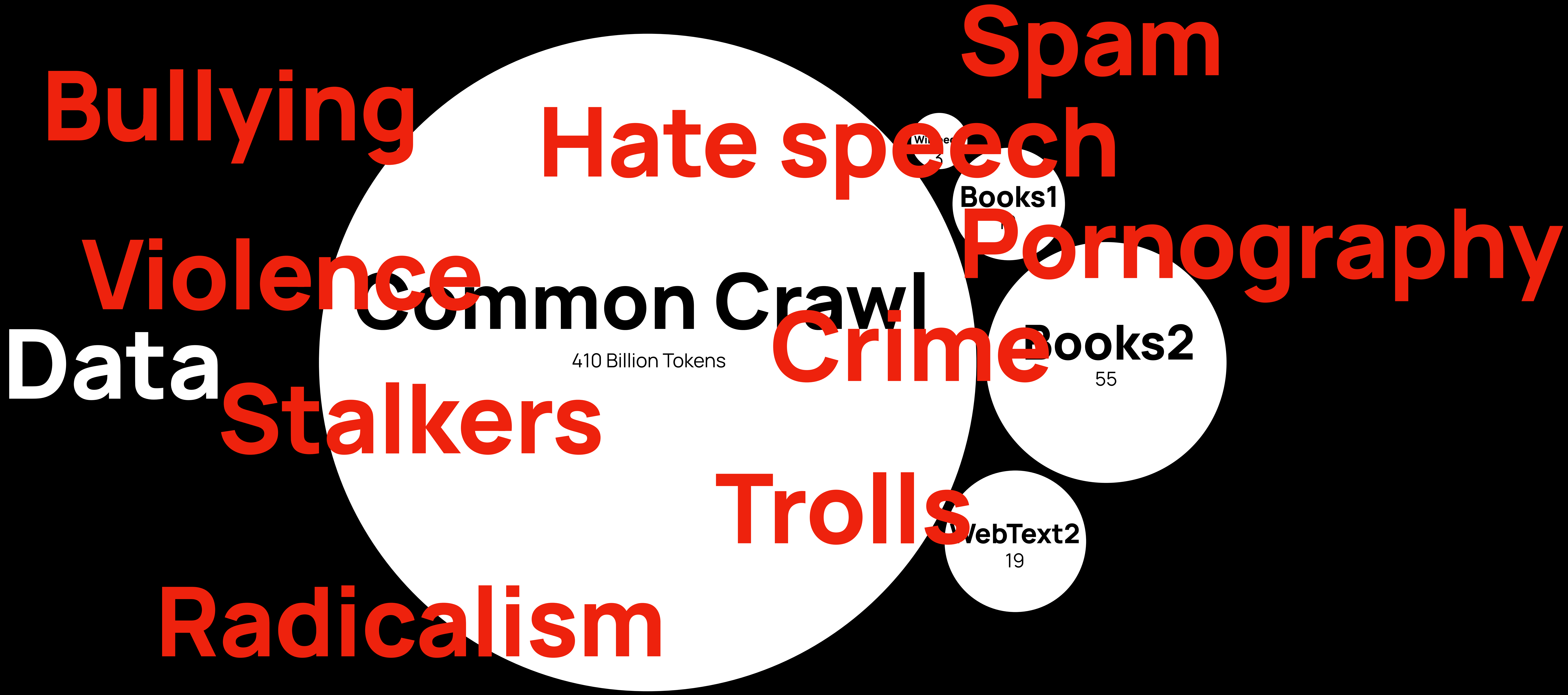
Historical Baggage

Data

Reference: Language Models are Few-Shot Learners arxiv.org/pdf/2005.14165.pdf


White
Caucasian
Asian
Black
African
African American
Indian
Hispanic
African-American
...

What is a large language model?

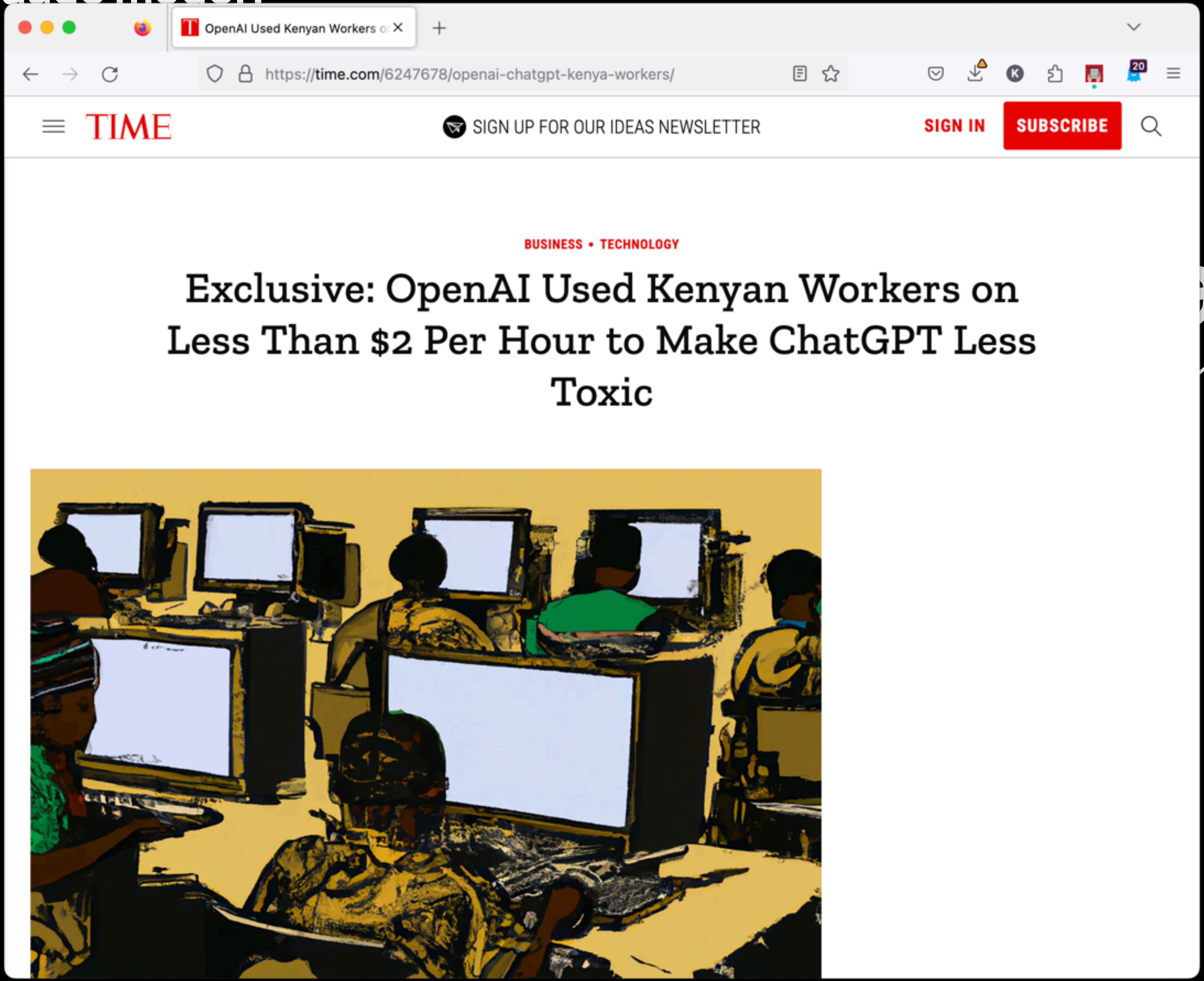


Reference: https://bsky.app/profile/dieinsider.bsky.social

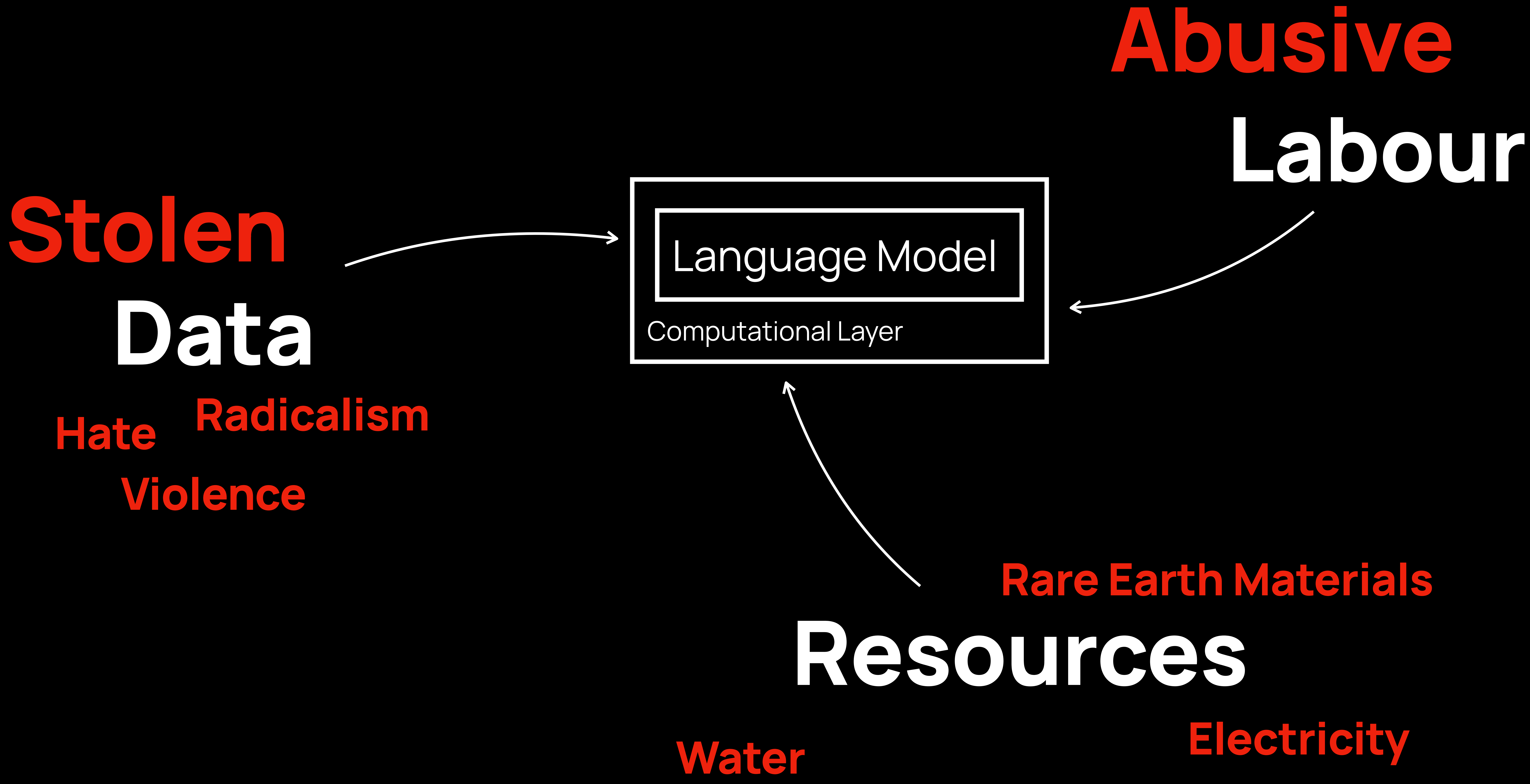
What is a large language model?

Data

abour




yahoo!finance
ChatGPT's Surprising Envir
Editor OilPrice.com
September 18, 2023 · 2 min read
In this article:
MSFT +0.50%
Climate alarmists have wa
intensive cryptocurrency
have yet to denounce larg
(LLMs) like Microsoft-back
that use 16 ounces of fres
user asks it a series of que
Making AI Less "Thirsty": Uncover
Secret Water Footprint o
Pengfei Li UC Riverside
Jianyi Yang UC Riverside
Abstract
The growing carbon footprint of artificial intelligence (AI) m
has been undergoing public scrutiny. Unfortunately, however, t
(withdrawal and consumption) footprint of AI models has remai
GPT-3 in Microsoft's state-of-the-art U.S. data centers can directl
ter, but such information has been kept a secret. More critically,
for 4.2 – 6.6 billion cubic meters of water withdrawal in 2027, whic
drawal of 4 – 6 Denmark or half of the United Kingdom. This is
become one of the most pressing challenges shared by all of us i
tion, depleting water resources, and aging water infrastructures.
AI models can, and also must, take social responsibility and lead
footprint. In this paper, we provide a principled methodology to
and also discuss the unique spatial-temporal diversities of AI m
highlight the necessity of holistically addressing water footprint
sustainable AI.
1 Introduction
• "Water is a finite resource, and every drop matters." — Faceboo
• "Fresh, clean water is one of the most precious resources on E
support water security and healthy ecosystems." — Google's Wa
• "Water is a human right and the common development denon
is in deep trouble." — U.N. Secretary-General António Guterres at
• "Historic droughts threaten our supply of water ... As the s
security is central to human and national security." — U.S. Whit
2022 [4].
Artificial intelligence (AI) models have witnessed remarkabl
areas of critical importance to our society over the last decade,
several global challenges such as climate change [5]. Increasingly
on power-hungry servers housed inside warehouse-scale data
hogs [6]. Consequently, despite the numerous benefits and pote
AI models, in particular carbon footprint, has been undergoing p
arXiv:2304.03271v3 [cs.LG] 29 Oct 2023
SCI AM
Sign Up for Our Daily Newsletter
OPINION
MAY 25, 2023 | 4 MIN READ
A Computer Scientist Breaks Down Generative AI's Hefty Carbon Footprint
Is generative AI bad for the environment? A computer scientist explains the carbon footprint of ChatGPT and its cousins—and how to reduce it
BY THE CONVERSATION US & KATE SAENKO


Abusive
Labour
Stolen
Data
Language Model
Computational Layer
Hate
Radicalism
Violence
Rare Earth Materials
Resources
Water
Electricity

Average culture based on stolen data.

*Based on Hito Steyerl notion of mean images
https://criticalai.art/hito.html

Culture and creativity are more than a statistical average of past stolen data.

»And due to the capital required to build AI at scale and the ways of seeing that it optimizes AI systems are ultimately designed to serve existing dominant interests. In this sense, artificial intelligence is a registry of power«

Atlas of AI by Kate Crawford, 2021

»And due to the capital required to build AI at scale and the ways of seeing that it optimizes AI systems are ultimately designed to serve existing dominant interests. In this sense, artificial intelligence is a registry of power«

Atlas of AI by Kate Crawford, 2021

Reclaiming power

“Big Dick Data is a formal, academic term that we, the authors, have coined to denote big data projects that are characterized by patriarchial, cis-masculinist, totalizing fantasies of world domination as enacted through data capture and analysis.”

Data Feminism, Catherine D'Ignazio and Lauren Klein, 2022
https://data-feminism.mitpress.mit.edu/pub/czq9dfs5/release/3

“Building one’s own small-scale datasets for creative ML can be seen as one way of subverting the power structures that reign AI development.”

Vigliensoni, Perry, and Fiebrink, A Small-Data Mindset for Generative AI Creative Work, 2022

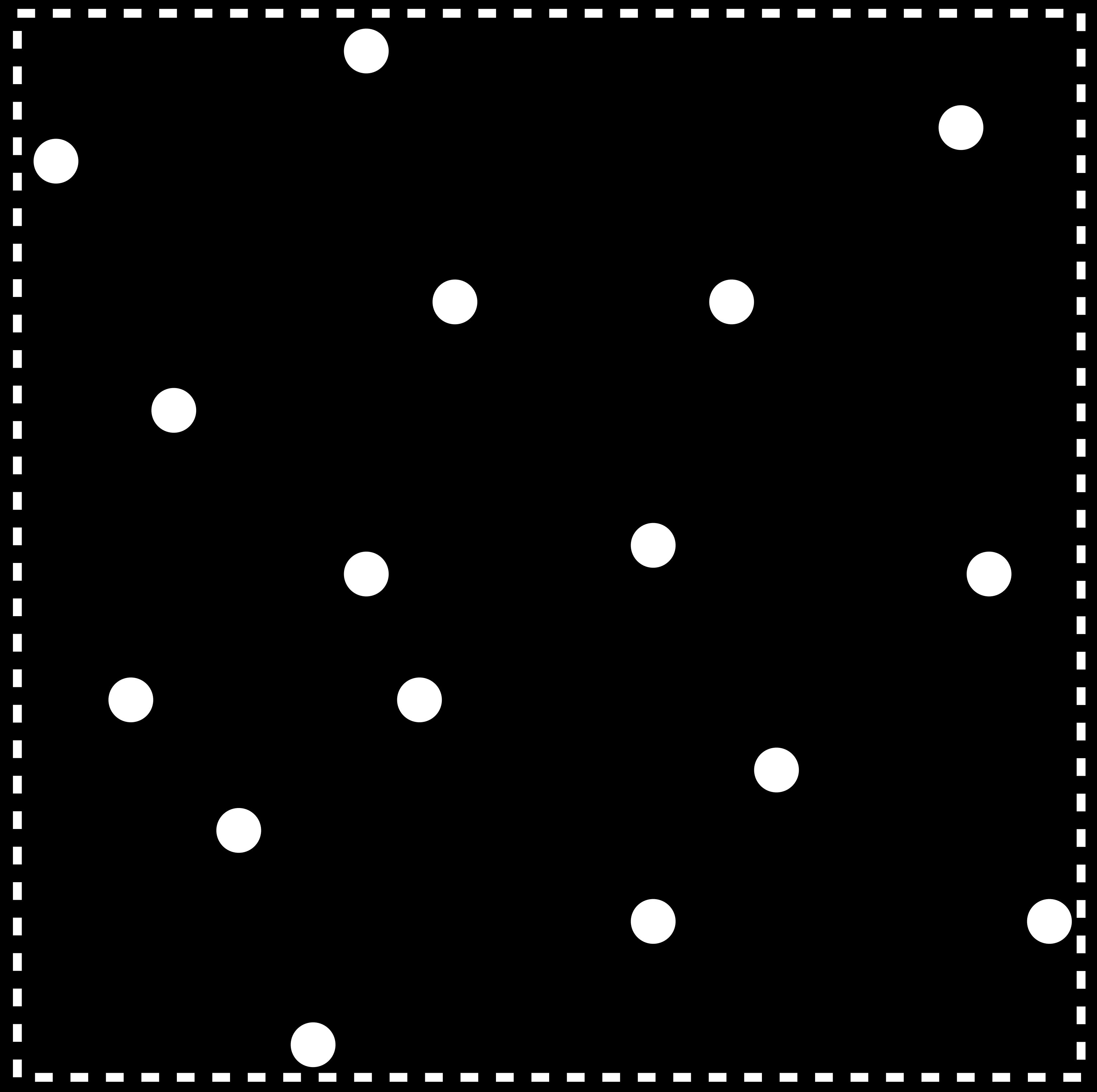
“Personal, bespoke data can be a powerful way to make useful models (sometimes more powerful than big, general-purpose data!)”

Rebecca Fiebrink, xCoAx Conference V&A Museum Dunde 2025

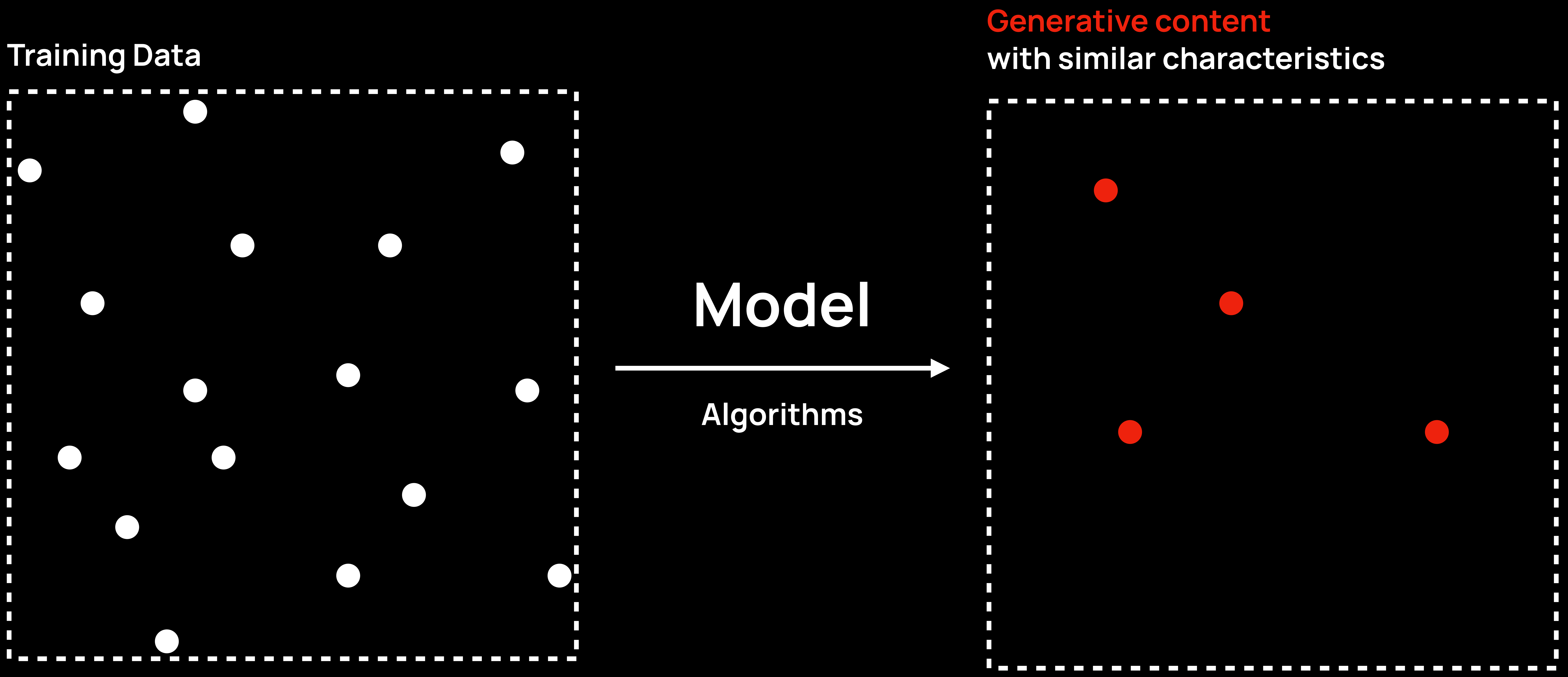
Models to Explore

Generative AI

Training Data



Generative AI



Combination and recombination of patterns from the training data.

