

# ImageGen

# Program

## Overview

- Intro
  - Disclaimer
  - Research Question
  - The Plan
  - Short history of Image generation
- Status Quo
  - Concept
  - Techniques
- Why we create?
  - Motivations
  - Means
- How we create?
  - Steps
  - Creative Fools
  - To the metal
  - Workflow
  - Rules
- The bigger picture
  - Automation
  - Future?
  - Pros / Cons
- Learning... Why?
  - Discussion

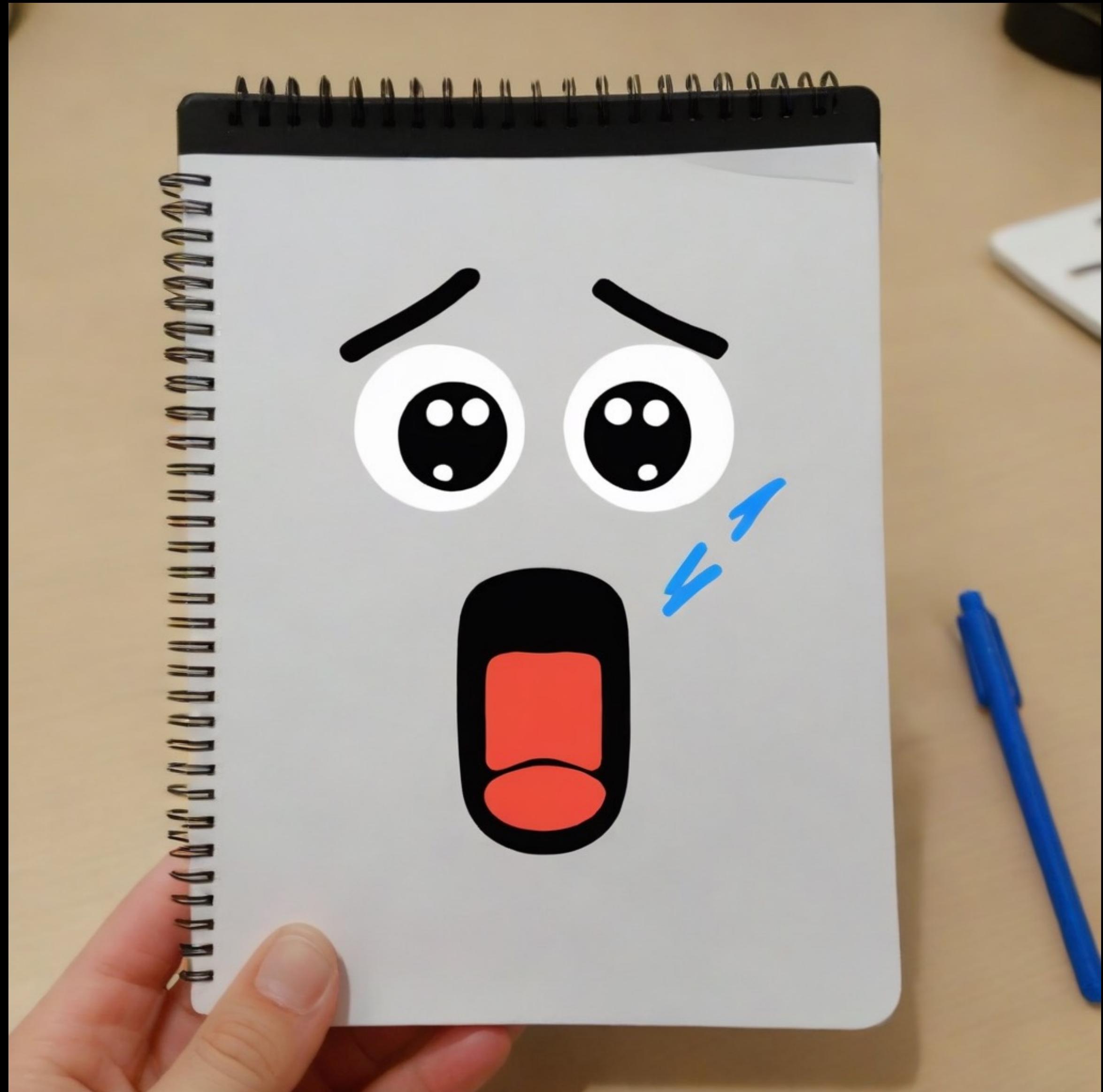
→ Intro

## Introduction - Disclaimer

- Focus on Image Gen.
- Focus on digital
- Findings are based on my own observation.
- We will consider image generative models as *fair use* for this presentation.

Keeping the motivation to learn  
in the age of one click image generation

## Introduction - The plan



→ “The problem with AI right now isn't that it's smart,  
it's that it's stupid in ways we can't always predict.”

*John Oliver*

Pretend I'm  
not even here.