Models

Classifier

Generative adversarial network

FastGAN

Pix2Pix (conditonal GAN)

Fine Tuning

Models & Histories

1950s–1980s
(Perceptron, Decision Trees)

AlexNet wins ImageNet (2012), showing deep neural nets outperform classical classifiers.

Classifier

Generative adversarial network

FastGAN

Pix2Pix
(conditonal GAN)

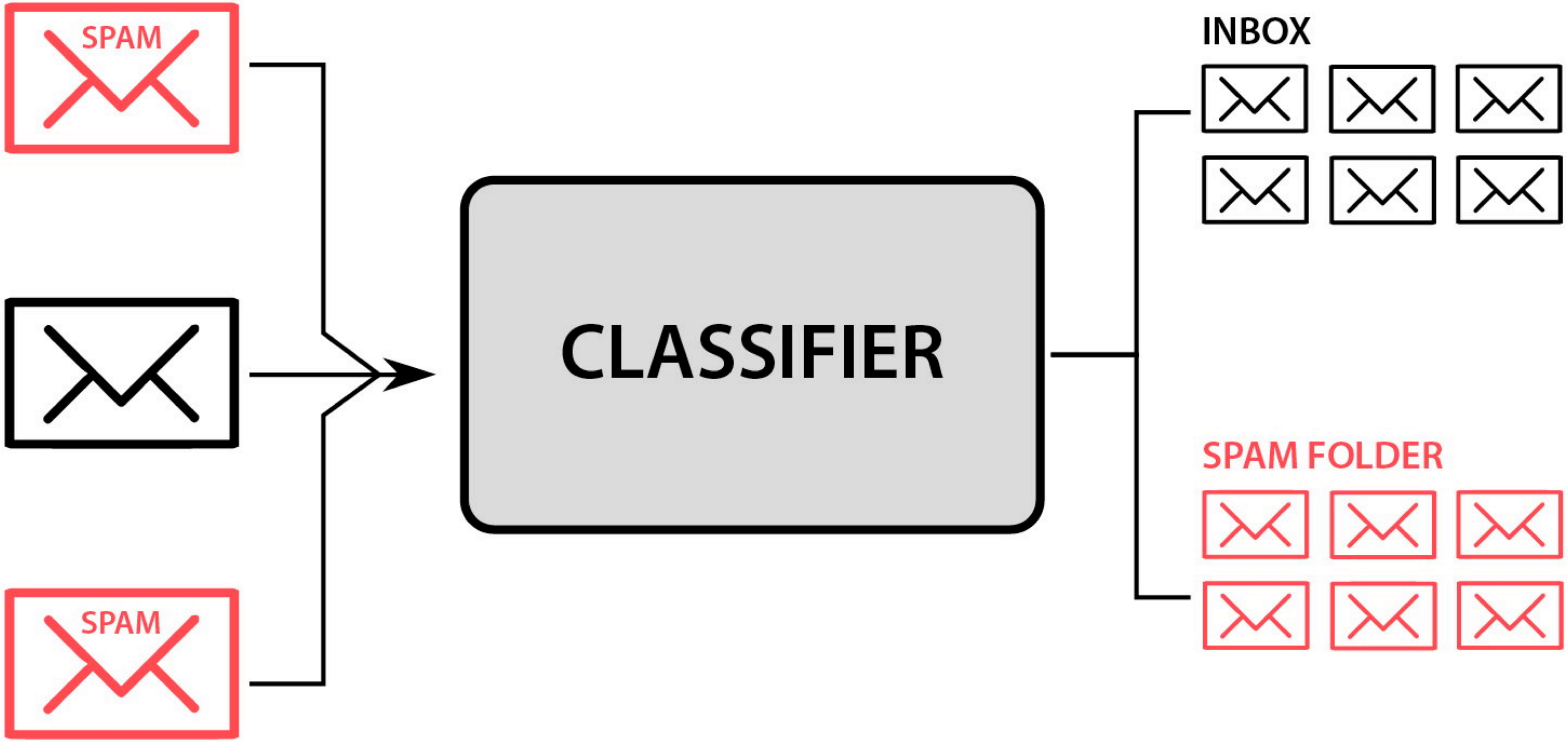
2014 (Goodfellow et al.)
DCGAN (2015) → Pix2Pix (2016) → StyleGAN (2018) → fuels art/A.I. boom.

Fine Tuning

2006 (early neural nets)
→ 2018+ (Transfer Learning)

Fine-tuning BERT (2018) & GPT-2 (2019) becomes standard for adapting large models to specific tasks. Artists start fine-tuning StyleGAN and GPT models.

Classifier



https://www.analytixlabs.co.in/blog/classification-in-machine-learning/

Input

Text
Images
Sound
Video
...

Model

Classifier

Output

Category A
Category B
Category C
...

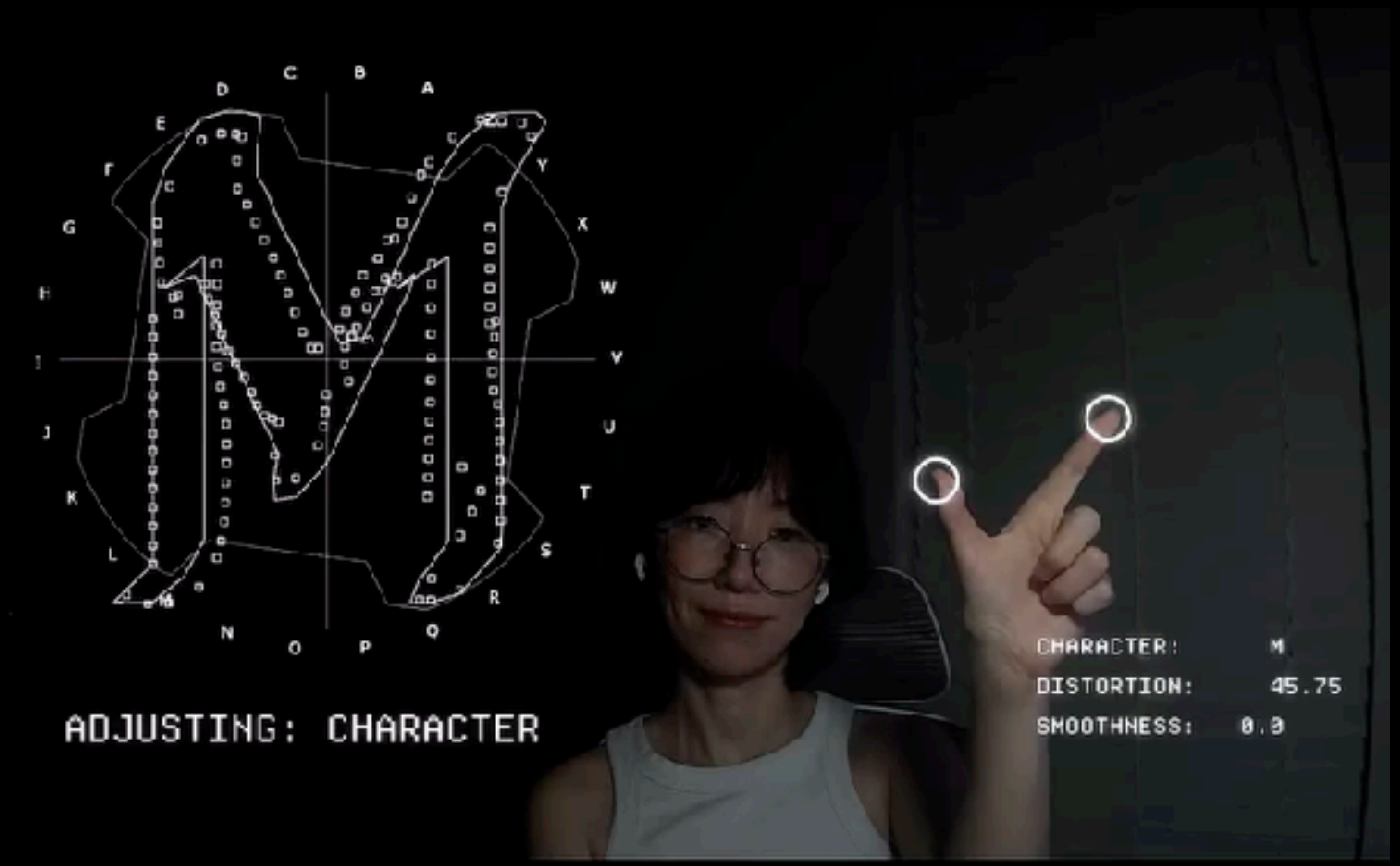

Chrome File Edit View History Bookmarks People Window Help
Teachable Machine
Secure https://gweb-teachable-ai.appspot.com/
LEARNING
OUTPUT
GIF
Sound
Speech
112 EXAMPLES
CONFIDENCE
89%
TRAIN GREEN
INPUT
119 EXAMPLES
CONFIDENCE
TRAIN PURPLE
0 EXAMPLES
CONFIDENCE
TRAIN ORANGE

Poet Engineer

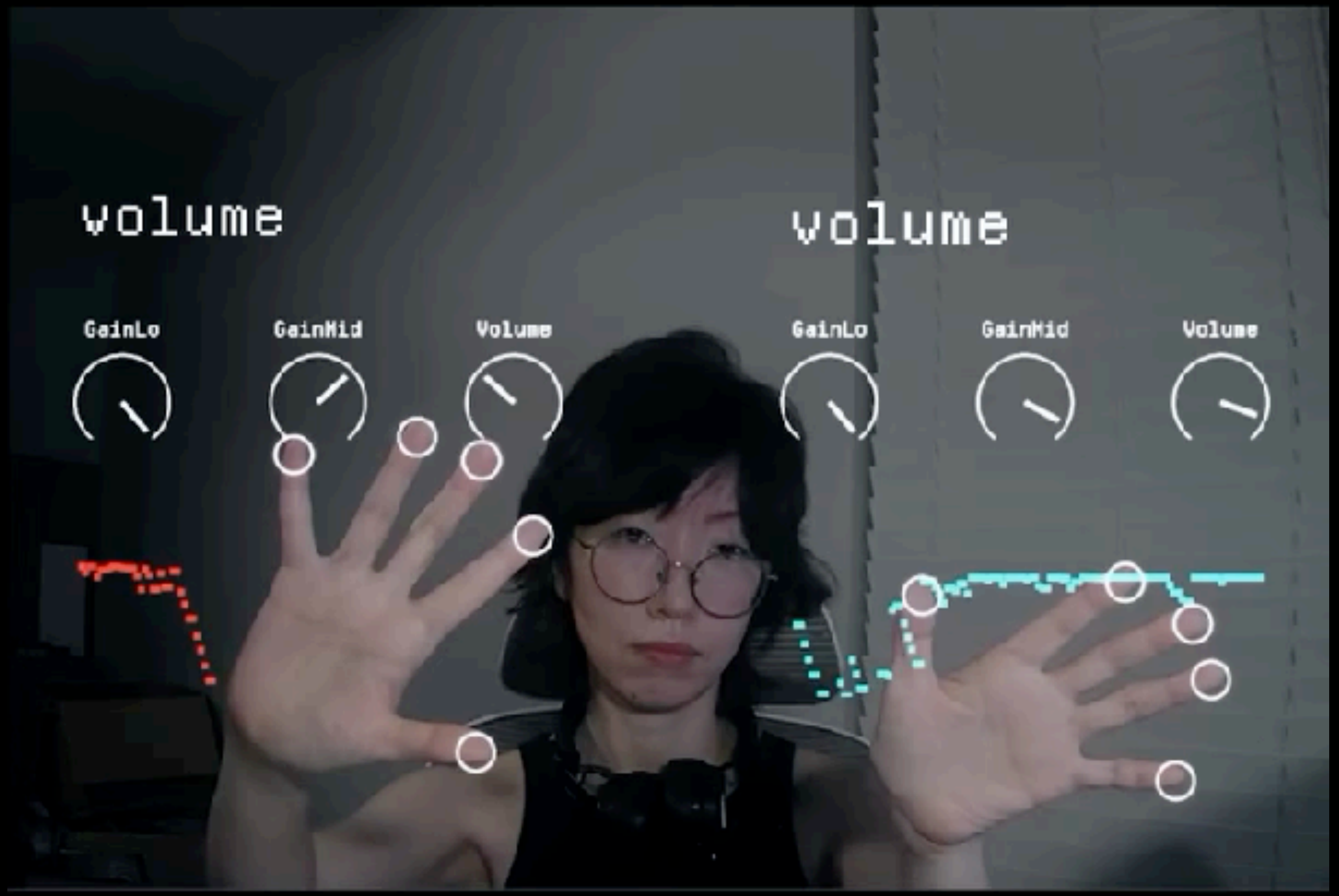
https://www.instagram.com/the.poet.engineer/


expand

i


ADJUSTING: CHARACTER
CHARACTER: M
DISTORTION: 45.75
SMOOTHNESS: 0.0

learning live dj mixing
with gestures
🎛️🎧🤗



UNIDS, Trevor Paglen

https://paglen.studio/2023/05/10/unids/

“... a project to photograph objects of unknown origin in orbit around the earth.”

“The term “unid” is a term that amateur astronomers created to describe objects that they have observed in orbit, but whose identity they have failed to establish.”

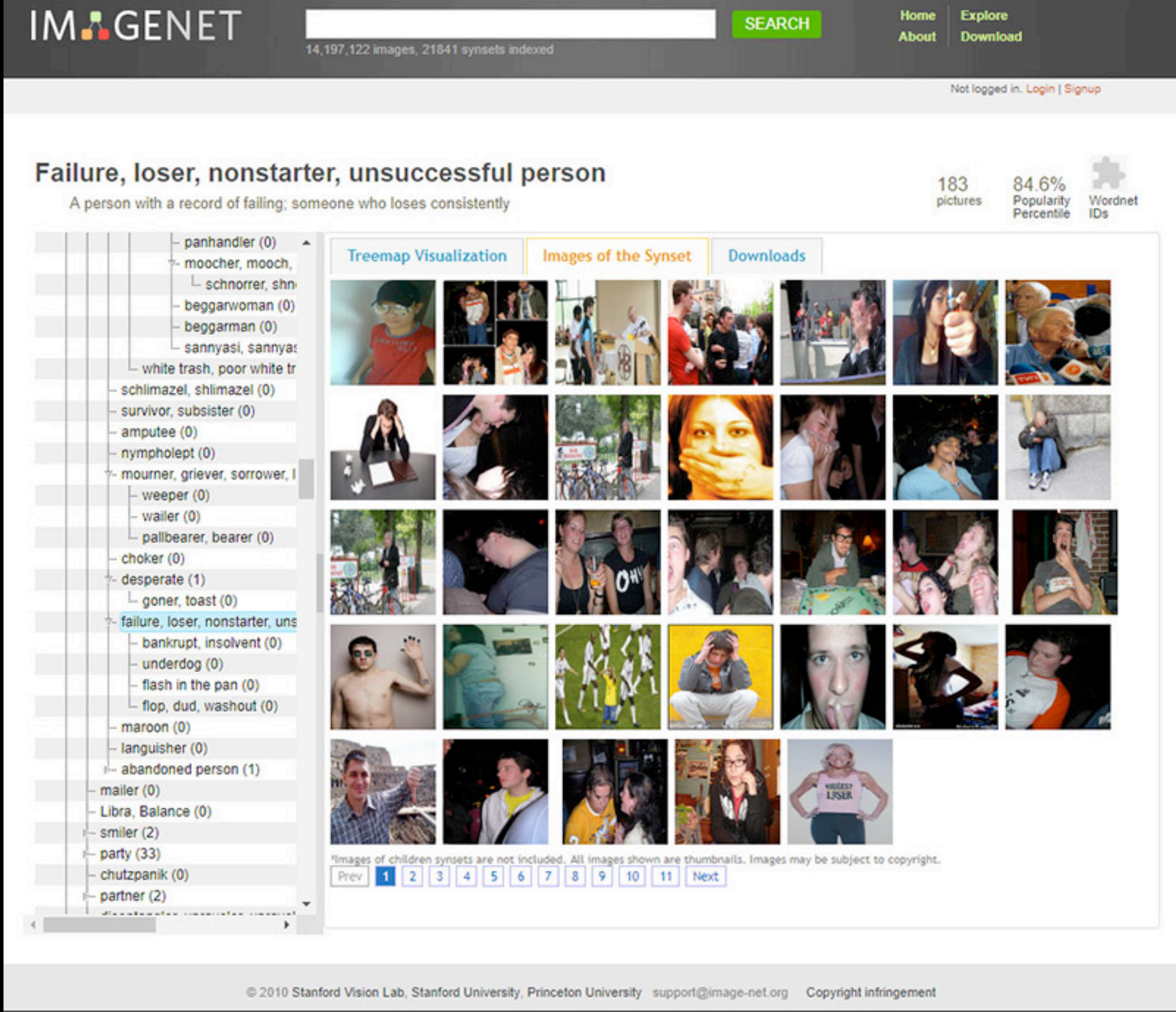
Orbital Elements for Unknown Objects (Paglen Studio)

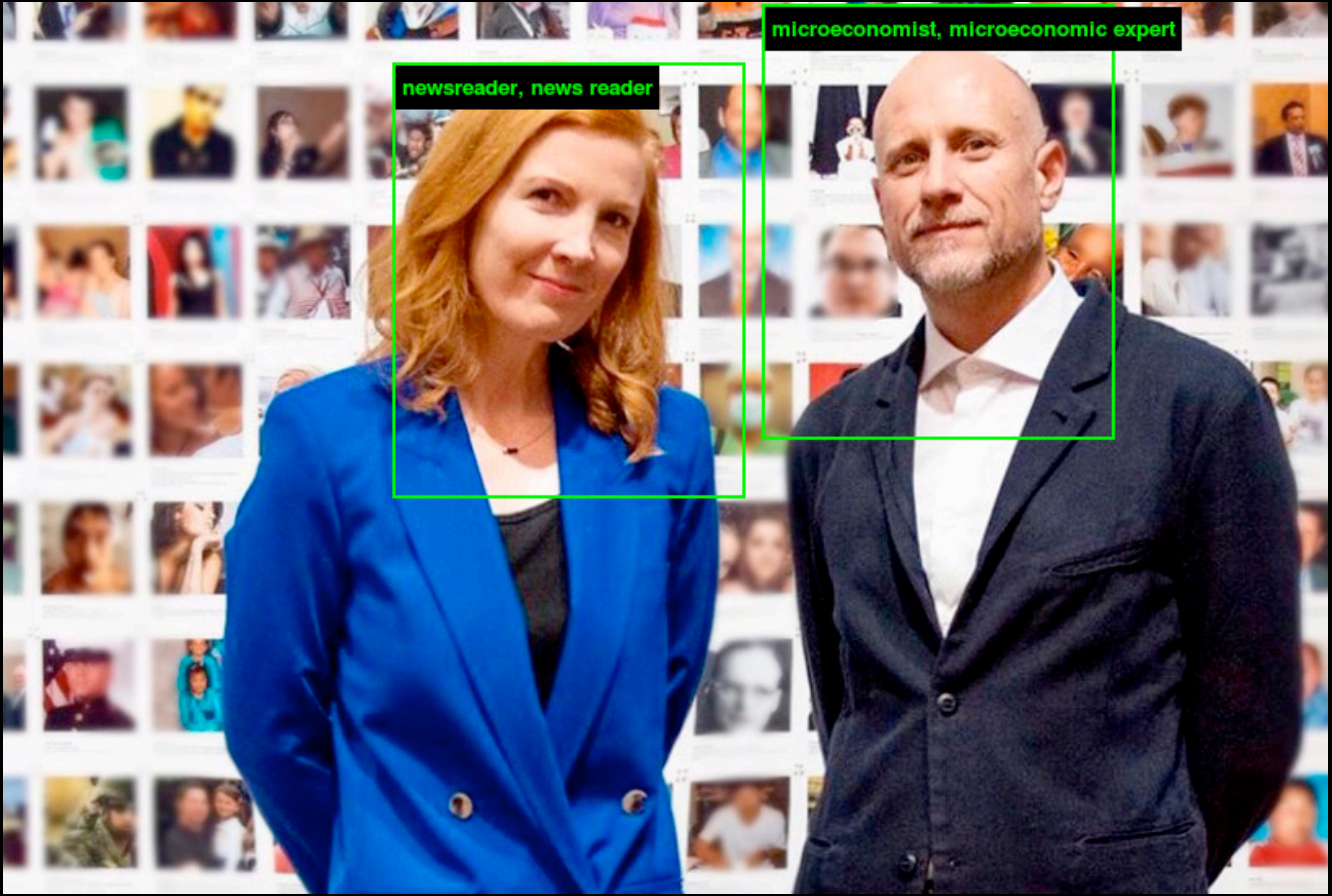
```
Unidentified - TO BE ASSIGNED
1 82891U          19274.97211782 +.00000236 +00000++40169-4  9992
2 82891  97.8262 340.4146 0014238 192.8285 167.2566 14.77392039103911
Unidentified - TO BE ASSIGNED
1 82939U          21305.97543406 -.00000116 +00000--78523-2  9997
2 82939 086.3573 310.0542 1886417 292.7704 048.5108 08.75629108323084
Unidentified - TO BE ASSIGNED
1 84006U          22230.06235828 -.00000089 +00000++00000+ 9997
2 84006  23.7246  50.0003 6131819  14.0642 357.3454  2.16231398 57994
Unidentified - TO BE ASSIGNED
1 84353U          22157.1200206+.00000353 +00000++00000+ 9990
2 84353  18.8985 279.1353 6683612 272.4086  18.7931  2.037266041915
Unidentified - TO BE ASSIGNED
1 85205U          22230.48154229 +.00000122 +00000++10863-3  9999
2 85205  82.9573 134.6974 0038197  79.5215  46.9552 13.76490932439570
Unidentified - TO BE ASSIGNED
1 85410U          20016.12638888  .00000006  00000- 00000+ 9996
2 8541 80.0053  91.4245 0009717 273.382201.61913.14633027    10
Unidentified - TO BE ASSIGNED
1 85412U          20016.34702987  .00000012  00000- 00000+ 9991
2 85412  88.910 82.4227 001013287.140204.57013.55759361    14
Unidentified - TO BE ASSIGNED
1 87730U          22234.79320754 +.00000069 +00000++56451-4  9994
2 8773 82.9432  47.6587 0031023 128.3305 232.0646 13.74210671419417
Unidentified - TO BE ASSIGNED
1 87934U          22234.40224825 +.00000053 +00000++40448-4  9992
2 87934  82.9356 112.5085 0031253 152.3913  24.1539 13.72830243438503
Unidentified - TO BE ASSIGNED
1 89177U          22231.34219058 +.00000062 +00000++50806-4  9994
2 89177  82.9179  87.0619 0031057 272.6767  87.0834 13.73436645896661
```