# Introduction - Short history of Image generation - First Neural "AI" Art 2015-2021



## Introduction - Short history of Image generation - First Neural "AI" Art 2015-2021



## Introduction - Short history of Image generation - First Neural "AI" Art 2015-2021

Coarse styles  
 $(4^2 - 8^2)$



Middle styles  
 $(16^2 - 32^2)$



Fine styles  
 $(64^2 - 1024^2)$



## Introduction - Short history of Image generation - Second Wave (2022-)



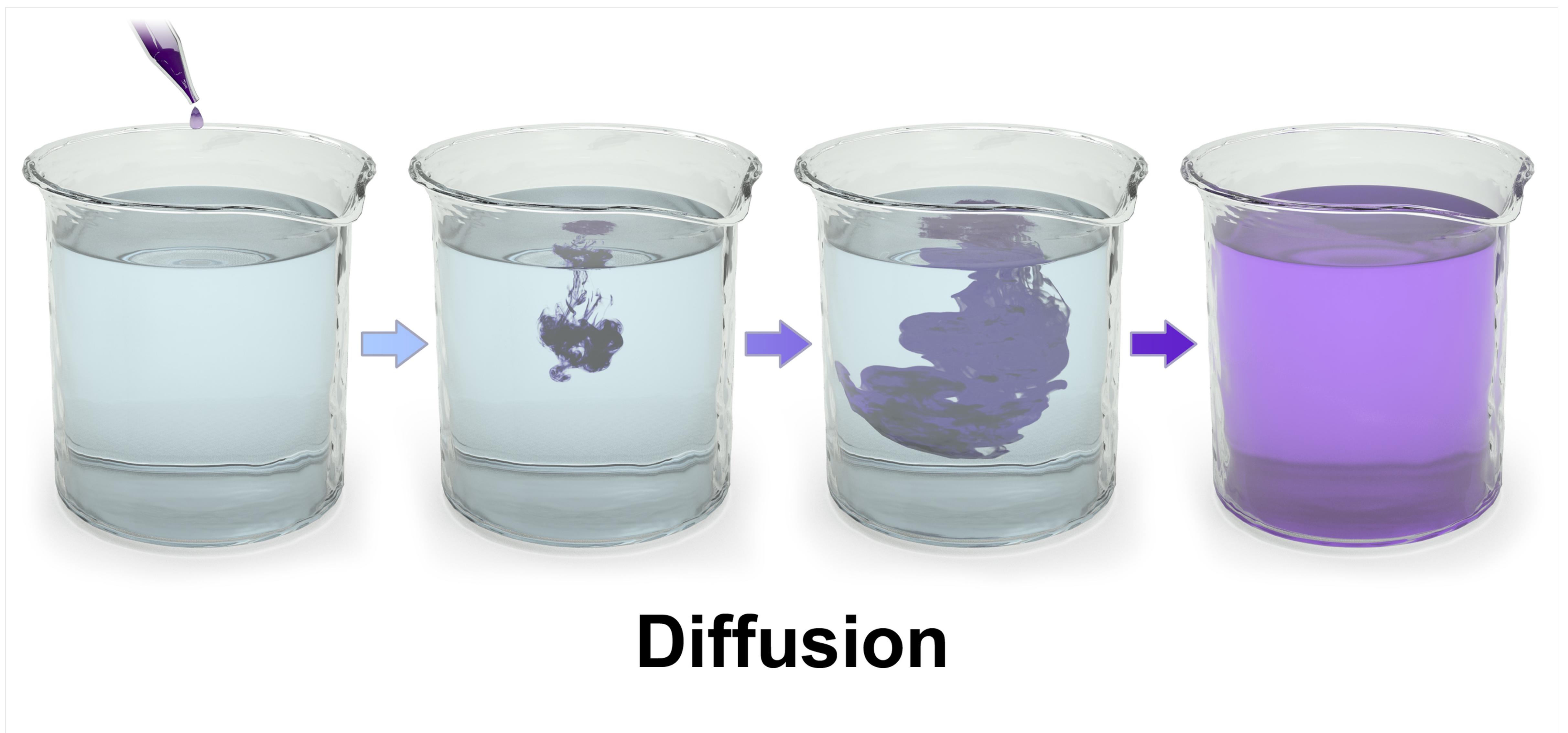
## Introduction - Short history of Image generation - Second Wave (2022-)



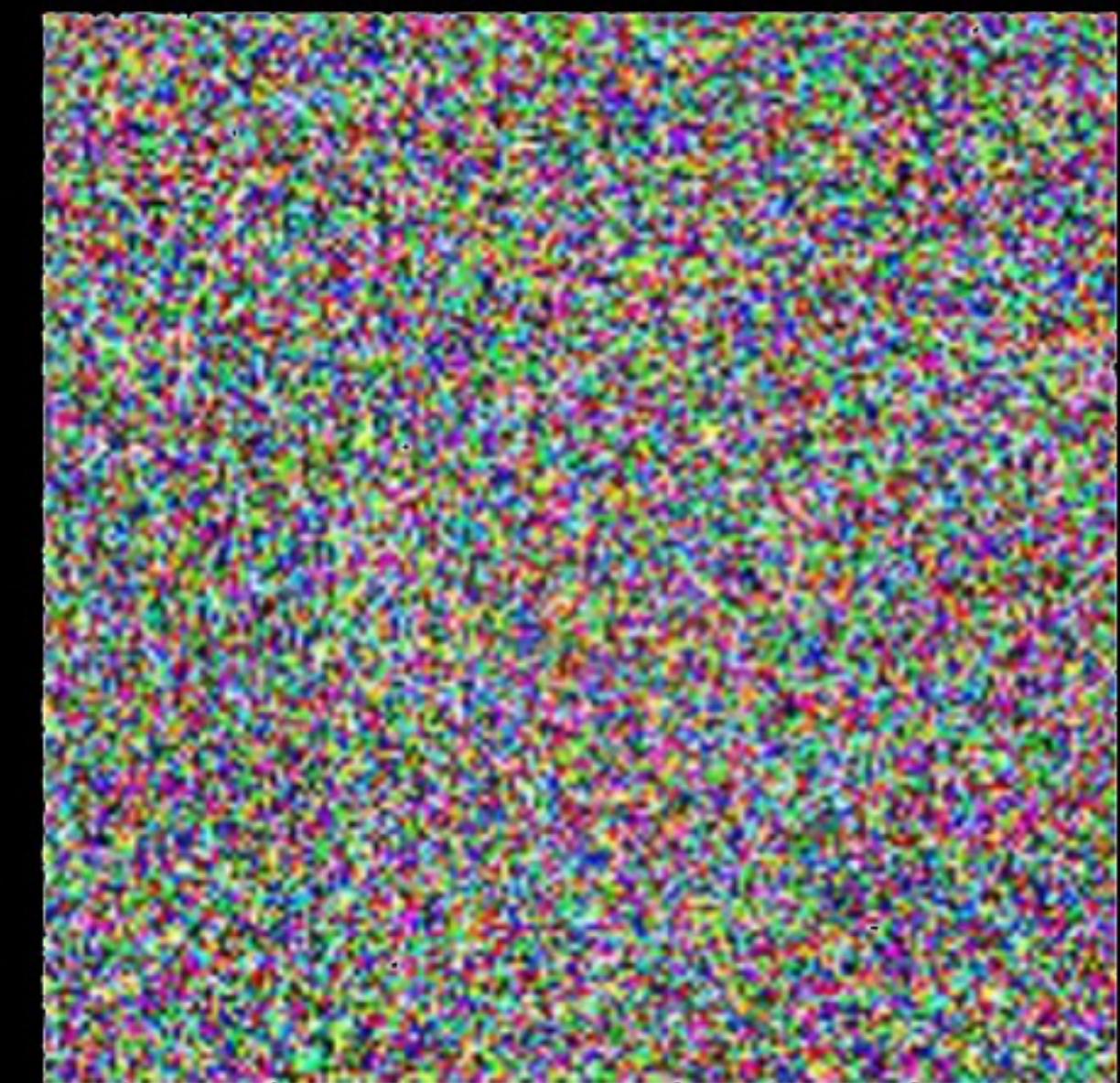
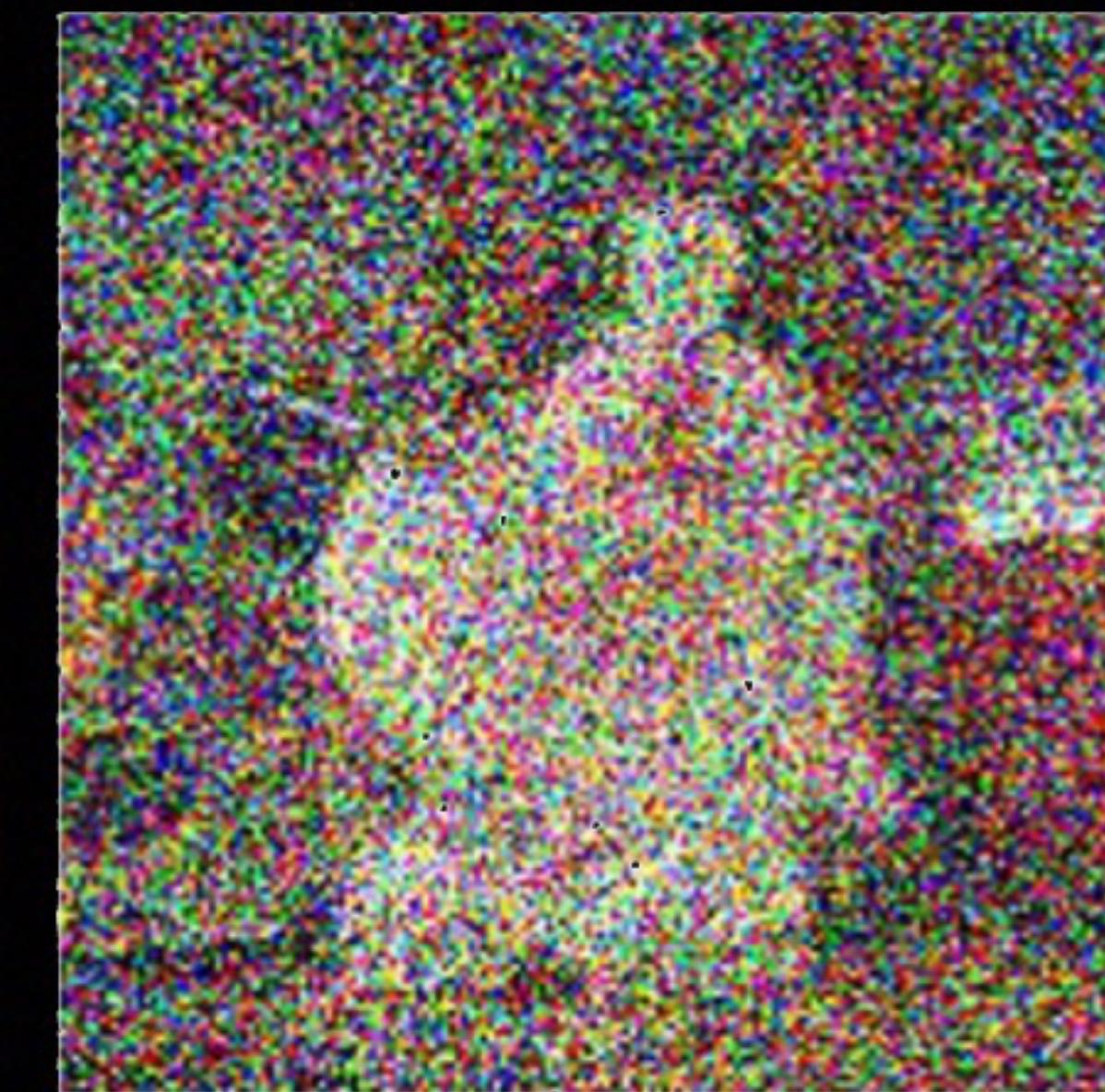
→ Status Quo