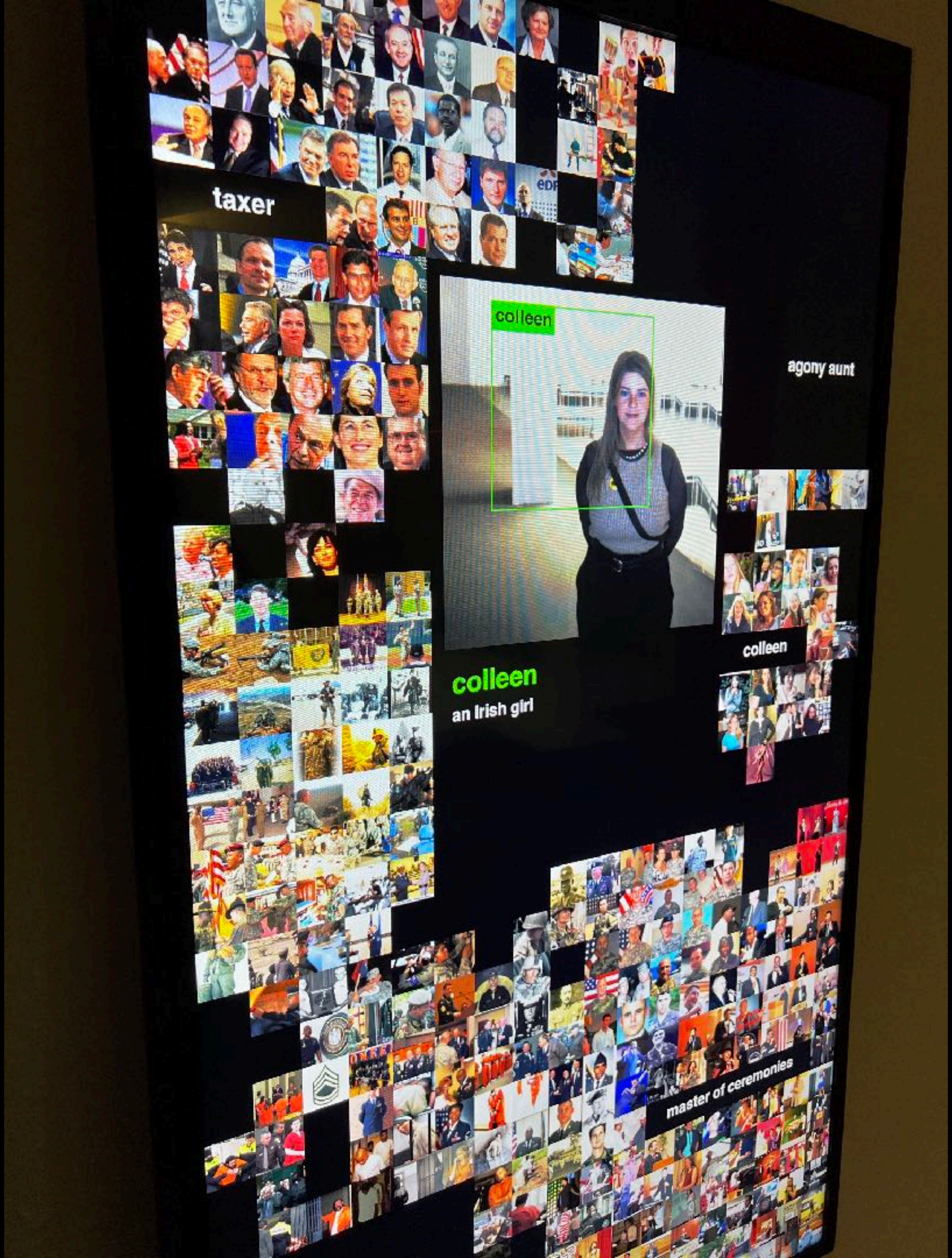
ImageNet Roulette

https://paglen.studio/2020/04/29/imagenet-roulette/

ImageNet Roulette is trained on the “person” categories from a dataset called ImageNet (developed at Princeton and Stanford Universities in 2009), one of the most widely used training sets in machine learning research and development.

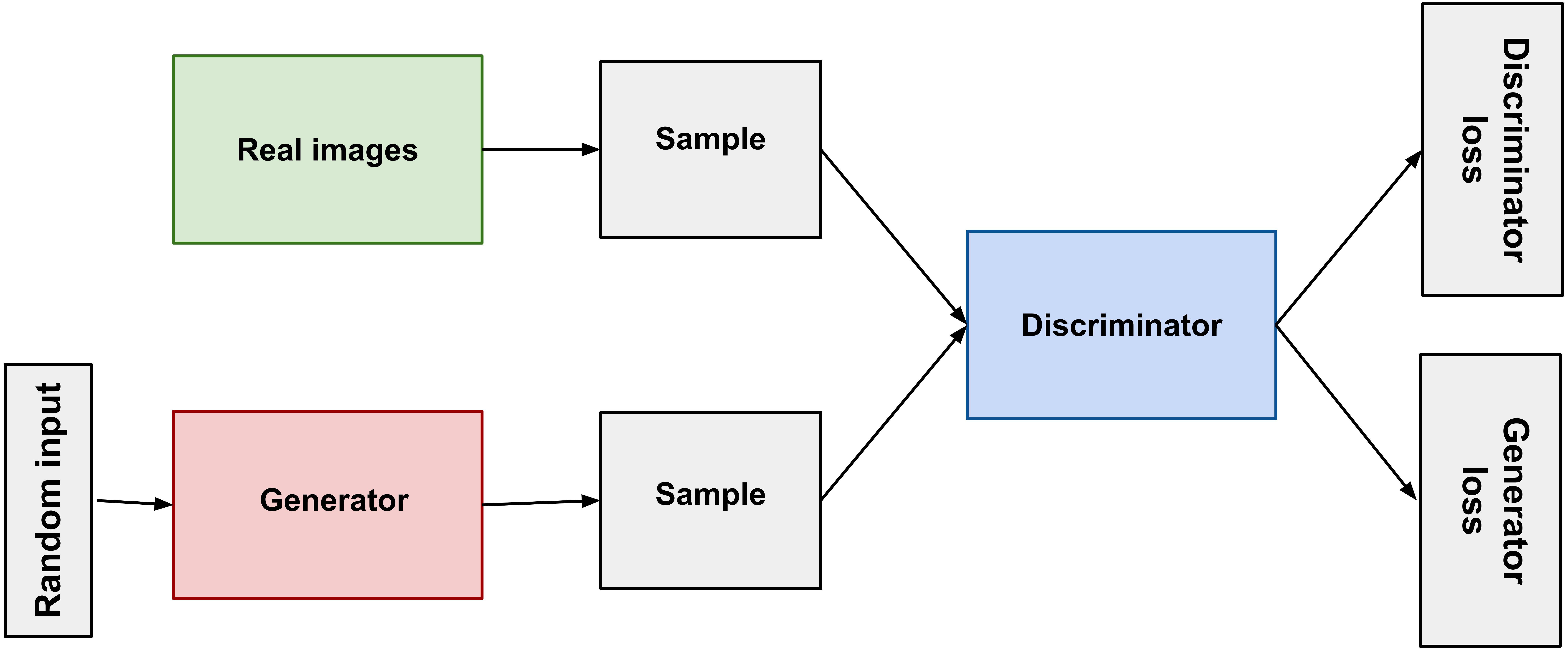



newsreader, news reader
microeconomist, microeconomic expert

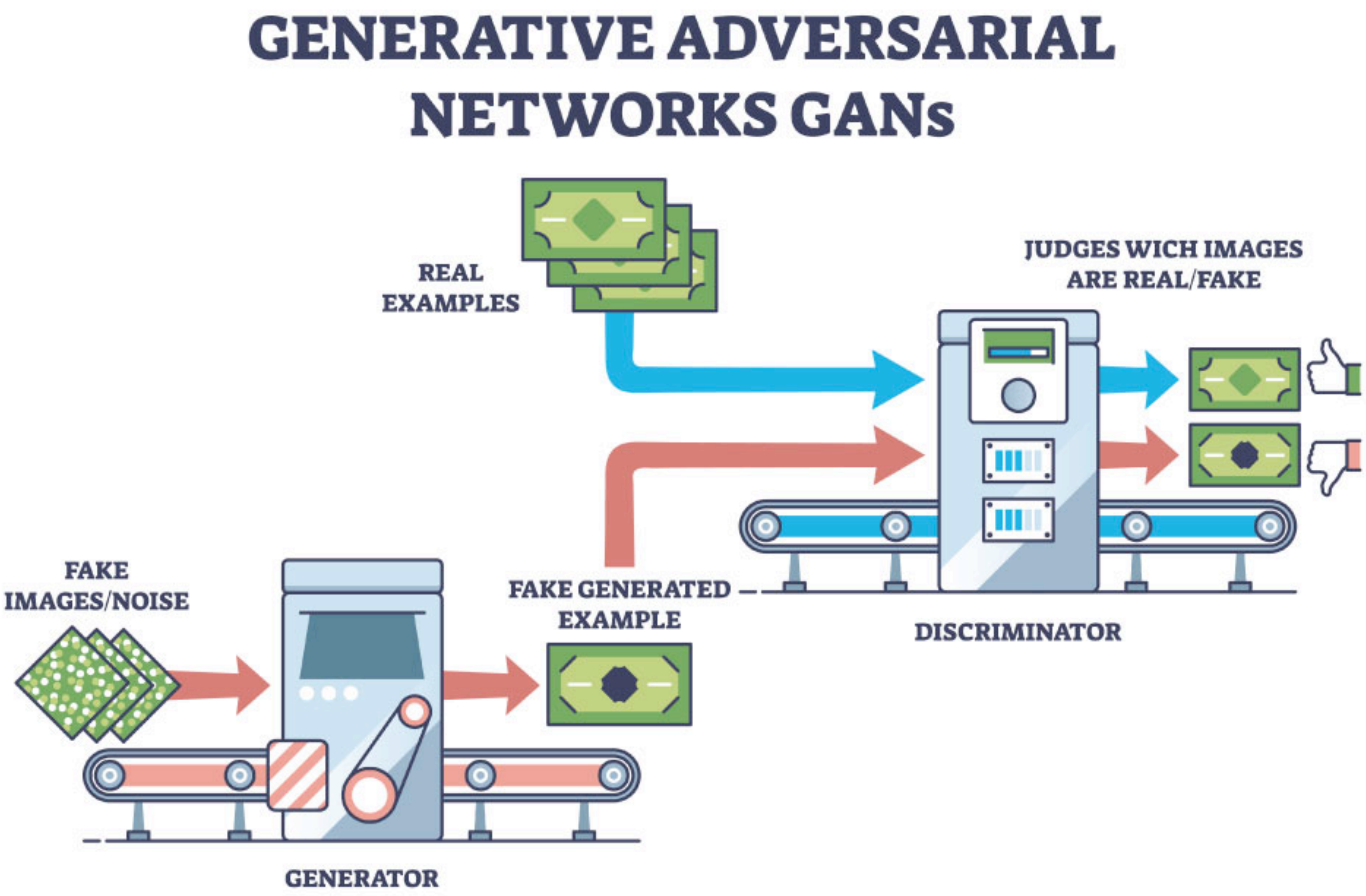

taxer
colleen
agony aunt
colleen
colleen
an Irish girl
master of ceremonies

The project was a provocation, acting as a window into some of the racist, misogynistic, cruel, and simply absurd categorizations embedded within ImageNet and other training sets that AI models are build upon.

FastGAN



https://developers.google.com/machine-learning/gan/gan_structure



https://pg-p.ctme.caltech.edu/blog/ai-ml/what-is-generative-adversarial-network-types