## Status Quo - Concepts - Diffusion





## Status Quo - Concepts - Diffusion



Step

1



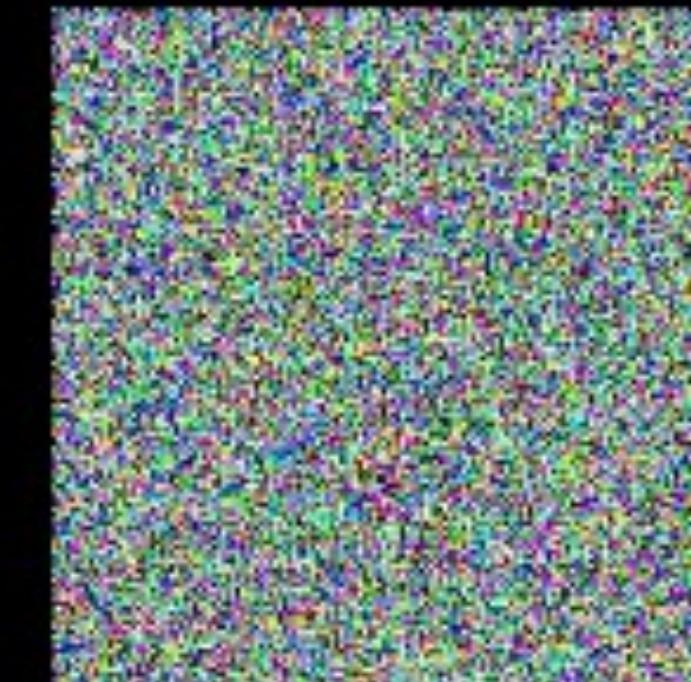
2



3



4

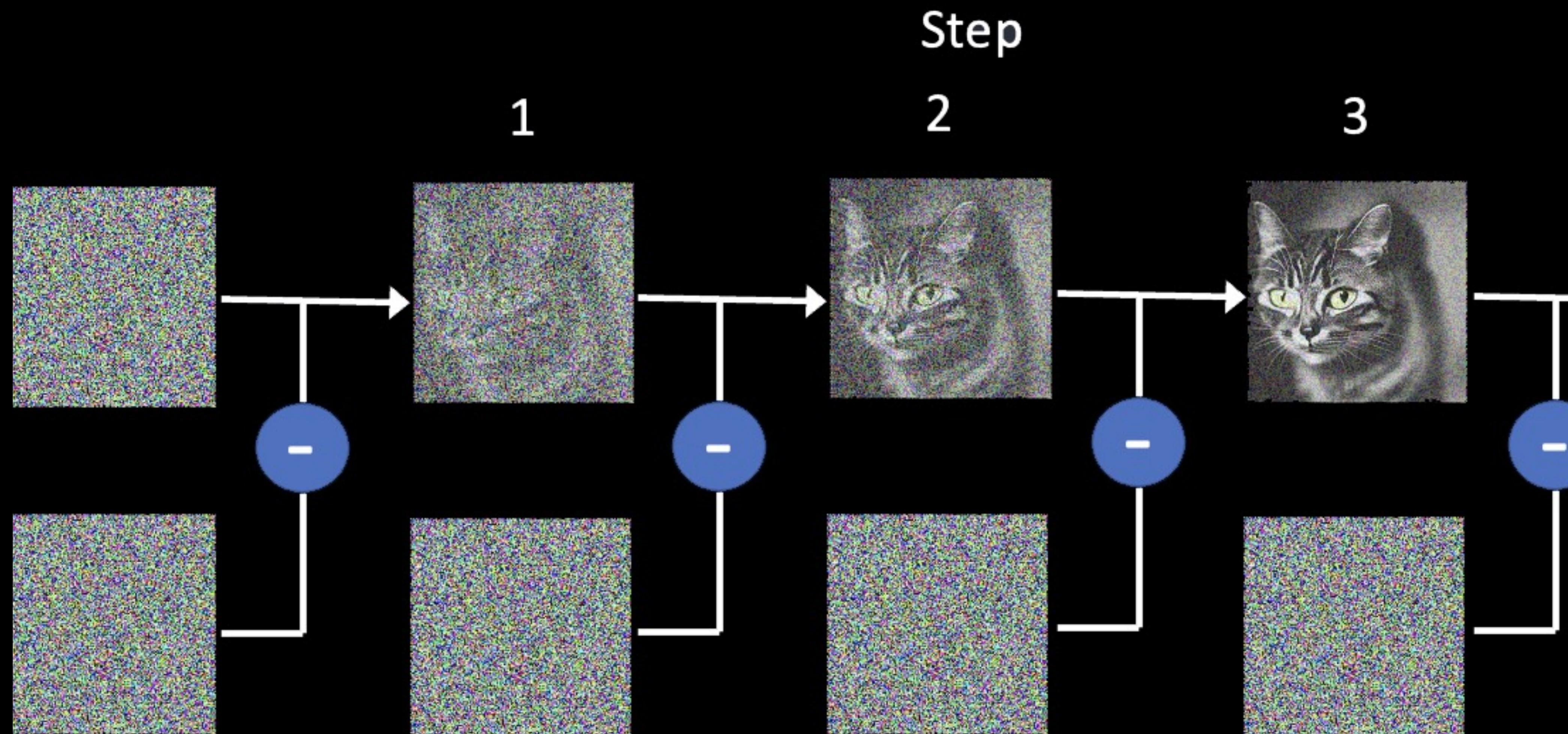


Noise 1

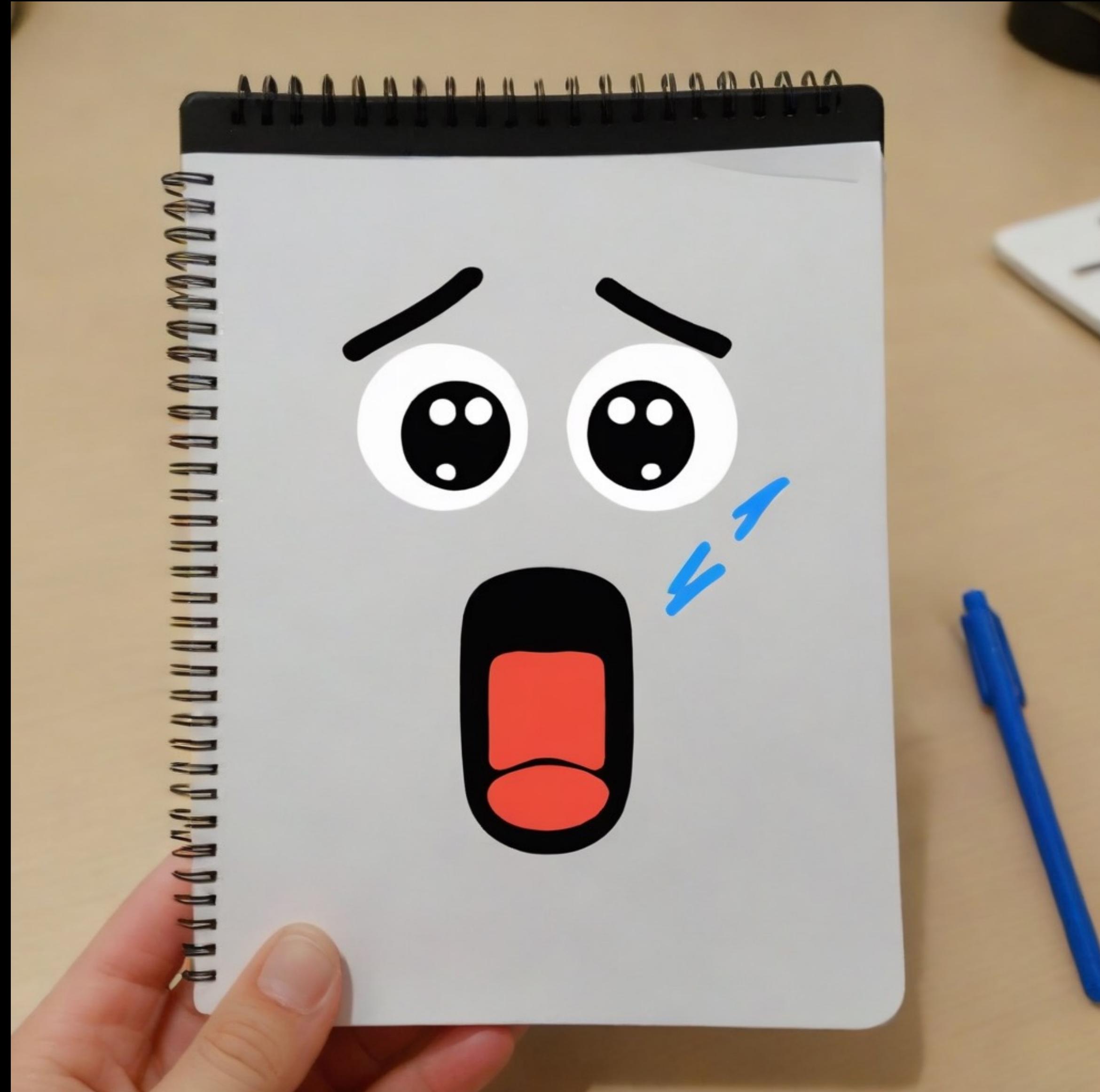
Noise 2

Noise 3

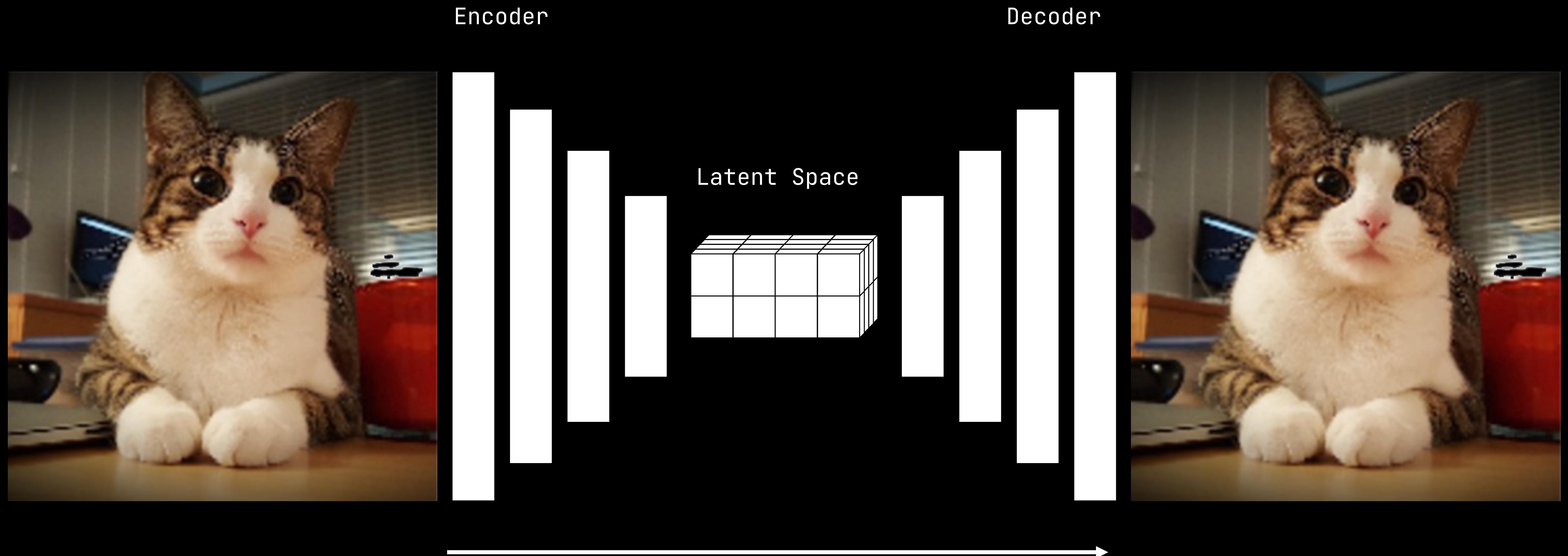
Noise 4



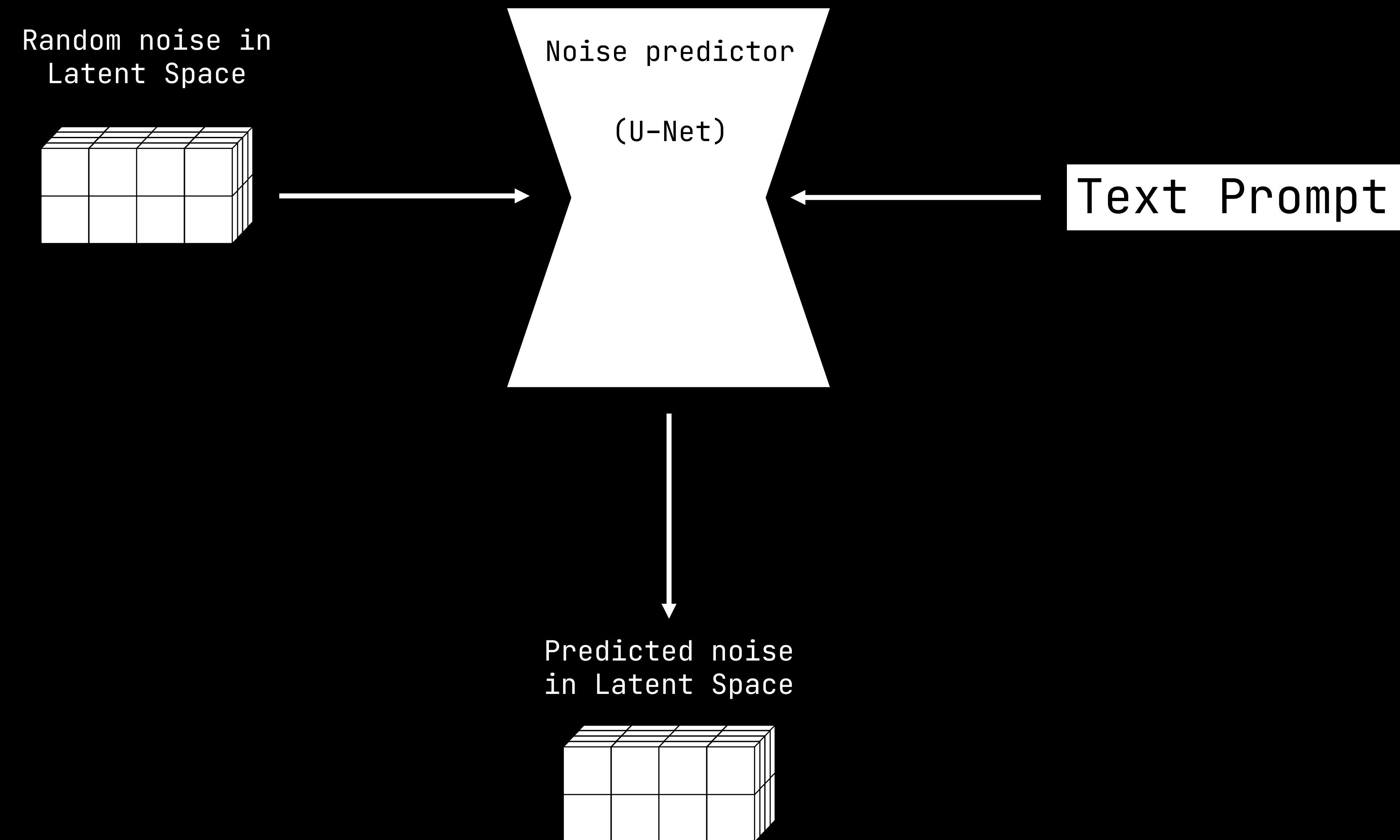
# Status Quo - Concepts - Diffusion



## Status Quo - Techniques - Encoding



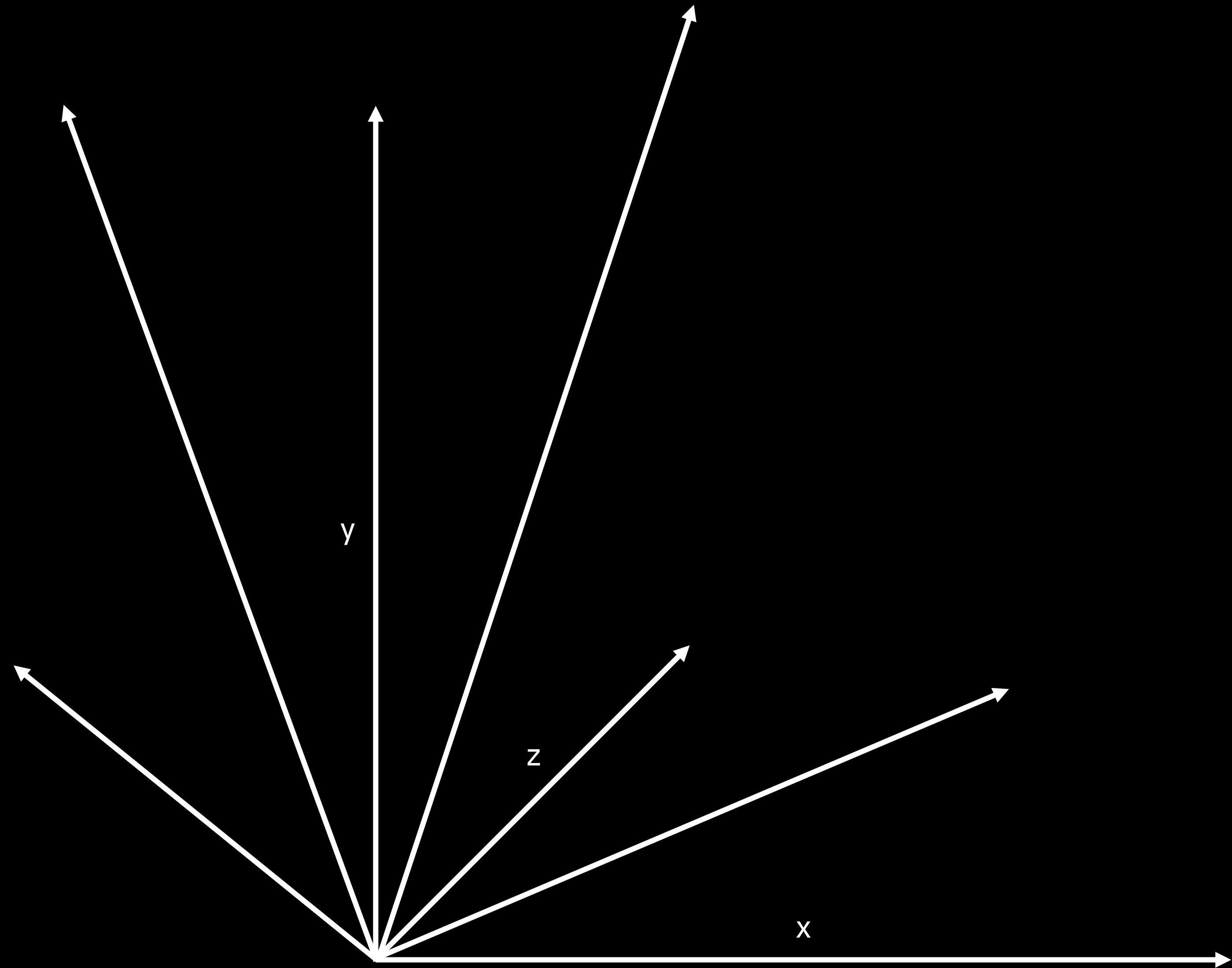
## Status Quo - Techniques - Text conditioning



## Status Quo - Techniques - Diffusion



## Status Quo - Techniques - input - Latent Space (dimensionality)





- “The