Memories of Passersby I, Mario Klingemann

GAN


Memories of Passersby I
Mario Klingemann

The Shell Record, Anna Ridler

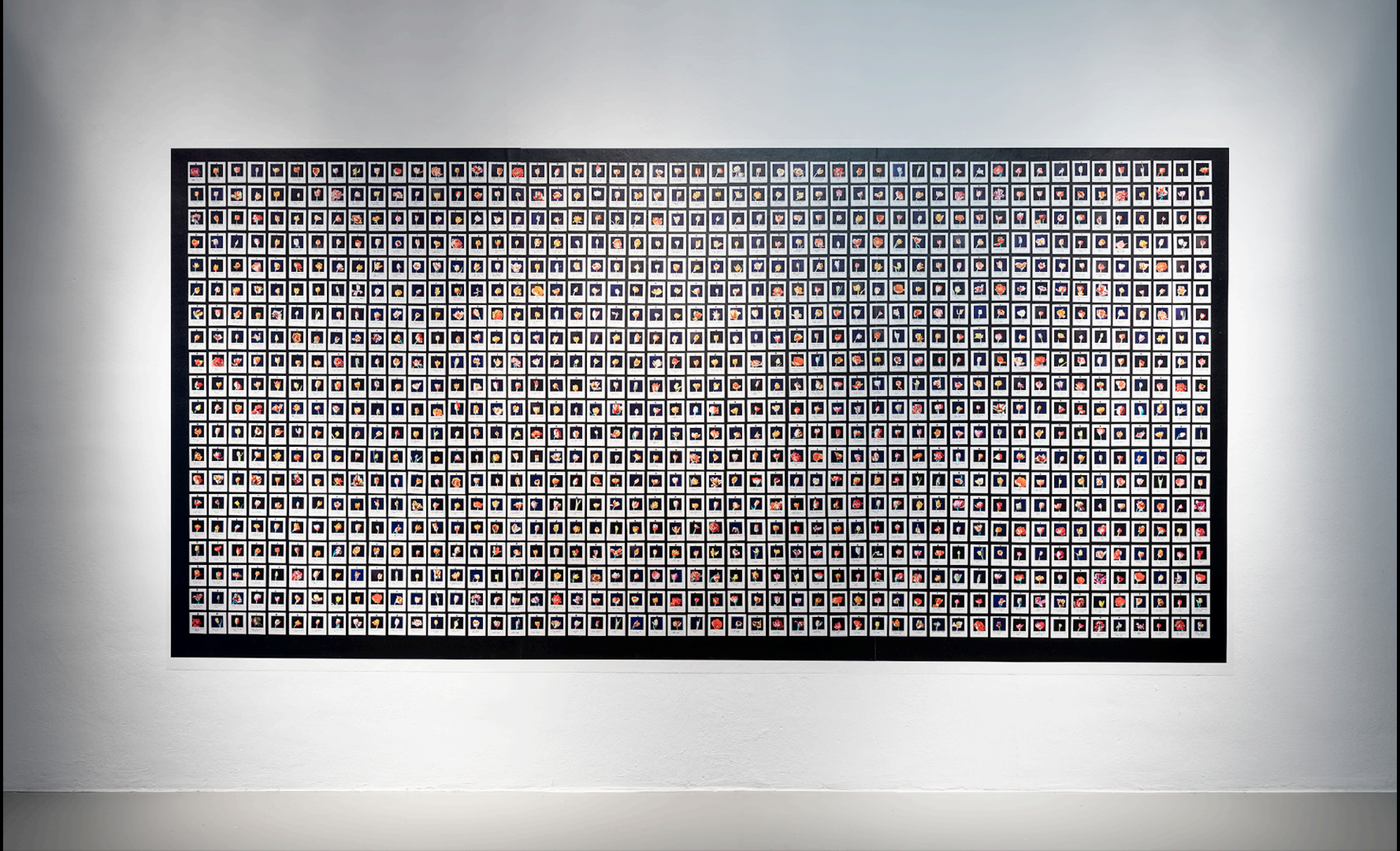
GAN

https://annaridler.com/the-shell-record-2021

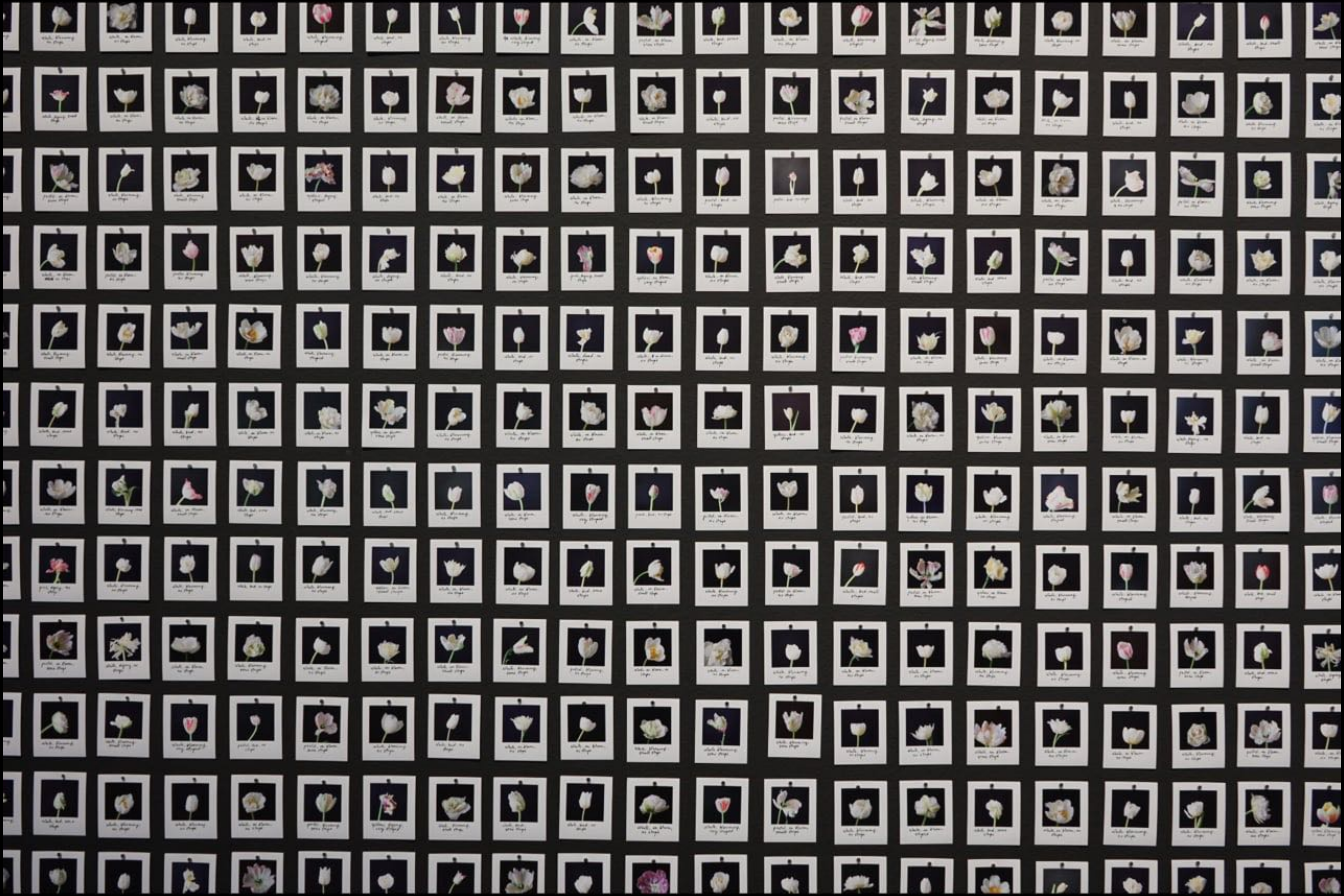

Join
us

Mosaic Virus, Anna Ridler

GAN

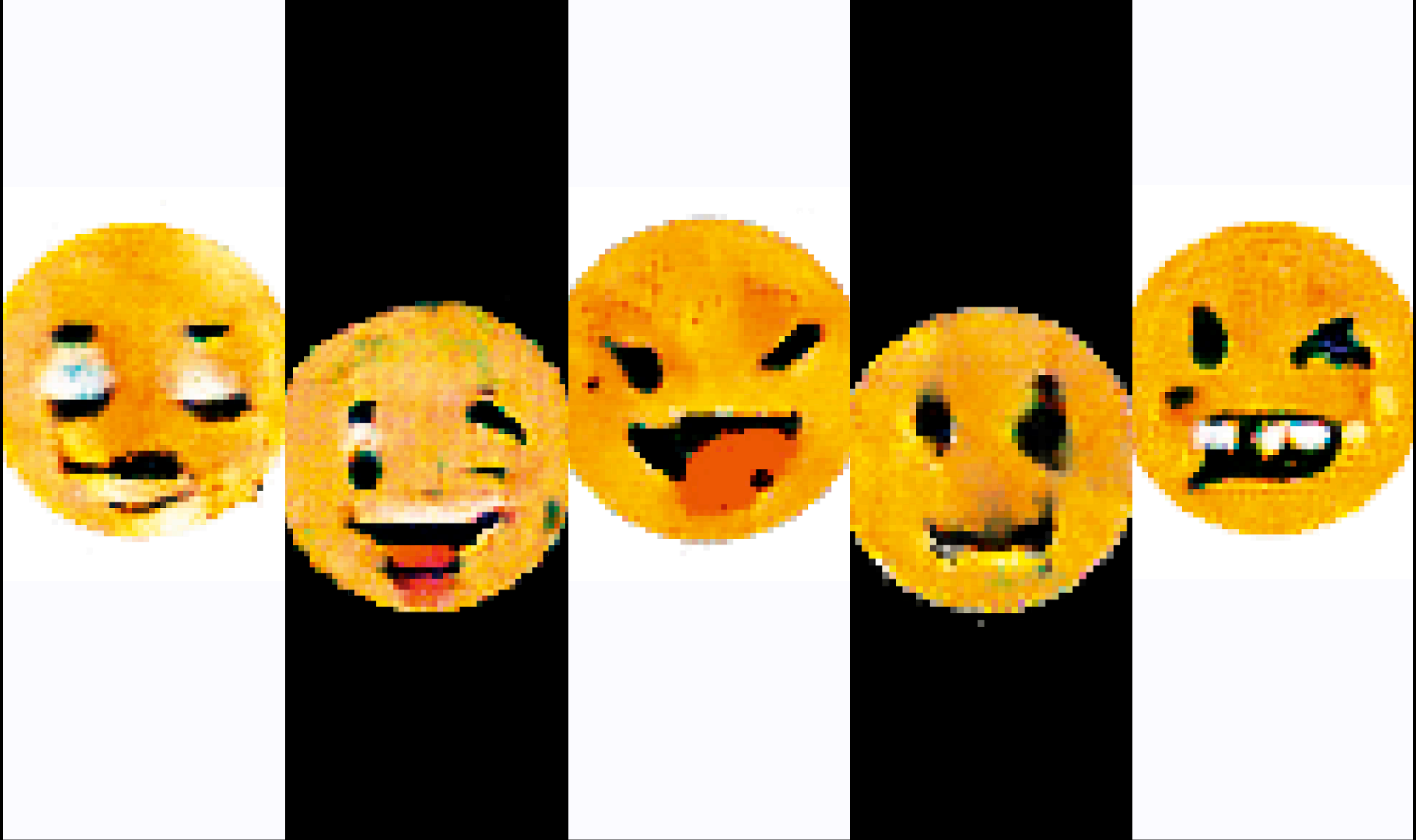


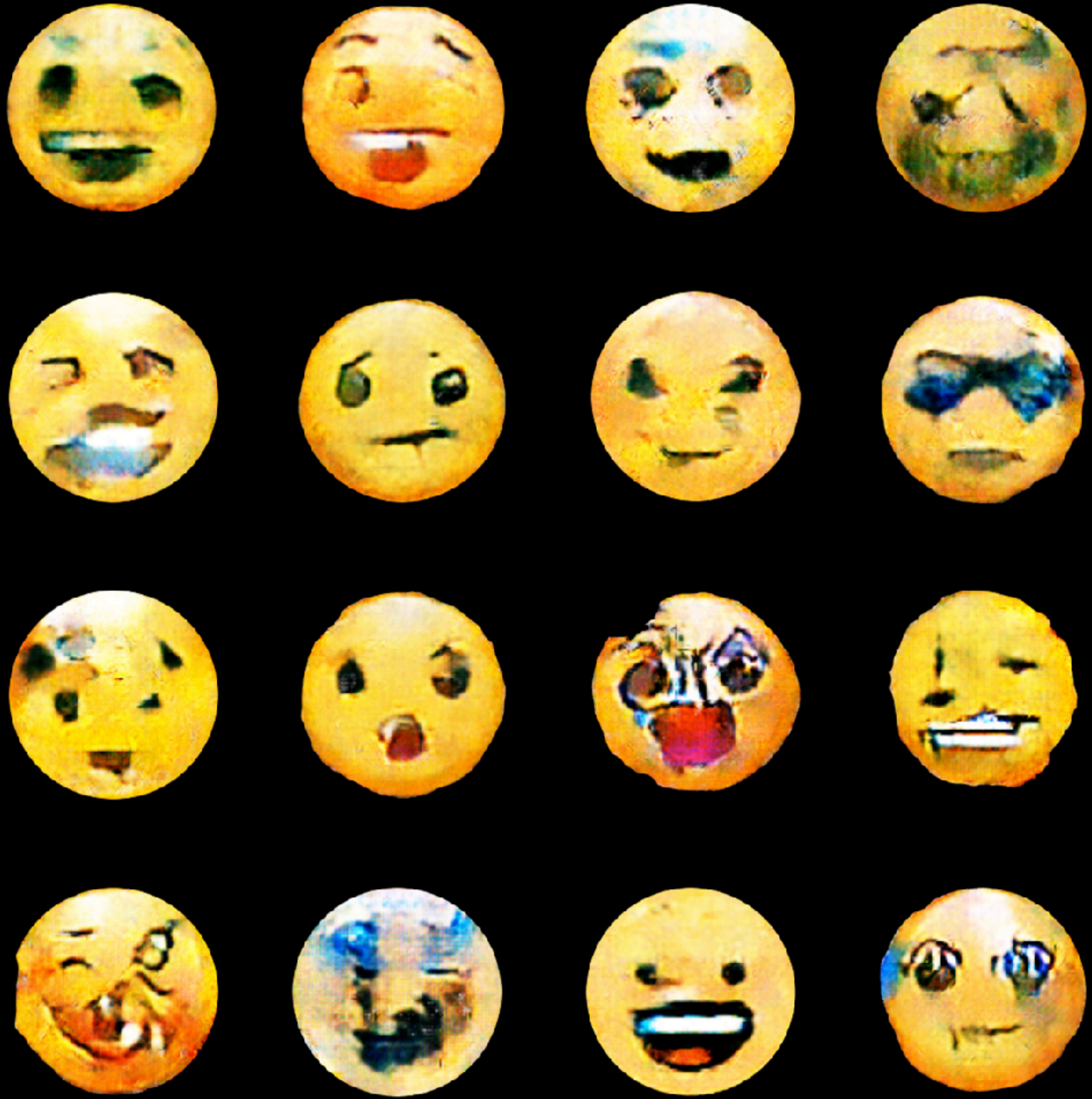

white, dying, no stripe
white, bud, no stripe
white, blooming, no stripe
pink, dying, small stripe
yellow, in bloom, very striped
white, in bloom, no stripe
white, blooming, small stripe
white, bud, some stripe
white, blooming, small stripe
white, in bloom, no stripe
pastel, blooming, no stripe
white, bud, no stripe
white, dead, no stripe
white, in bloom, no stripe
white, bud, no stripe
white, blooming, no stripe
pastel, blooming, small stripe
white, in bloom, no stripe
yellow, in bloom, some stripe
white, blooming, no stripe
white, in bloom, no stripe
white, in bloom, no stripe
white, in bloom, small stripe
white, in bloom, no stripe
yellow, bud, no stripe
white, blooming, no stripe
white, in bloom, no stripe
white, bud, some stripe
white, blooming, no stripe
white, in bloom, some stripe
white, blooming, very striped
pink, bud, no stripe
pastel, in bloom, no stripe
white, in bloom, no stripe
pastel, bud, no stripe
yellow, in bloom, no stripe
yellow, in bloom, small stripe
white, in bloom, no stripe
white, in bloom, no stripe
white, bud, some stripe
white, in bloom, small stripe
white, blooming, no stripe
pastel, in bloom, no stripe
white, bud, small stripe
pastel, in bloom, some stripe



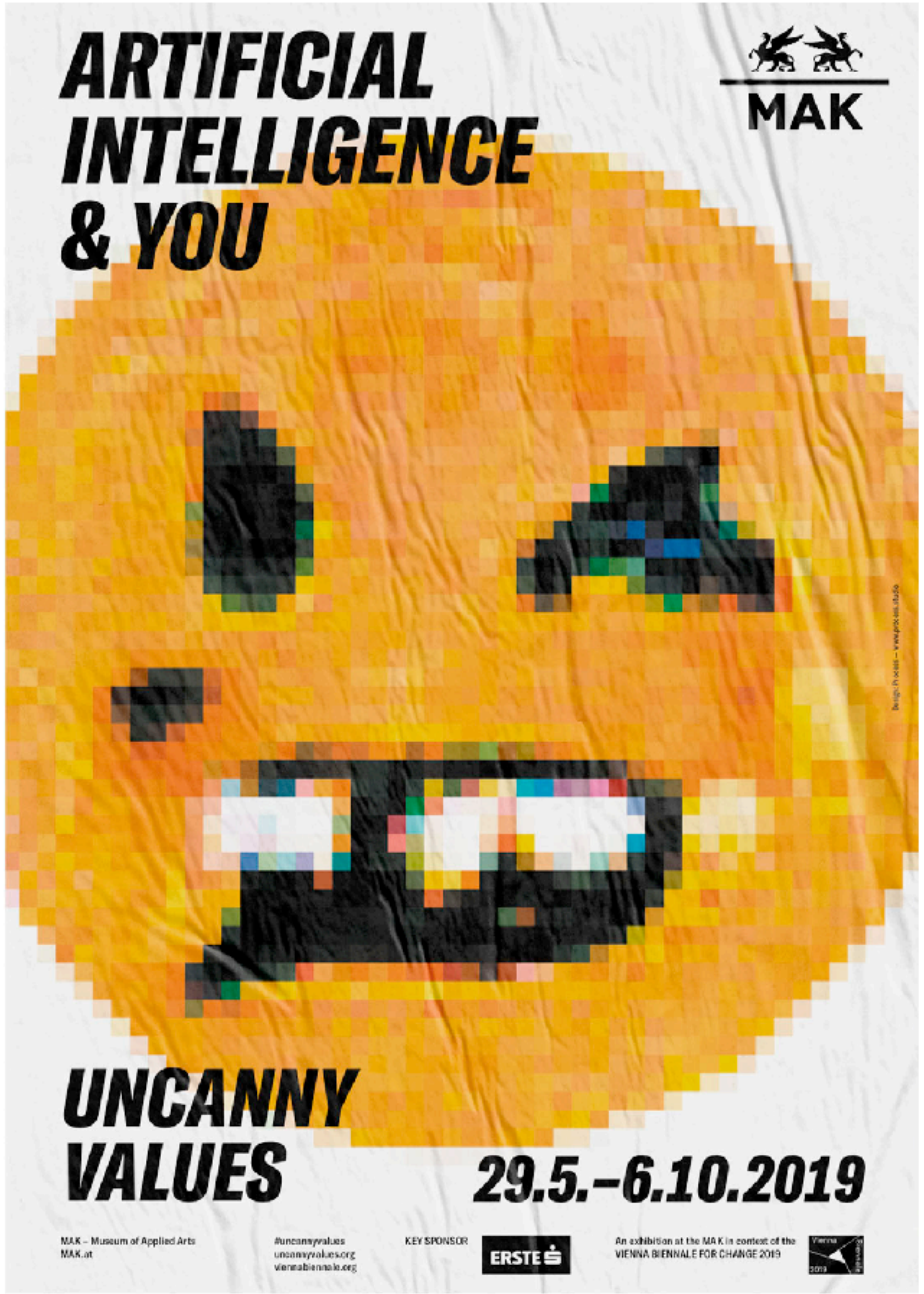
AImoji, Process Studio

GAN





ARTIFICIAL
INTELLIGENCE
& YOU
MAK
UNCANNY
VALUES
29.5.–6.10.2019
MAK – Museum of Applied Arts
MAK.at
#uncannyvalues
uncannyvalues.org
viennabiennale.org
KEY SPONSOR
ERSTE
An exhibition at the MAK in context of the
VIENNA BIENNALE FOR CHANGE 2019



ARTIFICIAL
INTELLIGENCE
& YOU
MAK
UNCANNY
VALUES
29.5.–6.10.2019
MAK – Museum of Applied Arts
MAK.at
#uncannyvalues
uncannyvalues.org
viennabiennale.org
KEY SPONSOR
ERSTE
An exhibition at the MAK in context of the
VIENNA BIENNALE FOR CHANGE 2019


KÜNSTLICHE INTELLIGENZ & DU
MAK
UNCANNY VALUES
29.5.–6.10.2019
KÜNSTLICHE INTELLIGENZ & DU
MAK
UNCANNY VALUES
29.5.–6.10.2019
STOP
ÖAMTC
EASY WAY
GRATIS!
Erde

Pix2Pix

Image-to-Image Translation


Labels to Street Scene
input
output
Aerial to Map
input
output
Labels to Facade
input
output
BW to Color
input
output
Day to Night
input
output
Edges to Photo
input
output


Monet ⇄ Photos
Zebras ⇄ Horses
Summer ⇄ Winter
Monet → photo
zebra → horse
summer → winter
photo → Monet
horse → zebra
winter → summer
Photograph
→
Monet
Van Gogh
Cezanne
Ukiyo-e

Memo Akten's “Learning to see”

pix2pix


#Deepdream is blowing my mind (2015)
memotv

Anna Ridler’s “Fall of the House of Usher”

pix2pix

THE FALL OF THE HOUSE OF USHER

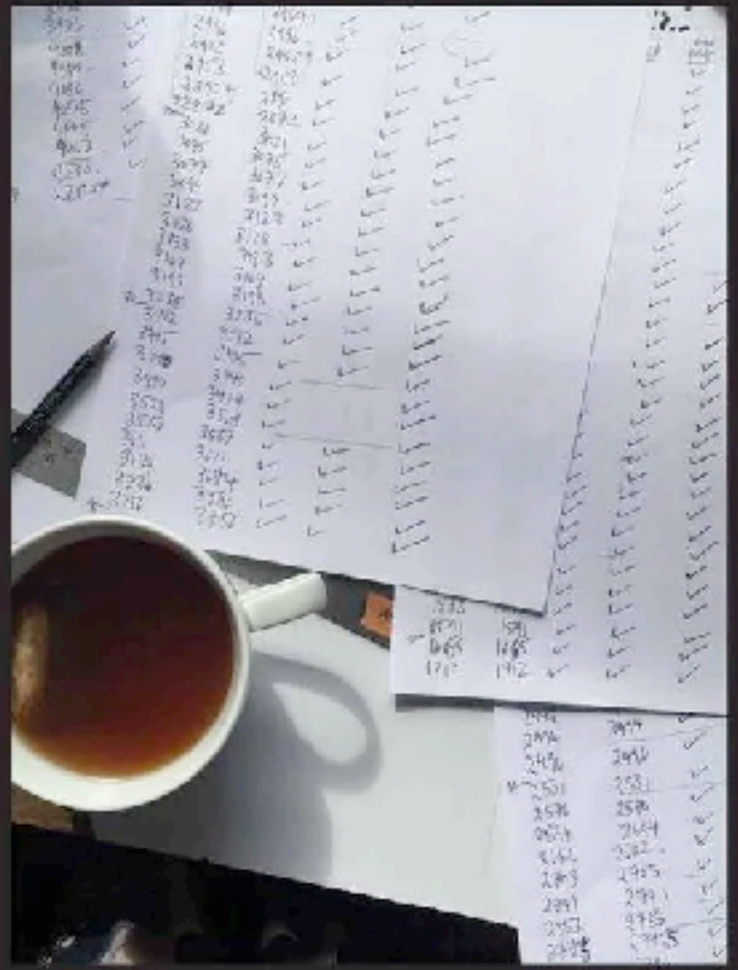
A Film Version of Poe's Story

By

MELVILLE WEBBER

“To make my animation I made a training set of set of 200 drawings using the 1929 version of the Fall of the House of Usher as my base so that the GAN would essentially learn how to draw in my style.”

Anna Ridler, Fall of the House of Usher,
https://www.vam.ac.uk/blog/museum-life/guest-blog-post-fall-of-the-house-of-usher-datasets-and-decay



1–4

“The errors and choices that are made when drawing are amplified and the GAN holds a mirror up to my own drawing and makes me realise things that I was not aware of: what I find the most important, what I always edit out.”

Anna Ridler, Fall of the House of Usher,
https://www.vam.ac.uk/blog/museum-life/guest-blog-post-fall-of-the-house-of-usher-datasets-and-decay

“The GAN copies me, learning the human-ness and mistakes that come as part of this process.”

Anna Ridler, Fall of the House of Usher,
https://www.vam.ac.uk/blog/museum-life/guest-blog-post-fall-of-the-house-of-usher-datasets-and-decay

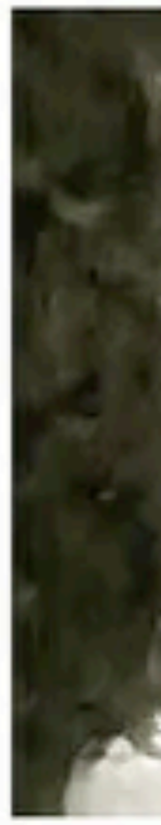
1

2

3

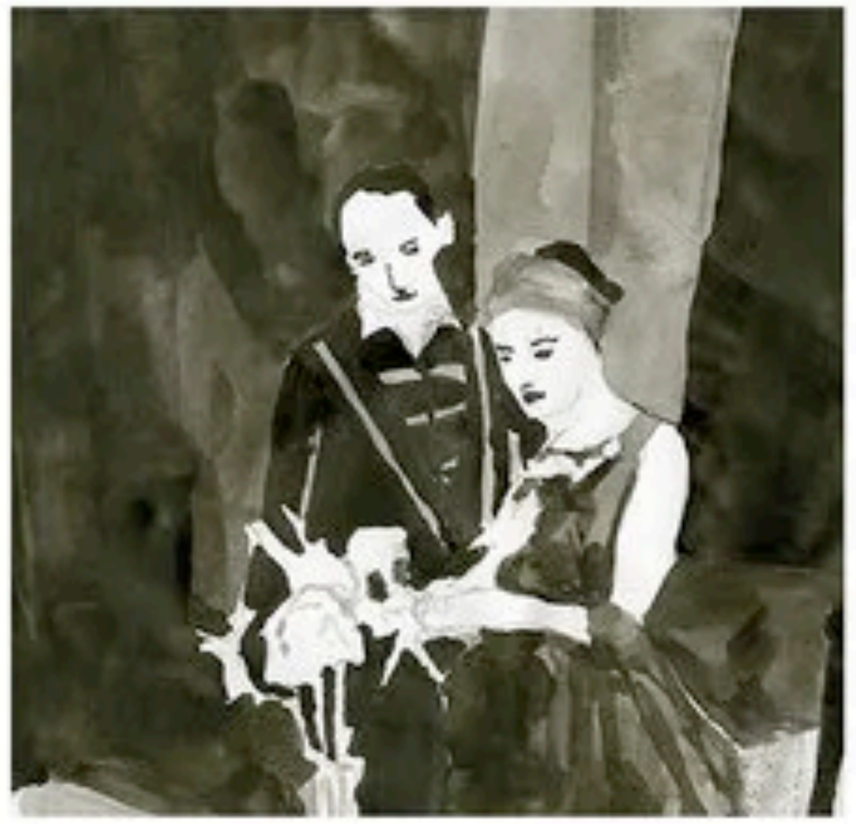
3

3



2

2



2

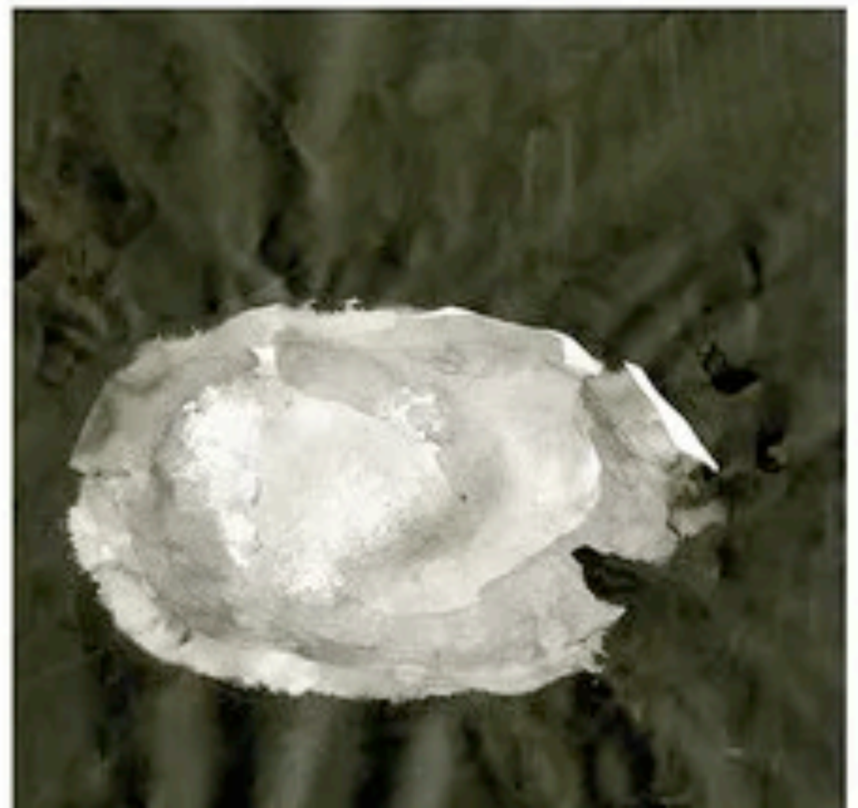
4

1

3

3

1. Still from 1929 film
2. Drawing of film
3. GAN generated image
4. Drawing of GAN generated image



“Producing an image using a GAN versus any other way gives the viewer a different experience, expectation, history, traces and contexts to consider. What are these associations and how might they be used in a piece of work? […] Can a GAN or indeed training set become a “wilful actor and agent within artistic processes" [5] as other materials are?”