perceptual relativity argument is that different objects can appear to have different characteristics (e.g. shape) depending on the observer's perspective”

*George Berkeley*

- Encoding
- Embedding
- Retrieval
- *Compression*

## Status Quo - Concepts - Capsule



# Status Quo - Concepts

## Food101

**guacamole** (90.1%) Ranked 1 out of 101 labels



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

## SUN397

**television studio** (90.2%) Ranked 1 out of 397 labels



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

## Youtube-BB

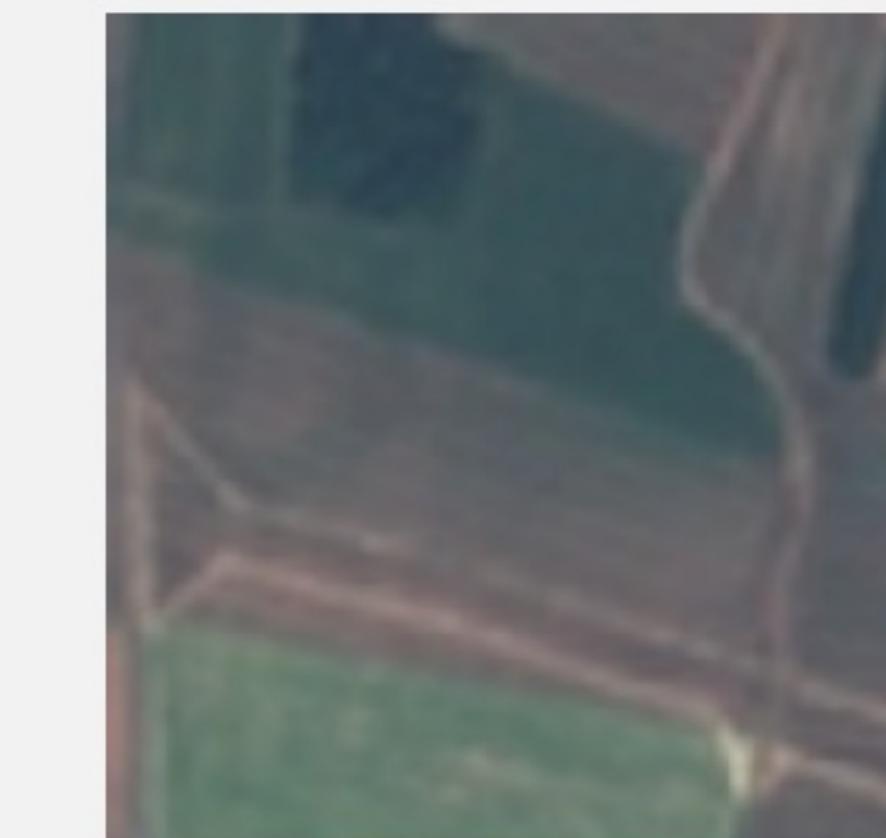
**airplane, person** (89.0%) Ranked 1 out of 23 labels