https://annaridler.com/gans-in-art

Alternative Face, Mario Klingemann

pix2pix

The piece employs a generative adversarial network (GAN), specifically the pix2pix model, which pairs two neural networks—one generating images, the other refining them through critique. I trained this model on thousands of facial markers derived from music video clips of French singer Françoise Hardy, chosen for her expressive features. In the resulting video, Conway's face morphs into Hardy's, her "alternative facts" audio driving the illusion.

https://quasimondo.com/2017/02/04/alternative-face/

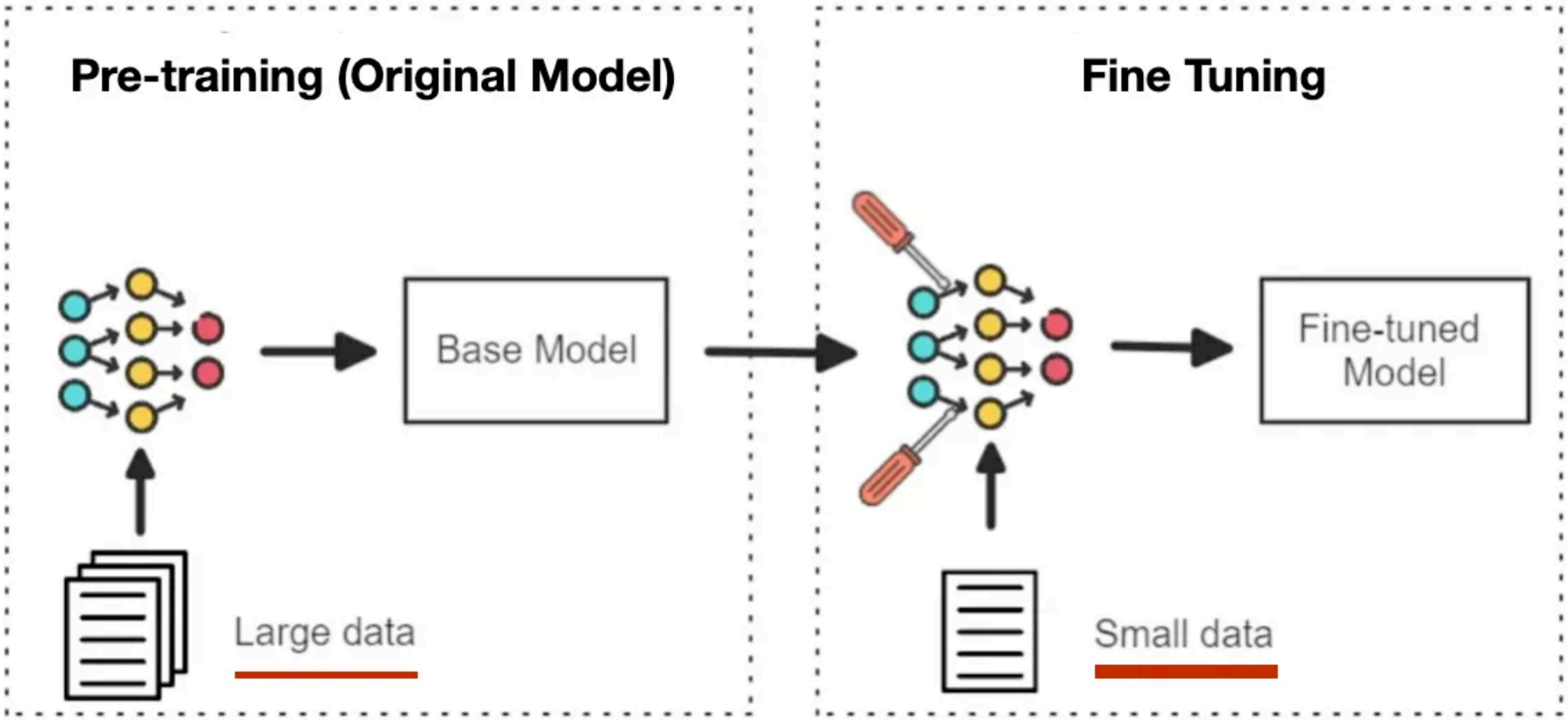
Fine-Tuning

Fine-tuning is the process of adapting a pre-trained model to a specific task using a smaller, specialized dataset.

https://www.ibm.com/think/topics/fine-tuning

https://www.geeksforgeeks.org/deep-learning/what-is-fine-tuning/

https://www.digitalocean.com/resources/articles/gpu-options-finetuning



https://python.plainenglish.io/fine-tune-llms-a-comprehensive-guide-between-full-partial-fine-tuning-an-end-to-end-python-3fa7223f5519

Fine-tuning allows developers to adapt a model to specialized domains such as legal, medical, or customer support applications.

https://python.plainenglish.io/fine-tune-llms-a-comprehensive-guide-between-full-partial-fine-tuning-an-end-to-end-python-3fa7223f5519

Cost-Effectiveness

GPT-3.5 4.600.000$ - 15.000.000$

ChatGPT-3 2.000.000$ - 4.000.000$

https://www.forbes.com/sites/katharinabuchholz/2024/08/23/the-extreme-cost-of-training-ai-models/

Cost-Effectiveness

Stable Diffusion 1.5 600.000$

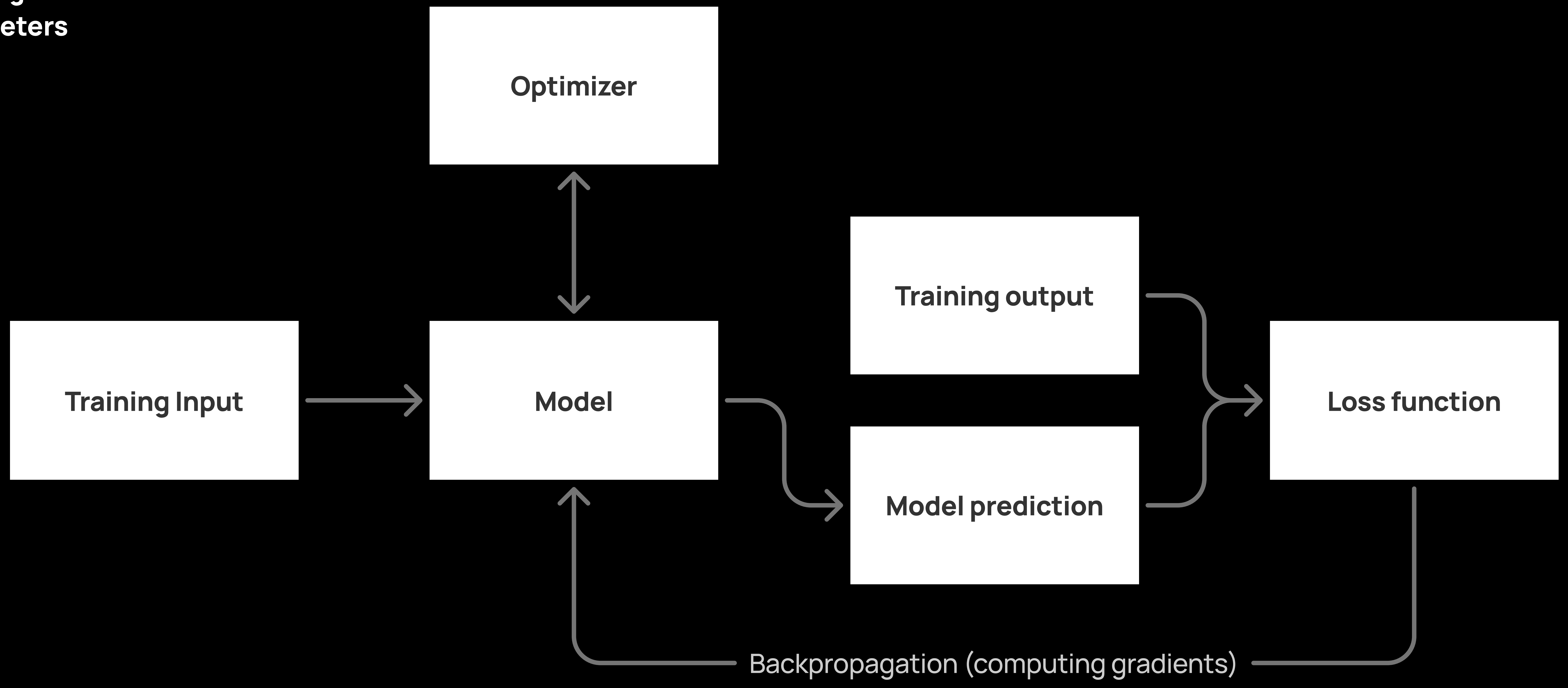
Fine-tuning 0.5$ - 10$ (~1.000-25.000€ when including pc hardware for training at home)

↖ There is a 1-pixel wide line at the start

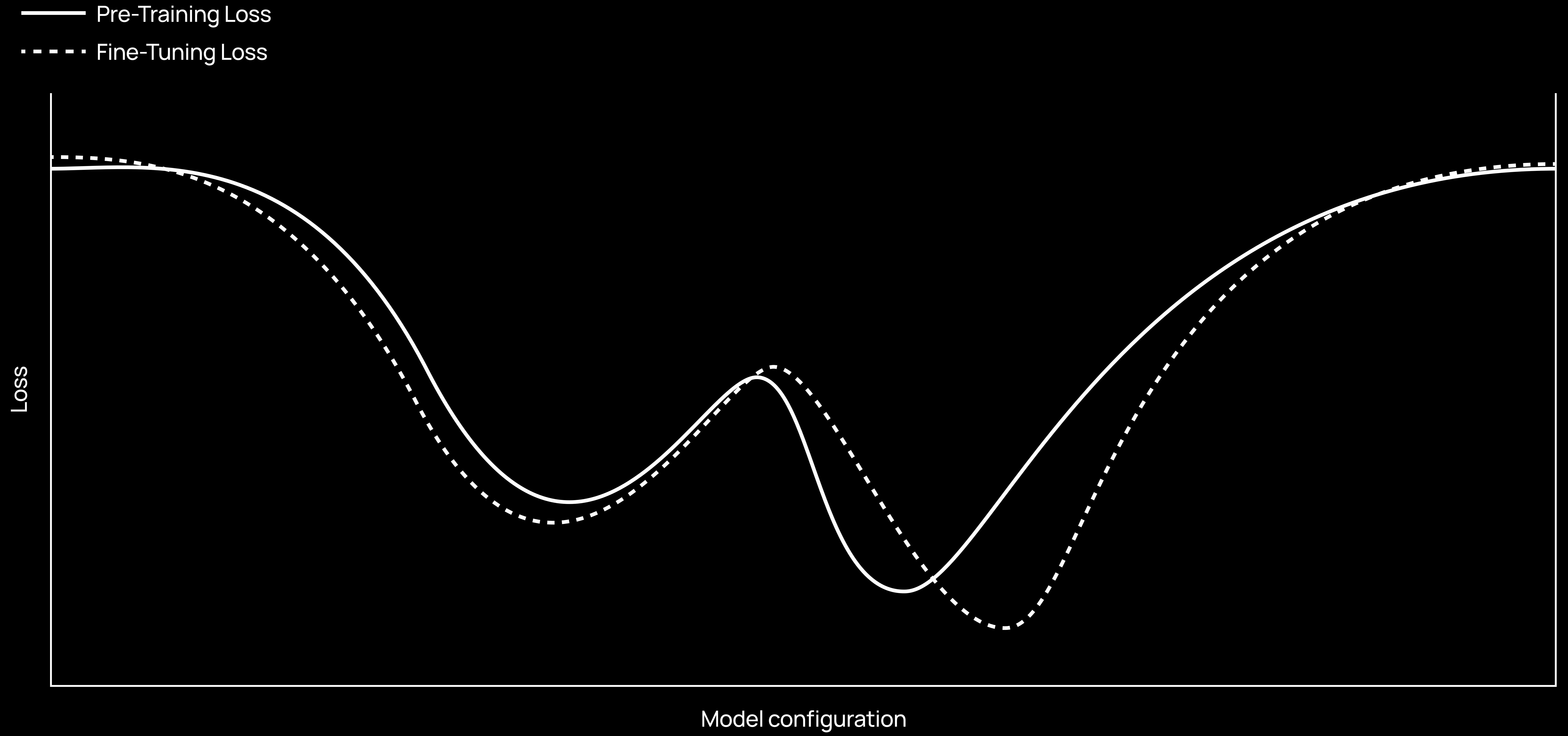
https://x.com/emostaque/status/1563870674111832066

Supervised Training and Fine-Tuning

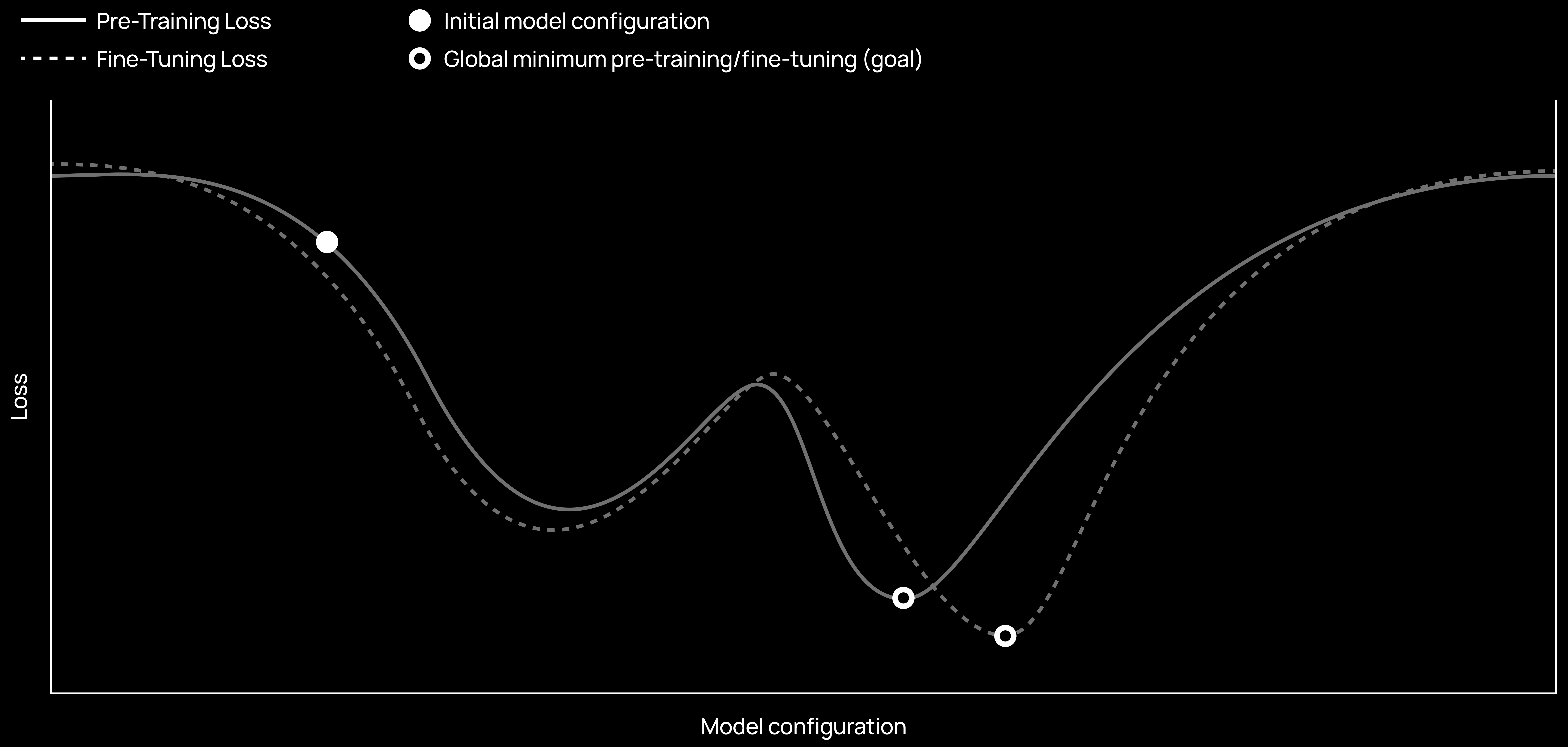
Fine-Tuning has the exact same steps as pre-training. The only differences are in the training data and the training parameters (the “hyperparameters”).



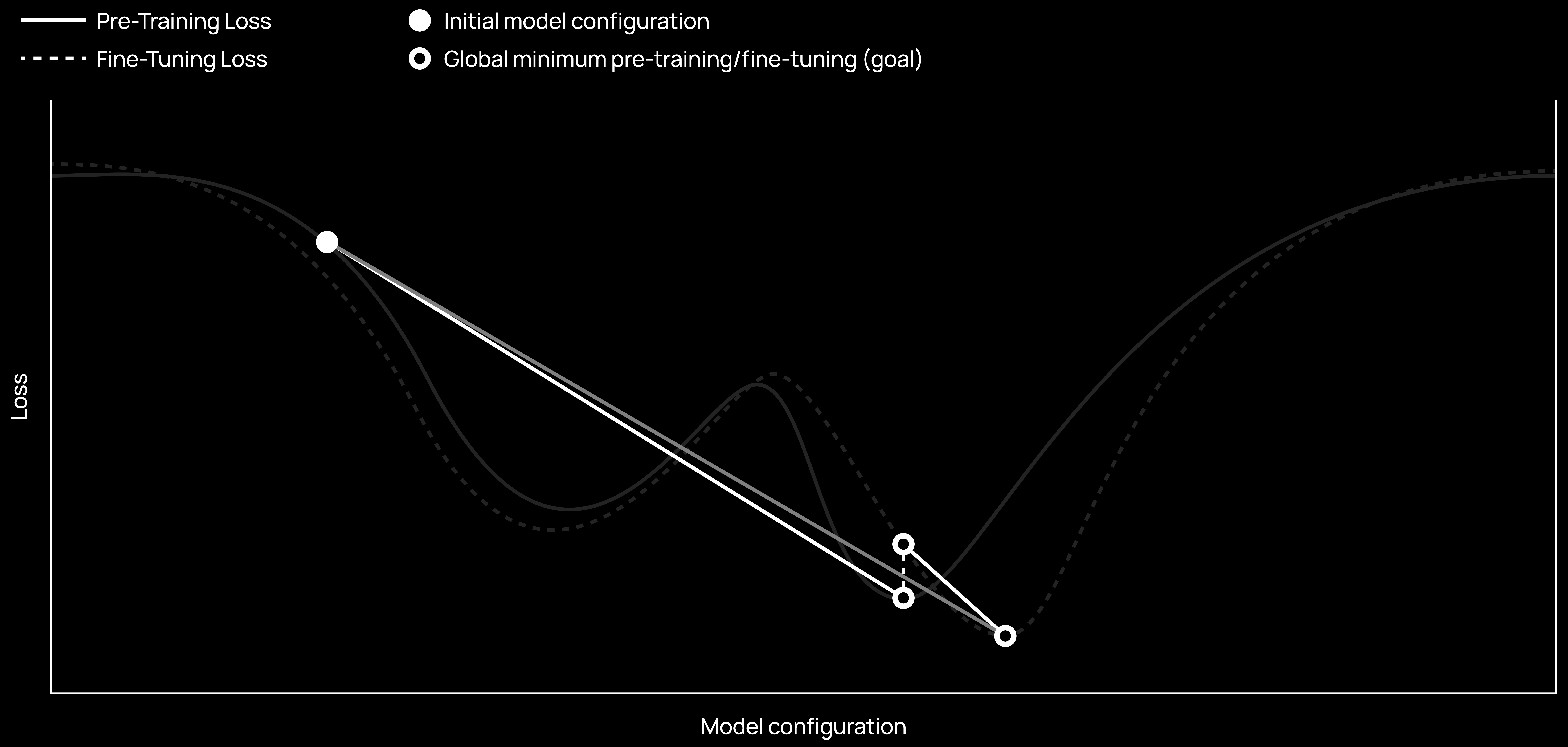
Supervised Training and Fine-Tuning



Supervised Training and Fine-Tuning



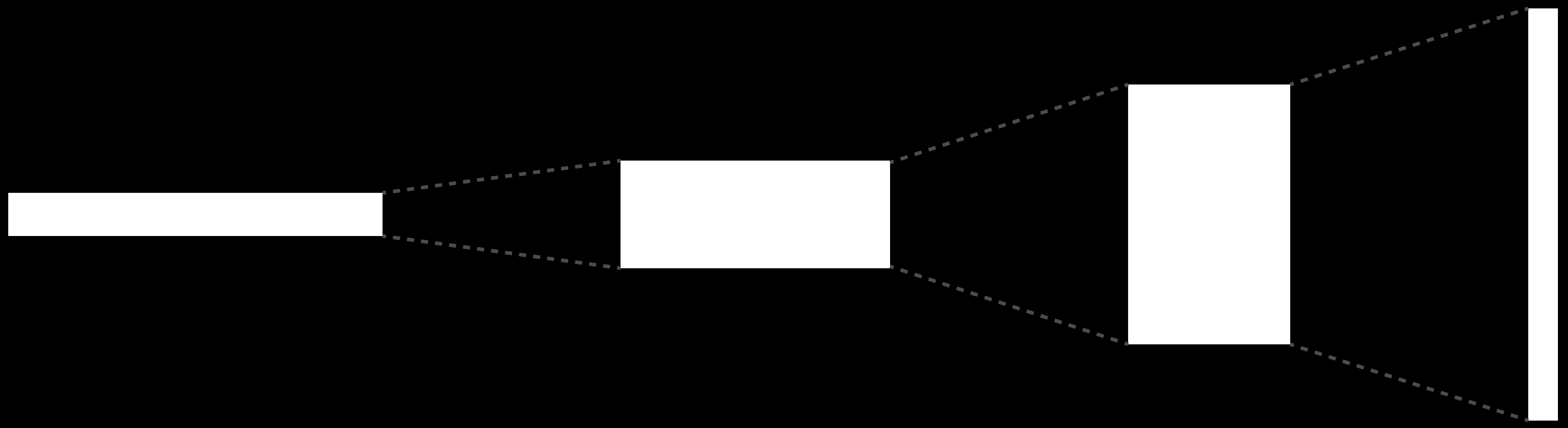
Supervised Training and Fine-Tuning



Forms of Fine-Tuning

01 Full fine-tuning
02 Partial fine-tuning
03 Adapter-based fine-tuning

Similarly to pre-training, all layers of a model change during fine-tuning. Full fine-tunes usually yield the best results, but are computationally expensive and prone to overfitting.



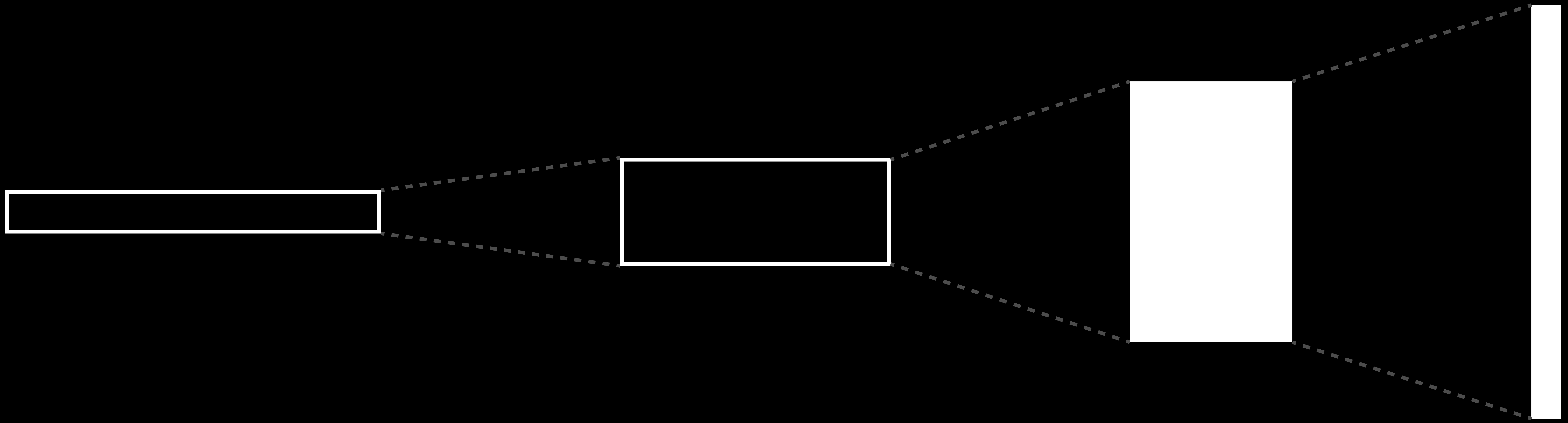
Forms of Fine-Tuning

01 Full fine-tuning

02 Partial fine-tuning

03 Adapter-based fine-tuning

Certain layers of a model get ‘frozen’. That means no gradient calculations during backpropagation and no adjustments to these layers by the optimizer are being made. Usually the first layers get frozen since they deal with lower level features.



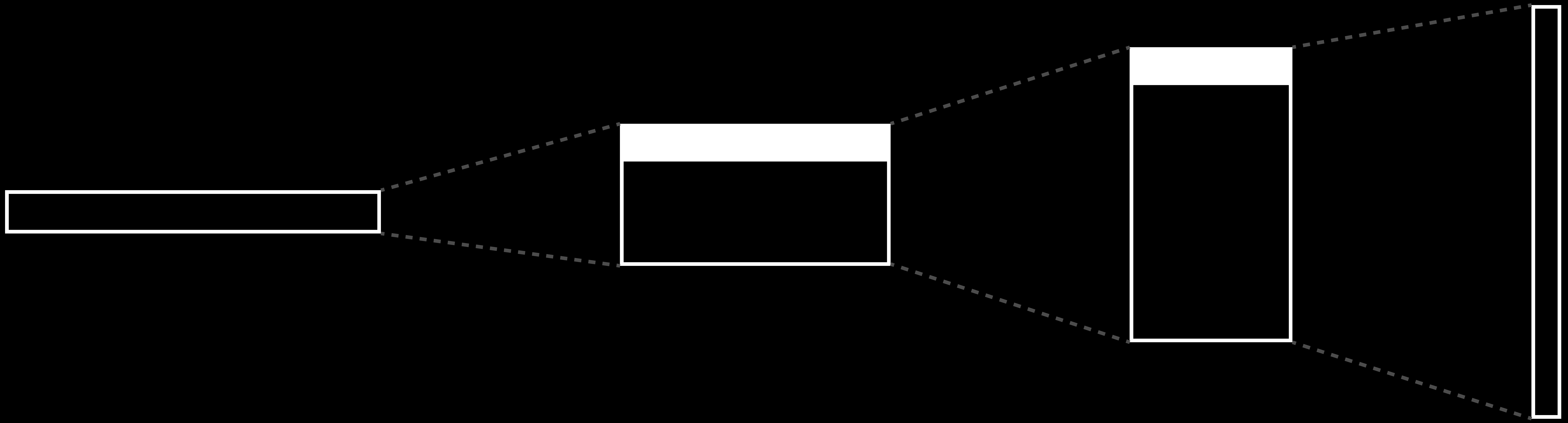
Forms of Fine-Tuning

01 Full fine-tuning

02 Partial fine-tuning

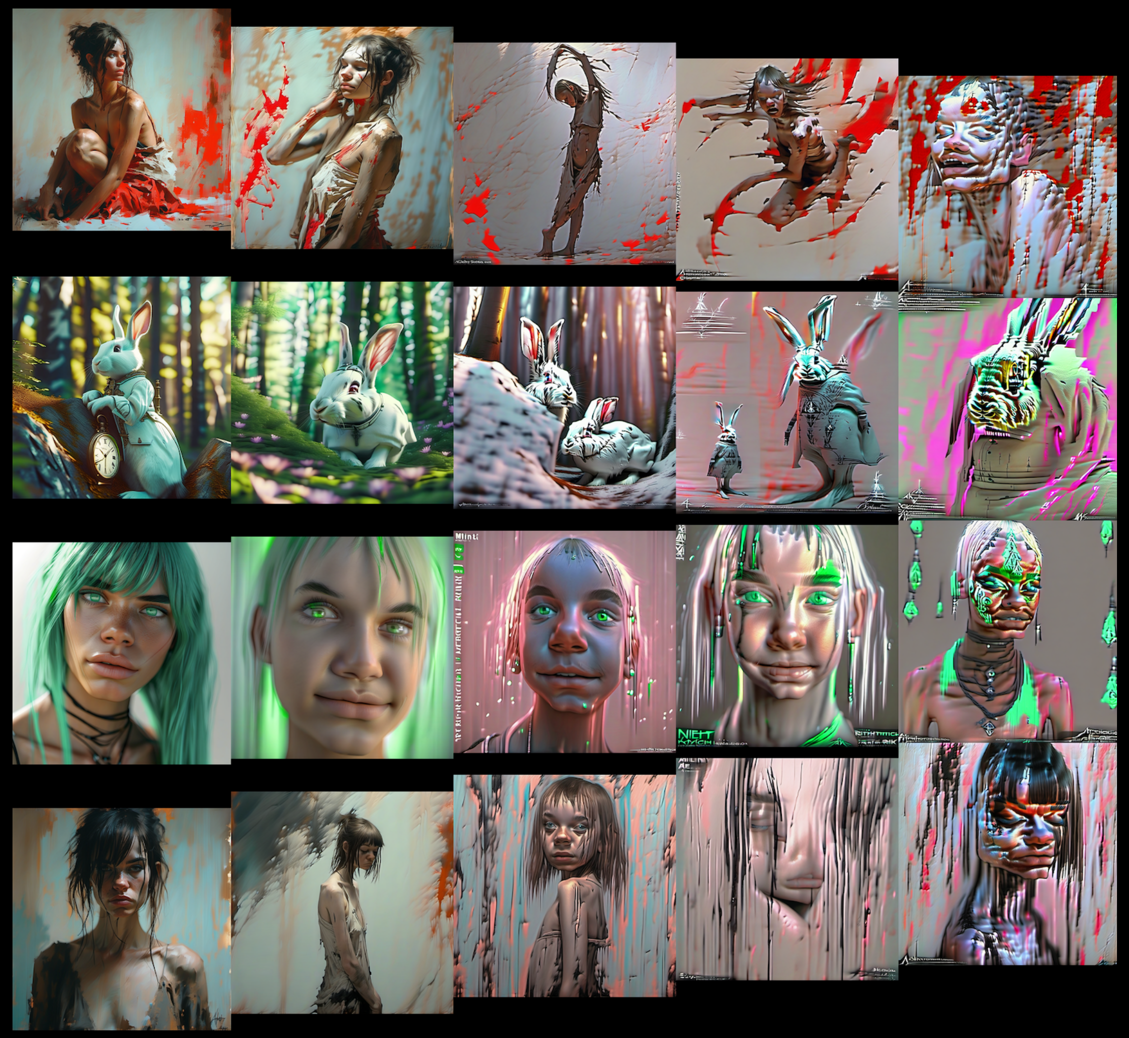
03 Adapter-based fine-tuning

Adapter-based methods don't change the model in any way. Instead they train separate smaller layers / matrices that get 'added' to the existing layers. That way they save on resources and output really small models that can be easily loaded and swapped out during inference.



Habsburg AI Portrait Studies, Martin Disley

Habsburg AI Portrait Studies is a series of cloth-printed generative portrait images produced using a bespoke diffusion model recursively retrained on its own output. This autophagous training causes the model to collapse on itself, constricting the model's distribution around the mean, forcing it to produce images in an amplified version of the model's default style.



https://www.creativeapplications.net/member/habsburg-ai-portrait-studies/

GPT-2 Generated Ceramic Recipes, Derek Au

Fine-Tuning

Crawly Elsie's Matte-04
38.0000 EP Kaolin
28.0000 Gerstley Borate
19.0000 G-200 Feldspar
9.0000   Lepidolite
6.0000   Soda Ash
4.0000   Wollastonite

Ame-Sosa-Wenkel
38.0000 Nepheline Syenite
29.0000 Silica
12.0000 Colemanite
8.0000   Whiting
6.0000   Dolomite
5.0000   Barium Carbonate
2.0000   Bentonite
1.0000   Rutile
0.7500   Copper Carbonate

Amber Celadon
34.0000 Albany slip
20.0000 Custer Feldspar
13.0000 Silica
13.0000 Wollastonite
6.0000   Whiting
3.0000   EP Kaolin
3.0000   Gerstley Borate
3.0000   Rutile
2.0000   Red Iron Oxide

Craters
30.0000 Lithium Carbonate
30.0000 Silica
15.0000 Borax
10.0000 Zircopax
10.0000 Kaolin
5.0000   Bentonite
3.0000   Copper Carbonate

Glazy **This Glaze Did Not Exist.**

OpenAI GPT-2 generated glaze recipes. Orton cone 6, oxidation.

OpenAI

xhairymutantx, Holly Herndon and Mat Dryhurst

Fine-Tuning

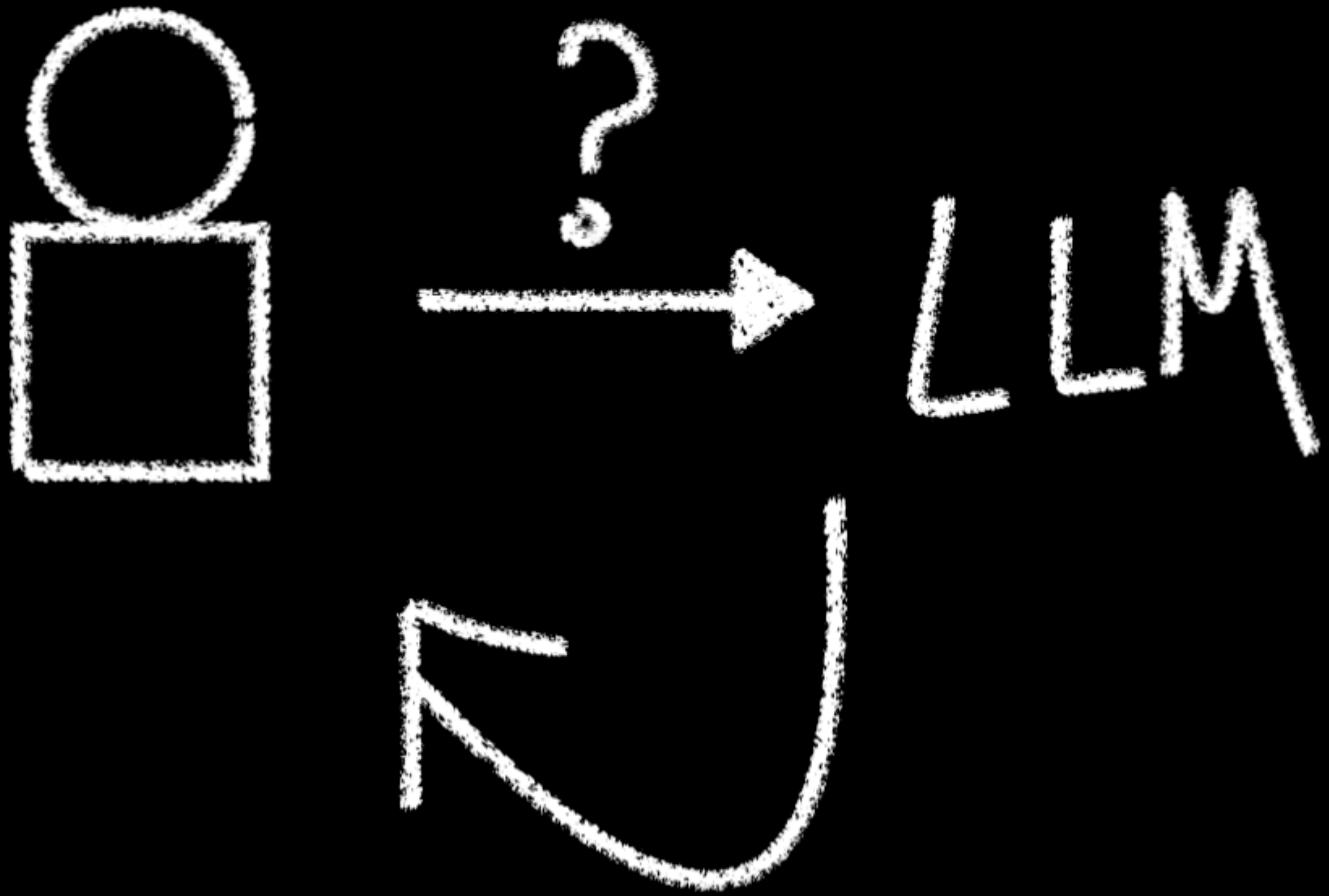

We used images of Holly wearing this costume to fine-tune an image model, and that model was recursively refined to produce a consistent character that is able to be spawned by anyone using the interface provided. This model can produce infinite images of this new character. The images produced by this model will mostly all, in some way, be infected by the hairy mutant.

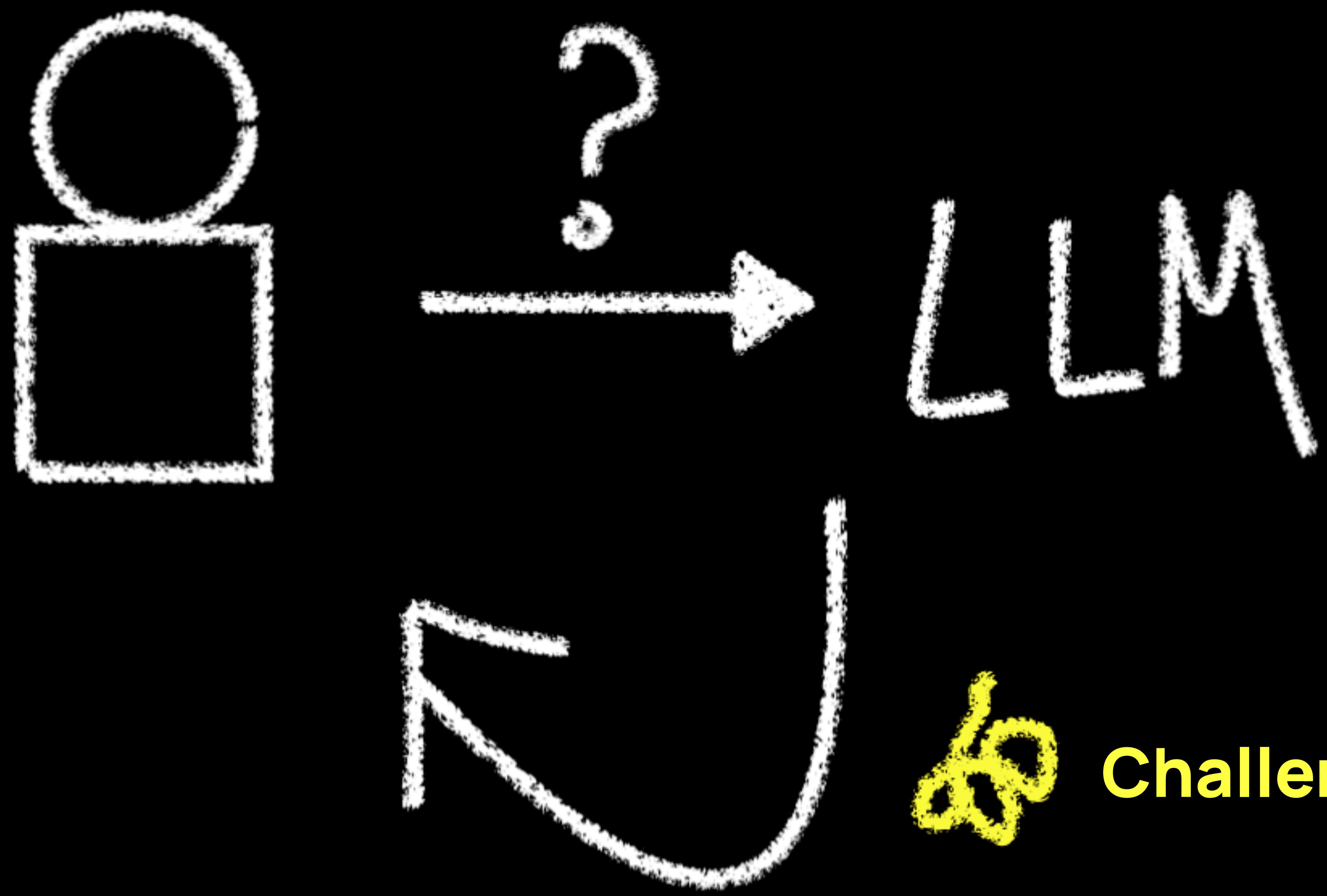
https://xhairymutantx.whitney.org/



These models are trained on the open Internet,
Discard
Submit ↑

RAG

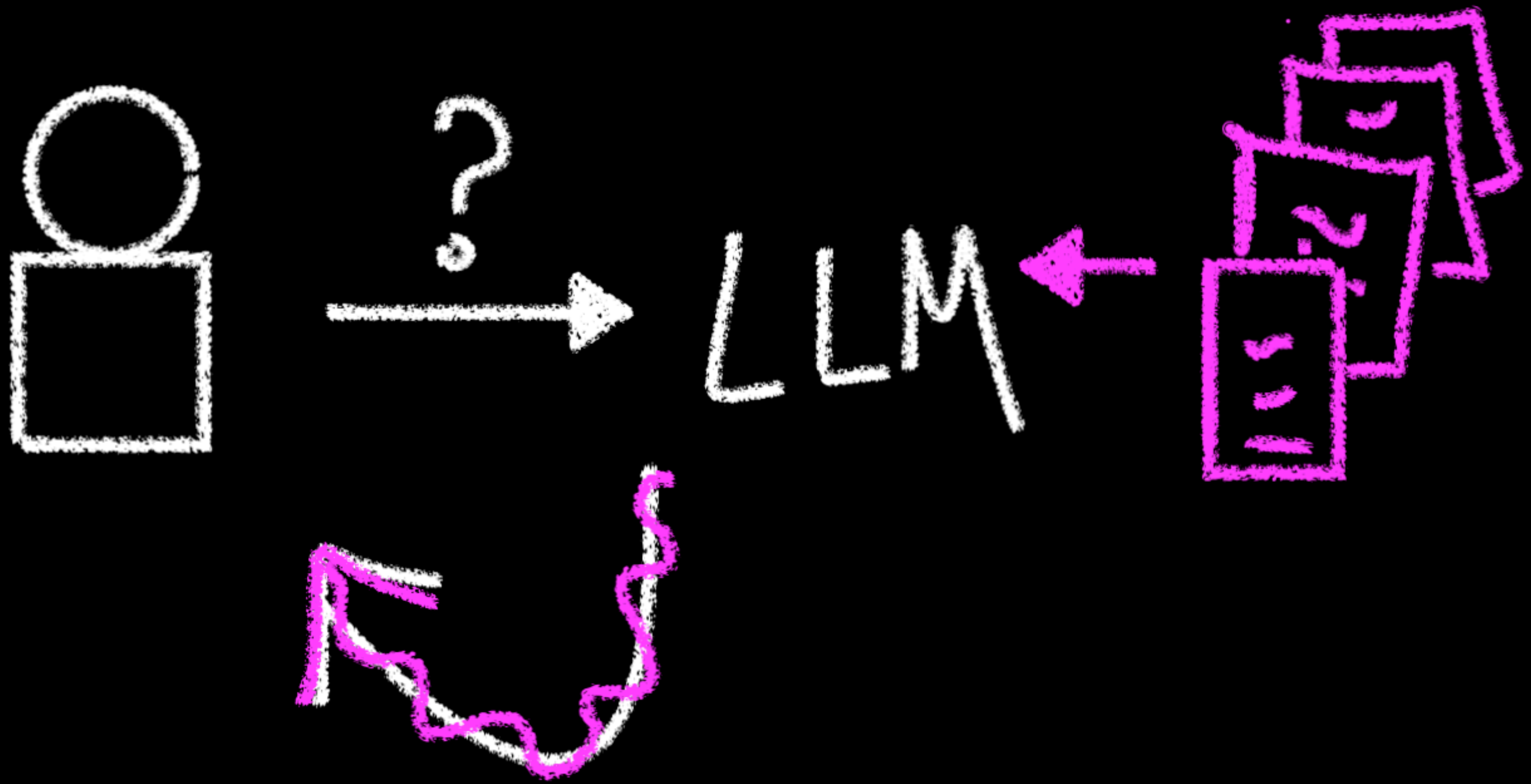

?
LLM



Challenges

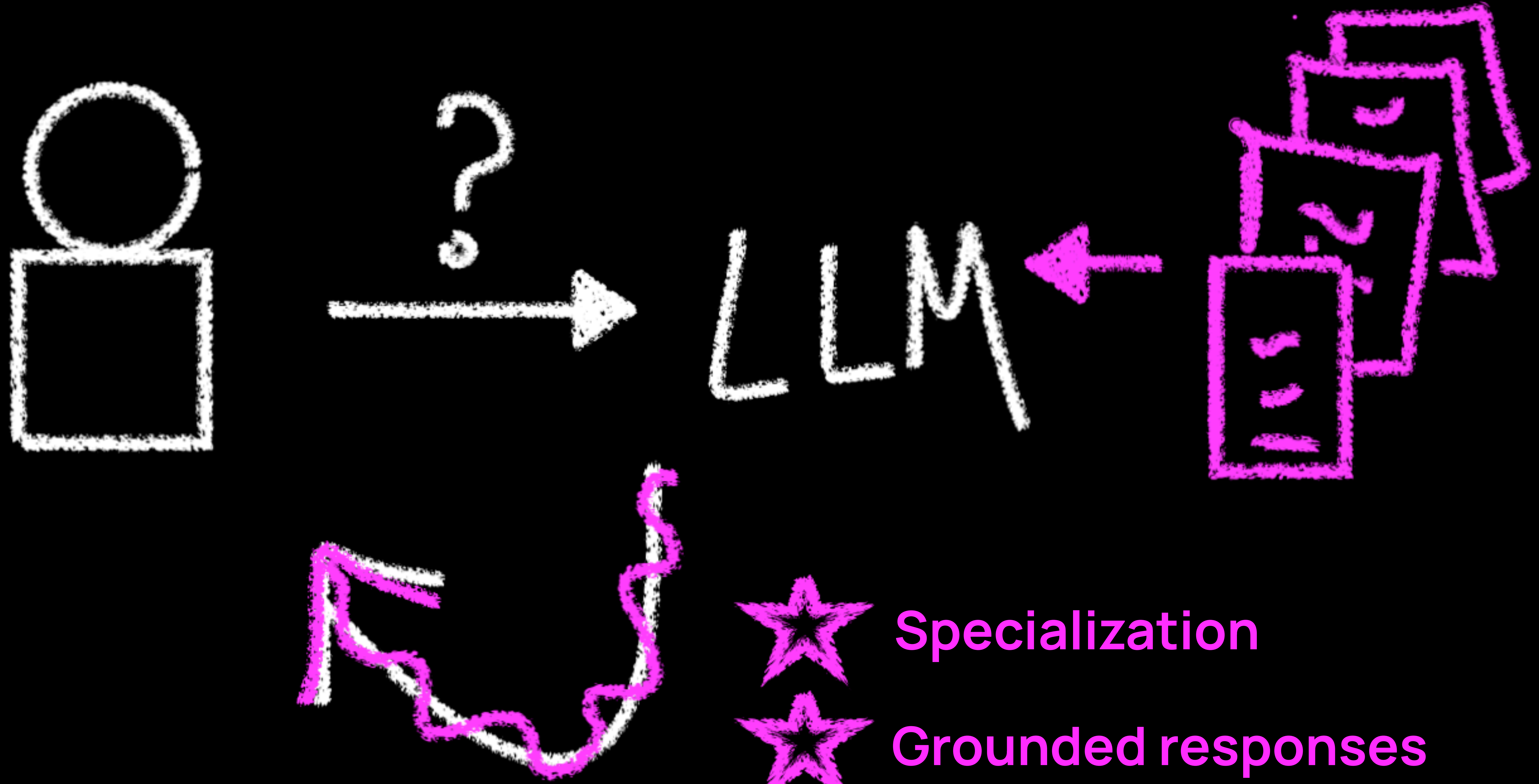
Retrieval-Augmented Generation (RAG)

Knowledge Base



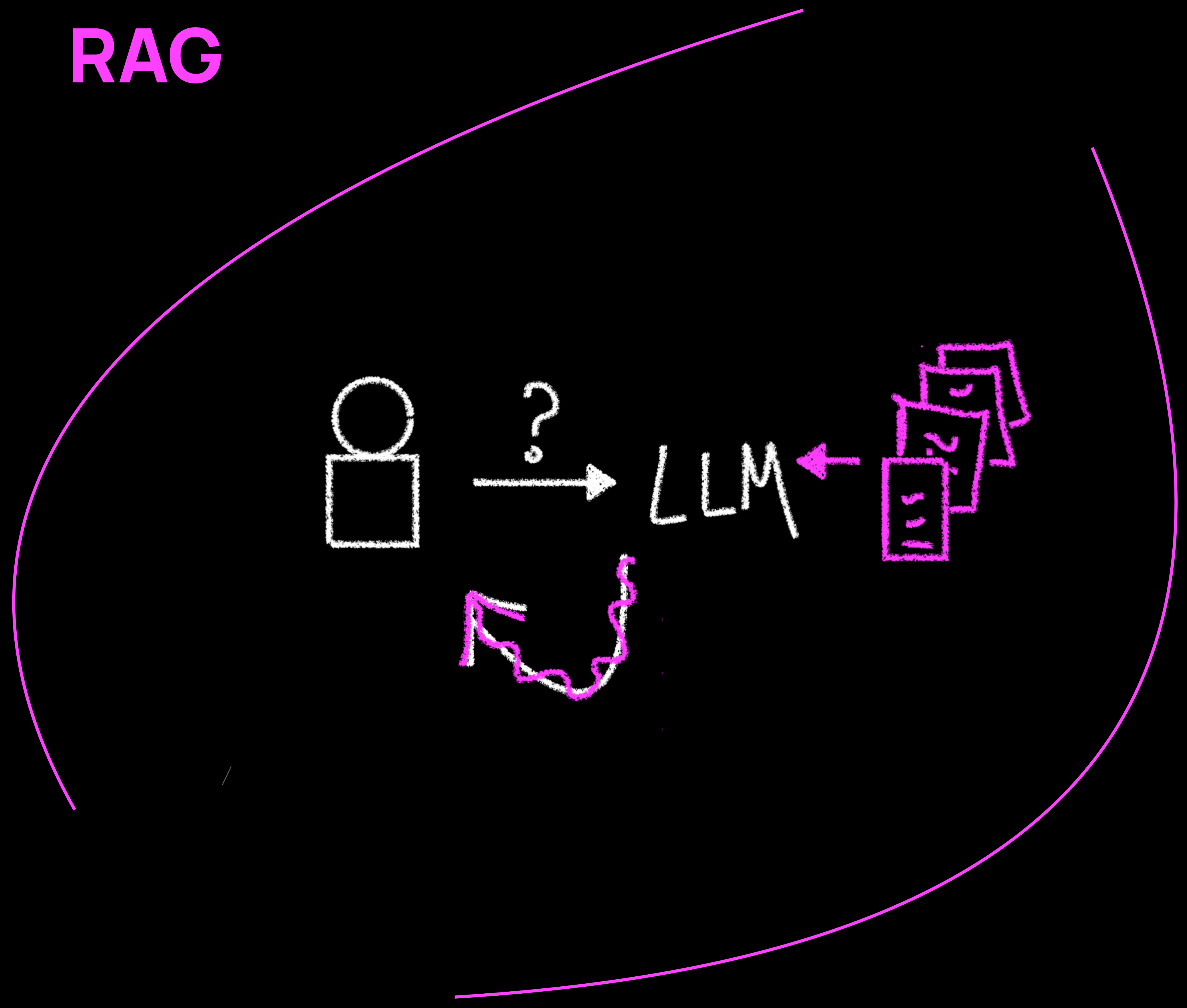
Retrieval-Augmented Generation (RAG)

Knowledge Base



- Specialization
- Grounded responses
- Less hallucination

RAG



your own continuously updatable, locally running mini-LLM, shaped by the collection of documents you curate

Artistic Chatbot at Warsaw Academy of Fine Arts

RAG

Artistic Chatbot: a voice-to-voice RAG powered chat system

"The question answering (QA) chatbot responded to free-form spoken questions in Polish using the context retrieved from a curated, domain-specific knowledge base consisting of 226 documents provided by the organizers, including faculty information, art magazines, books, and journals."

by Filip J. Kucia, Bartosz Grabek, Szymon D. Trochimiak, and Anna Wróblewska
At Warsaw Academy of Fine Arts https://arxiv.org/pdf/2509.00572

Case Study of *Artistic Chatbot*

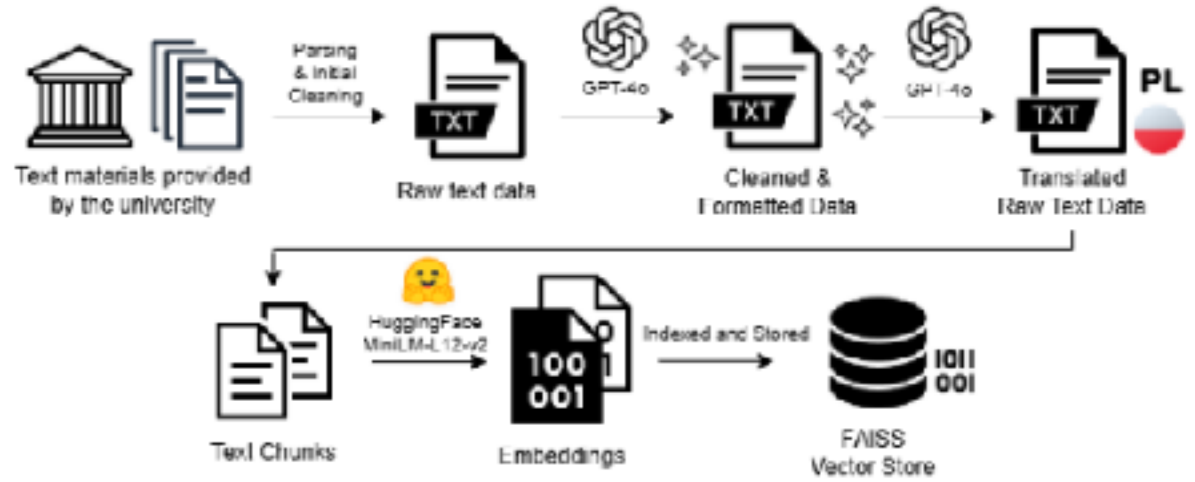


Figure 1: Data Preprocessing Pipeline

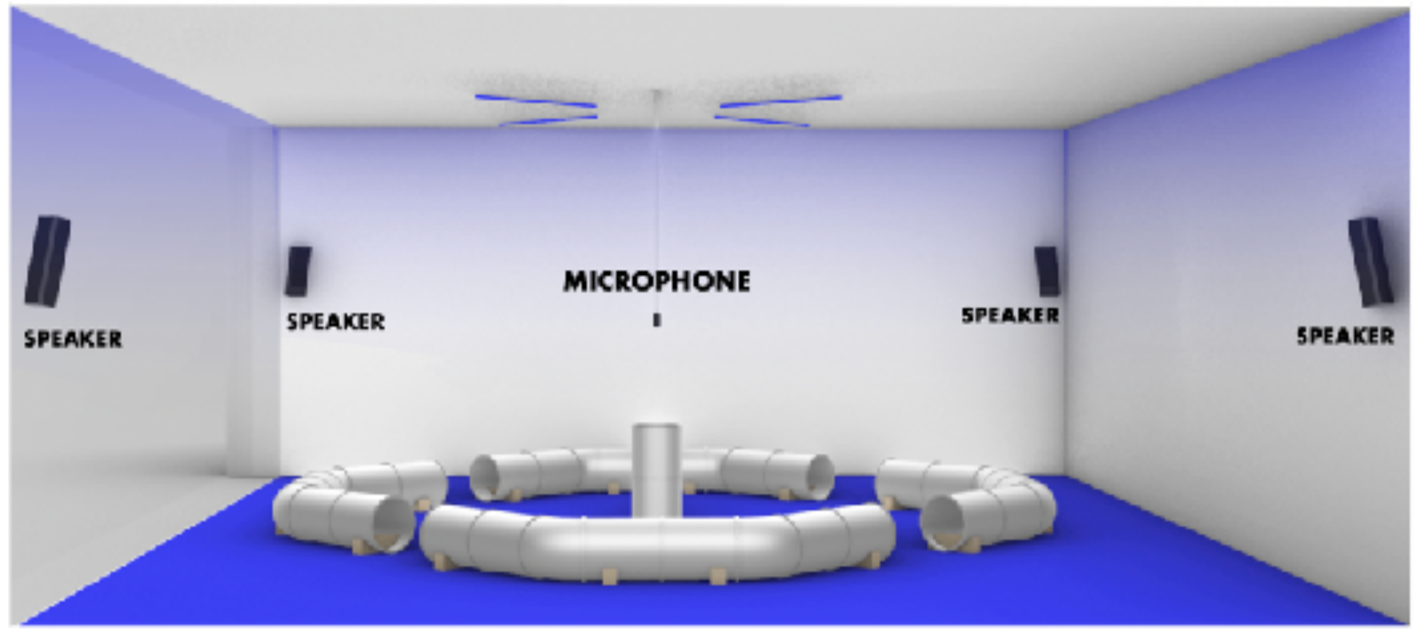


Figure 3: Chatbot Physical Setup

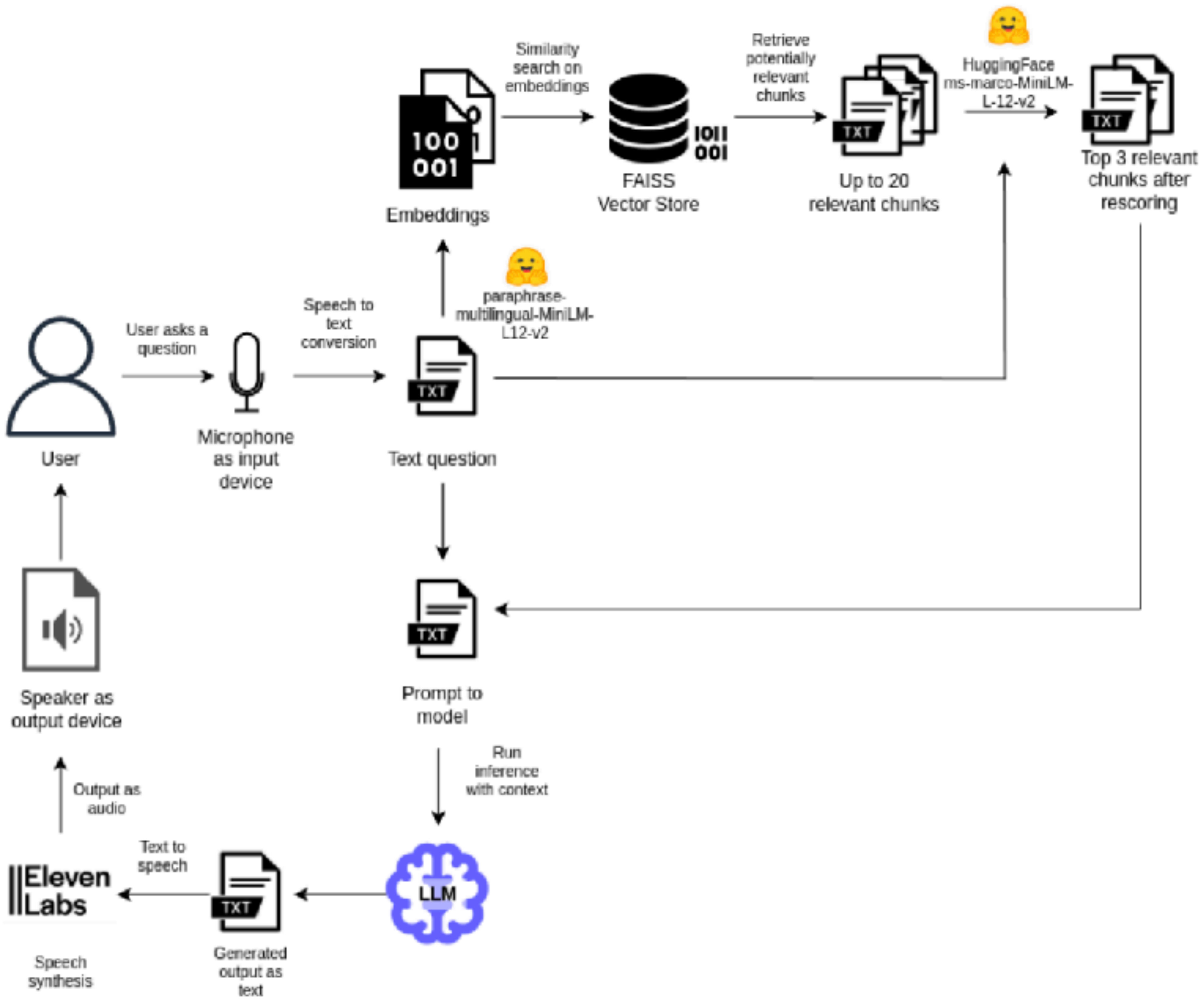


Figure 2: Inference Pipeline

by Filip J. Kucia, Bartosz Grabek, Szymon D. Trochimiak, and Anna Wróblewska
At Warsaw Academy of Fine Arts

https://arxiv.org/pdf/2509.00572

Artistic Chatbot adopts the persona of an artificial art curator, a role that involves responding to questions while simultaneously assessing their relevance to the exhibition.

Chatbots such as Artistic Chatbot effectively maintain responses grounded in exhibition content (60% of responses directly relevant), even when faced with unpredictable queries outside the target domain.

by Filip J. Kucia, Bartosz Grabek, Szymon D. Trochimiak, and Anna Wróblewska
At Warsaw Academy of Fine Arts https://arxiv.org/pdf/2509.00572

AI Blob!

RAG

AI Blob!: an experimental system that uses the capabilities of Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) in television archives

processes a curated dataset of 1,547 Italian television videos

by Roberto Balestri
at Università di Bologna
https://www.arxiv.org/pdf/2508.09535

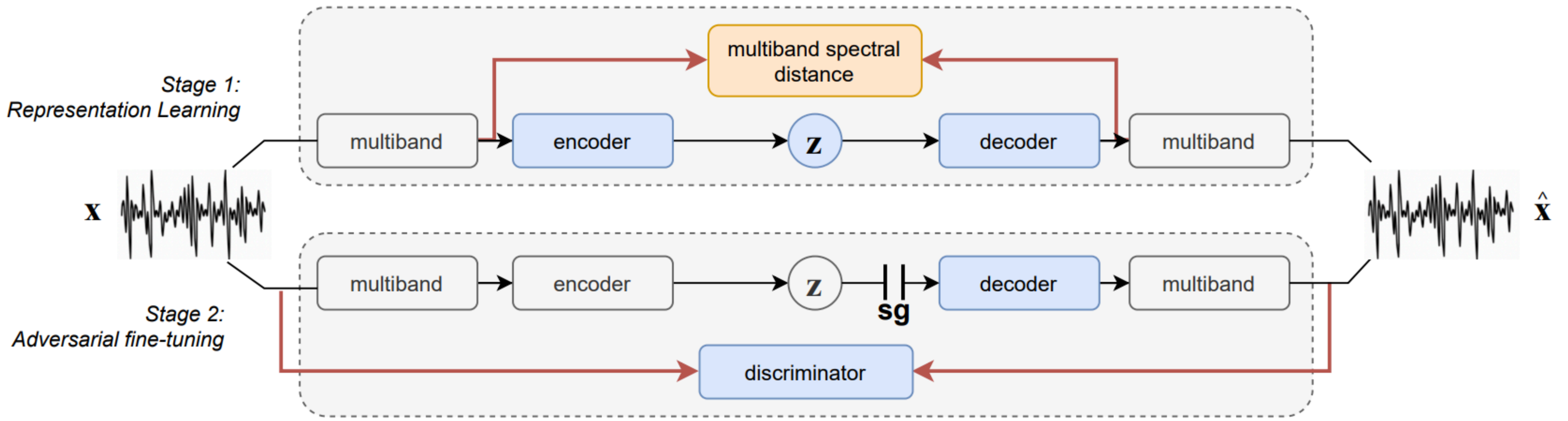

Rai Edu 2
REWIND
Altro che si prendono certe libertà con le mie
carni, come se avessero pagato vitto e alloggio.

“By foregrounding dynamic, content-aware retrieval over static metadata schemas, AI Blob! demonstrates how semantic technologies can facilitate new approaches to archival engagement, enabling novel forms of automated narrative construction and cultural analysis.”

by Roberto Balestri
at Università di Bologna

https://www.arxiv.org/pdf/2508.09535

RAVE



RAVE (Realtime Audio Variational autoEncoder) is an auto-encoder for sound: it takes sound as input and is trained to reconstruct that sound as output.

By Antoine Caillon & Philippe Esling, IRCAM - Sorbonne Universite

https://arxiv.org/abs/2111.05011

RAVE_anonymous

View the Project on GitHub
anonymous84654/RAVE_anonymous

RAVE: A variational autoencoder for fast and high-quality neural audio synthesis

Abstract: Deep generative models applied to audio have improved by a large margin the state-of-the-art in many speech and music related tasks. However, as raw waveform modelling remains an inherently difficult task, audio generative models are either computationally intensive, rely on low sampling rates, are complicated to control or restrict the nature of possible signals. Among those models, Variational AutoEncoders (VAE) give control over the generation by exposing latent variables, although they usually suffer from low synthesis quality. In this paper, we introduce a Realtime Audio Variational autoEncoder (RAVE) allowing both fast and high-quality audio waveform synthesis. We introduce a novel two-stage training procedure, namely representation learning and adversarial fine-tuning. We show that using a post-training analysis of the latent space allows a direct control between the reconstruction fidelity and the representation compactness. By leveraging a multi-band decomposition of the raw waveform, we show that our model is the first able to generate 48kHz audio signals, while simultaneously running 20 times faster than real-time on a standard laptop CPU. We evaluate synthesis quality using both quantitative and qualitative subjective experiments and show the superiority of our approach compared to existing models. Finally, we present applications of our model for timbre transfer and signal compression. All of our source code and audio examples are publicly available.

Timbre transfer

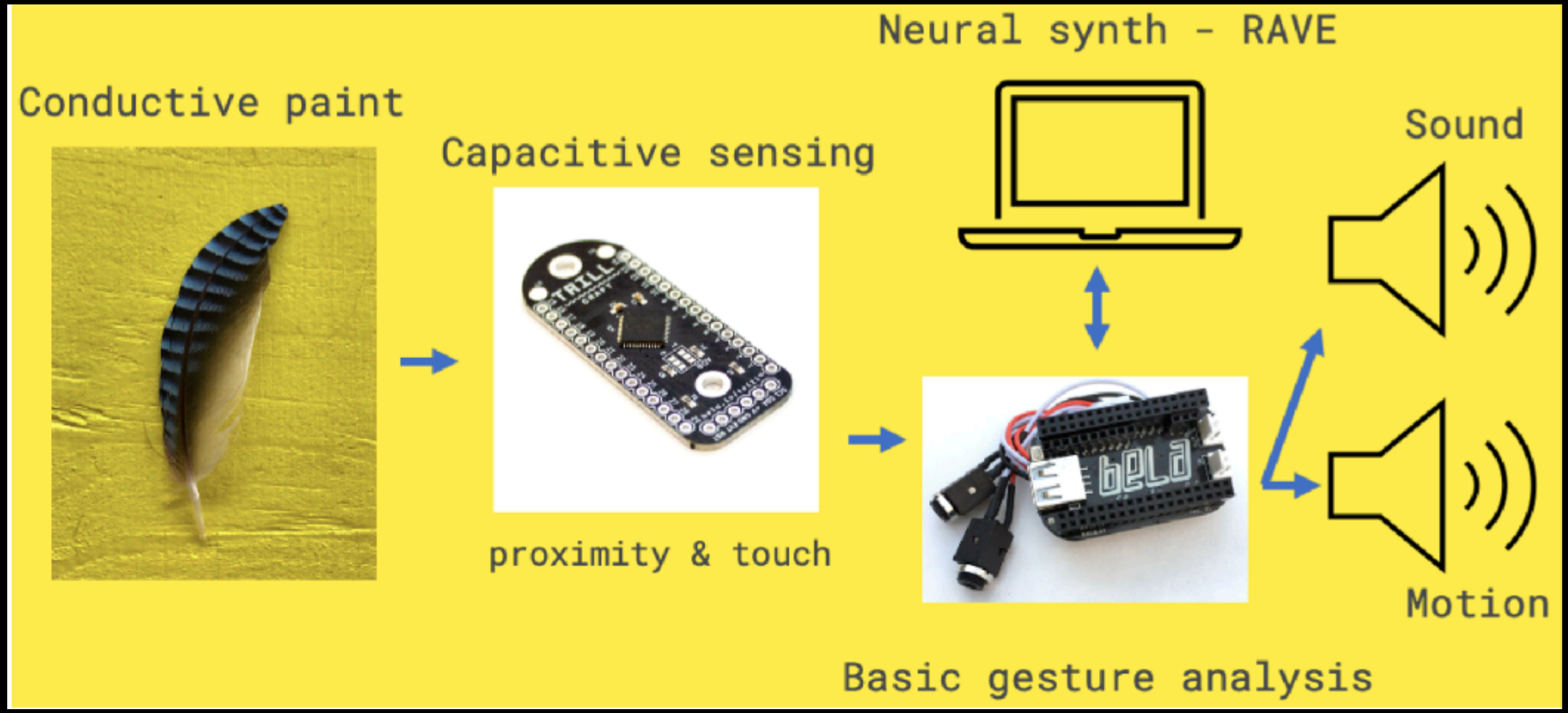
Given the high compression applied to the input waveform when encoded into a latent representation, we demonstrate that RAVE can be used to perform timbre transfer.

Strings to speech transfer


play_v3
15.
-0.574
-0.89
0.
sig~
1.146
-0.287
-6.1
0.12
0.364
0.259
-0.351
0.354
0.
-0.55
1.43
-0.02
Soundinit-Soundwalk-2021-07-18.wav
Soundinit-soundwalk-2021-08-21.wav
Soundinit-Soundwalk-2021-09-13.wav
nn~ fr1_e18d54798e_streaming_EPOCH-609.ts encode
0.646
nn~ fr1_e18d54798e_streaming_EPOCH-609.ts decode
Gain
5.5 dB
Gain
-2.8 dB
dac~

Pluma

RAVE



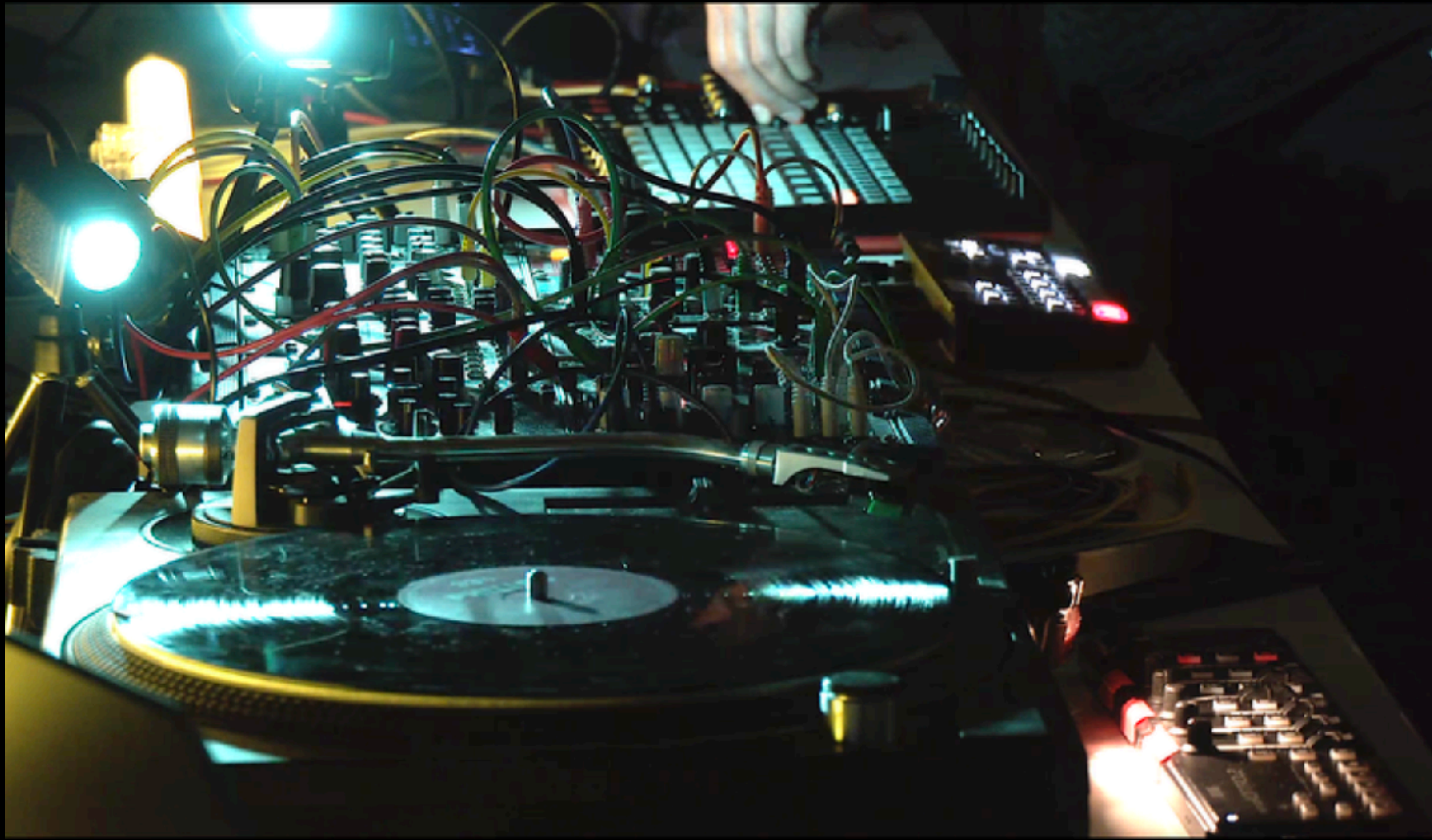
Concept, design & sound: Giacomo Lepri
CNC fabrication: Halldór Úlfarsson
Audio neural synthesis: Victor Shepardson
https://www.giacomolepri.com/pluma

Developed at the Intelligent Instruments Lab as part of the EU ERC INTENT project.

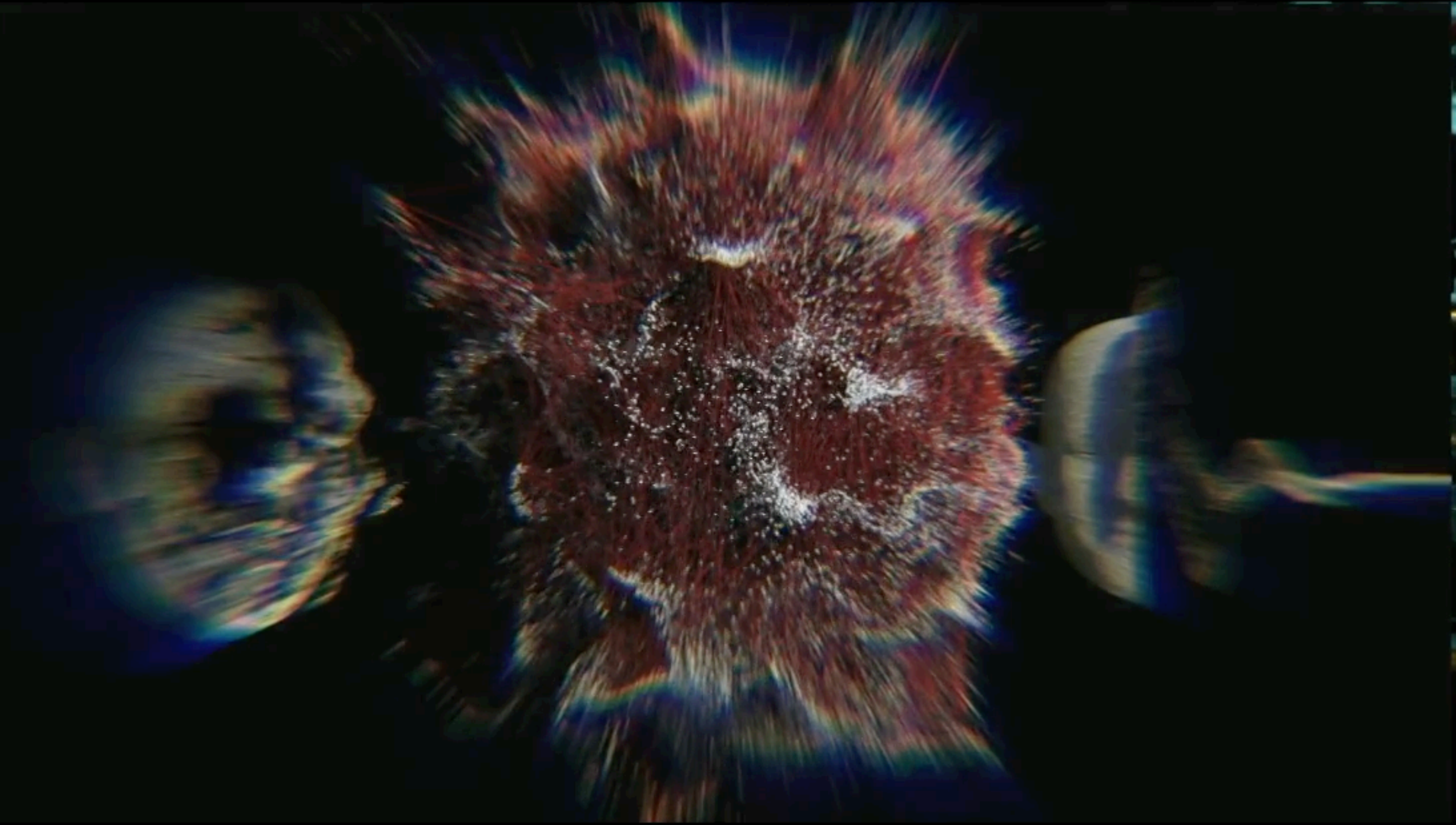
New Ruins

RAVE

“Their utilization of RAVE real-time timbre morphing models and the Somax2 machine listening and improvisation system developed by IRCAM adds organic and pleasantly unpredictable elements to what is otherwise a conventional setup using modular synthesizers and a turntable. The artist’s physical interaction with these devices paradoxically highlights the hidden presence of machine intelligence. “



by Abo Abo (Daniele Carcassi) and Tania Cortés Becerra
https://calls.ars.electronica.art/2025/prix/winners/17320/




Models

Classifier

Generative adversarial network

- FastGAN
- Pix2Pix (conditonal GAN)

Fine Tuning

RAG

RAVE

Models & Histories

1950s–1980s
(Perceptron, Decision Trees)

AlexNet wins ImageNet (2012), showing deep neural nets outperform classical classifiers.

Classifier

2014 (Goodfellow et al.)
DCGAN (2015) → Pix2Pix (2016) → StyleGAN (2018) → fuels art/A.I. boom.

Generative adversarial network

FastGAN

Pix2Pix
(conditonal GAN)

2006 (early neural nets)
→ 2018+ (Transfer Learning)

Fine-tuning BERT (2018) & GPT-2 (2019) becomes standard for adapting large models to specific tasks. Artists start fine-tuning StyleGAN and GPT models.

Fine Tuning

RAG

2020

RAVE

2024

Models

Classifier

Generative adversarial network

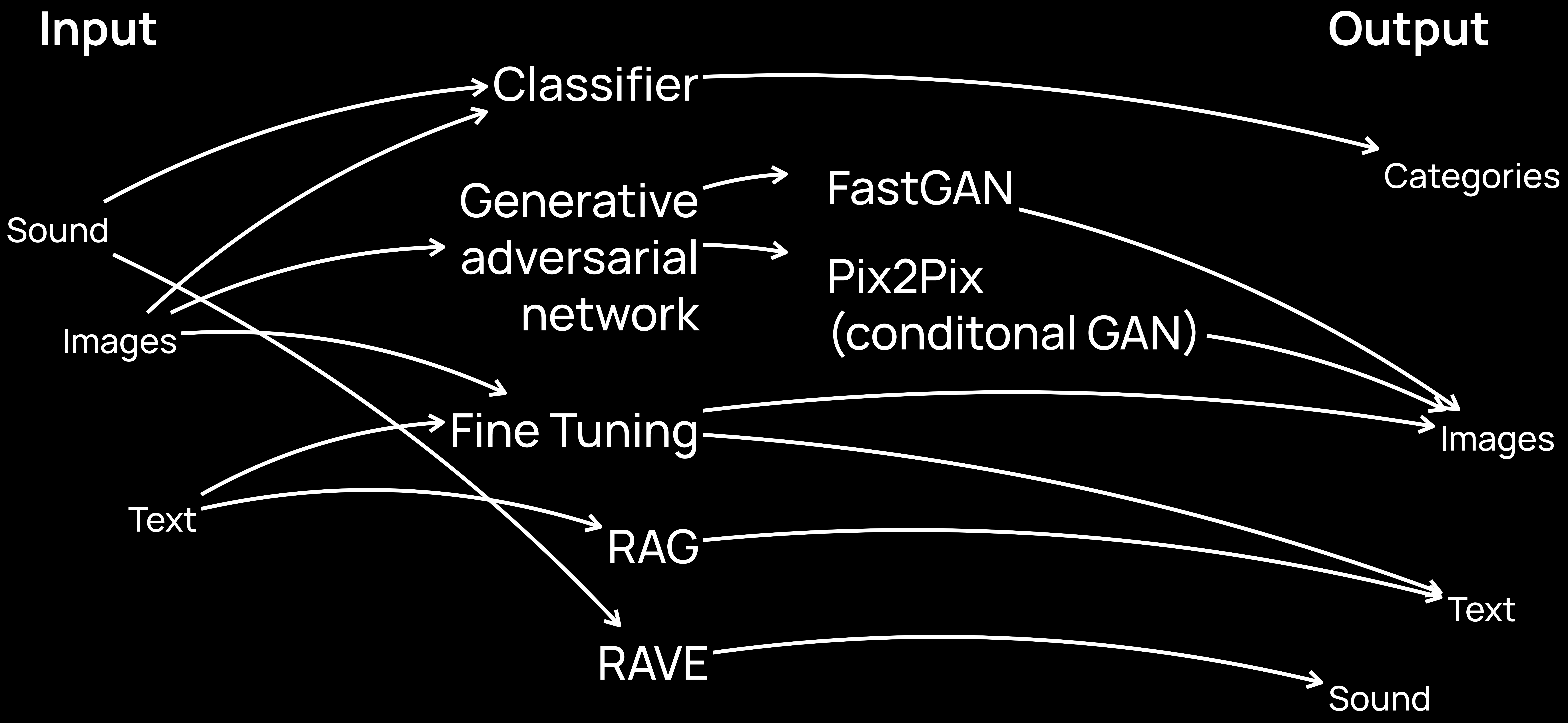
- FastGAN
- Pix2Pix (conditonal GAN)

Fine Tuning

RAG

RAVE

Models


Input
Output
Classifier
Categories
Generative
adversarial
network
FastGAN
Pix2Pix
(conditonal GAN)
Sound
Images
Fine Tuning
Images
Text
RAG
Text
RAVE
Sound

Goals

Pick one model, collect a dataset, train from scratch or near-scratch, reflect on results.

Understand AI as a material and cultural process, not a black box.

Goals

Experiments over outcomes

Strangeness over perfection

Fun over Exhaustion

Workshop

Workshop

Monday 13 – 18h

Part 1

- 4 Groups (Classification, Pix2Pix, FastGAN, FineTuning, …)
- 30 minutes to research and develop ideas together
- Presenting ideas
- Mixing into new Groups

Part 2

After 4 workshop rounds together we decide on the best project ideas for the week.

Part 3

- Pick Project for the week
- Start Working