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

## EuroSAT

**annual crop land** (46.5%) Ranked 4 out of 10 labels



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

txt2img

img2img

...2img

...2...

Status Quo - Concept - Focus on Image Generation



→ Why we create

- Emulation
- Recognition
- Expression
- Be part of something
- Process
- PASSION





- Access
- Tools
- Inspiration
- Education
- Skills?

→ “Creativity is intelligence having fun.”

*Albert Einstein*

→ How we create

How we create

NOBODY  
STARTS  
OUT  
ORIGINAL

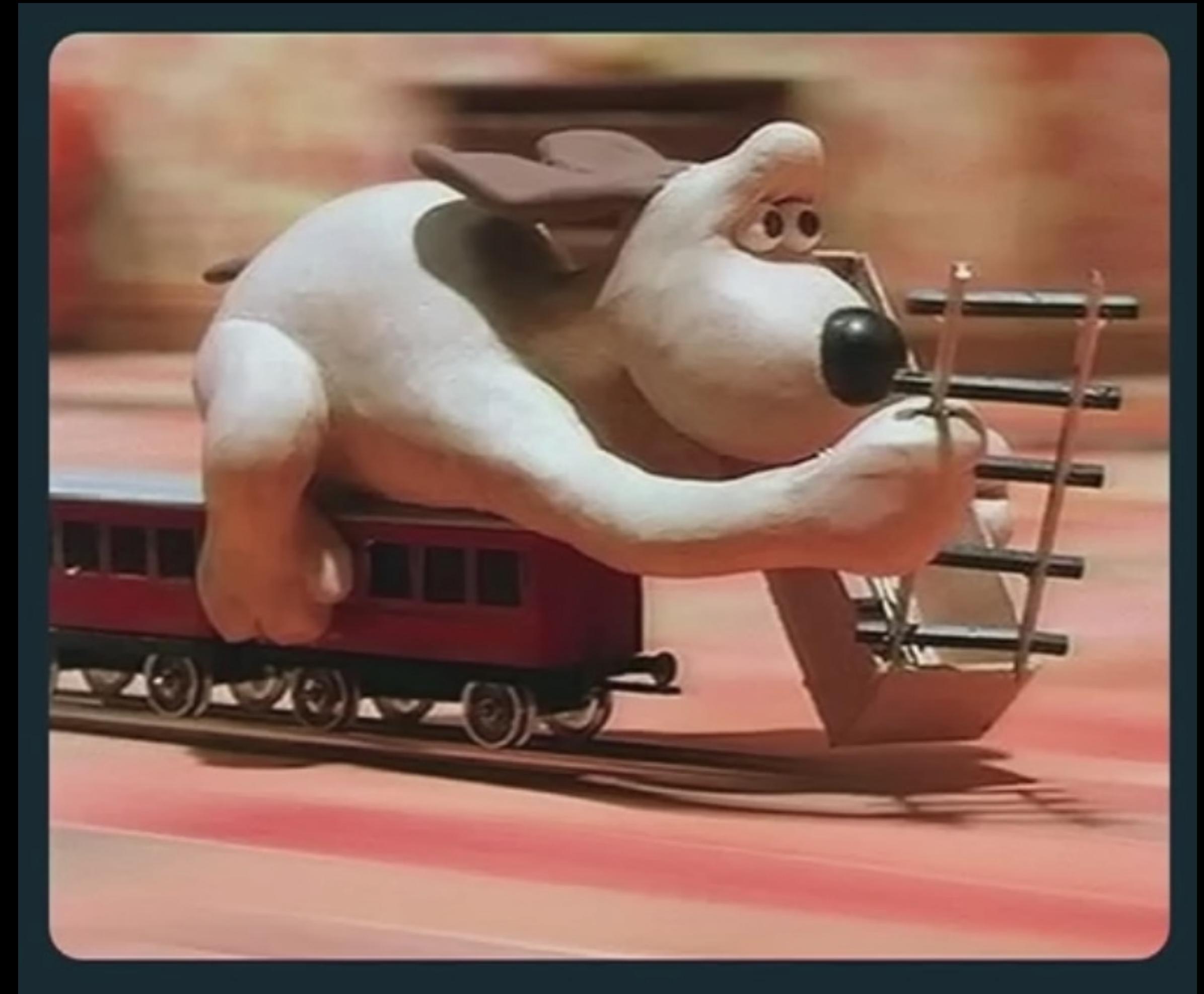


## How we create - Steps

- Research
- Moodboards
- Drafts
- Sketches
- Lofi concepts
- Hifi concepts
- Etc...

We used to search for content that was already created

Now we “create” the content that we’re searching.



## How we create - Searching

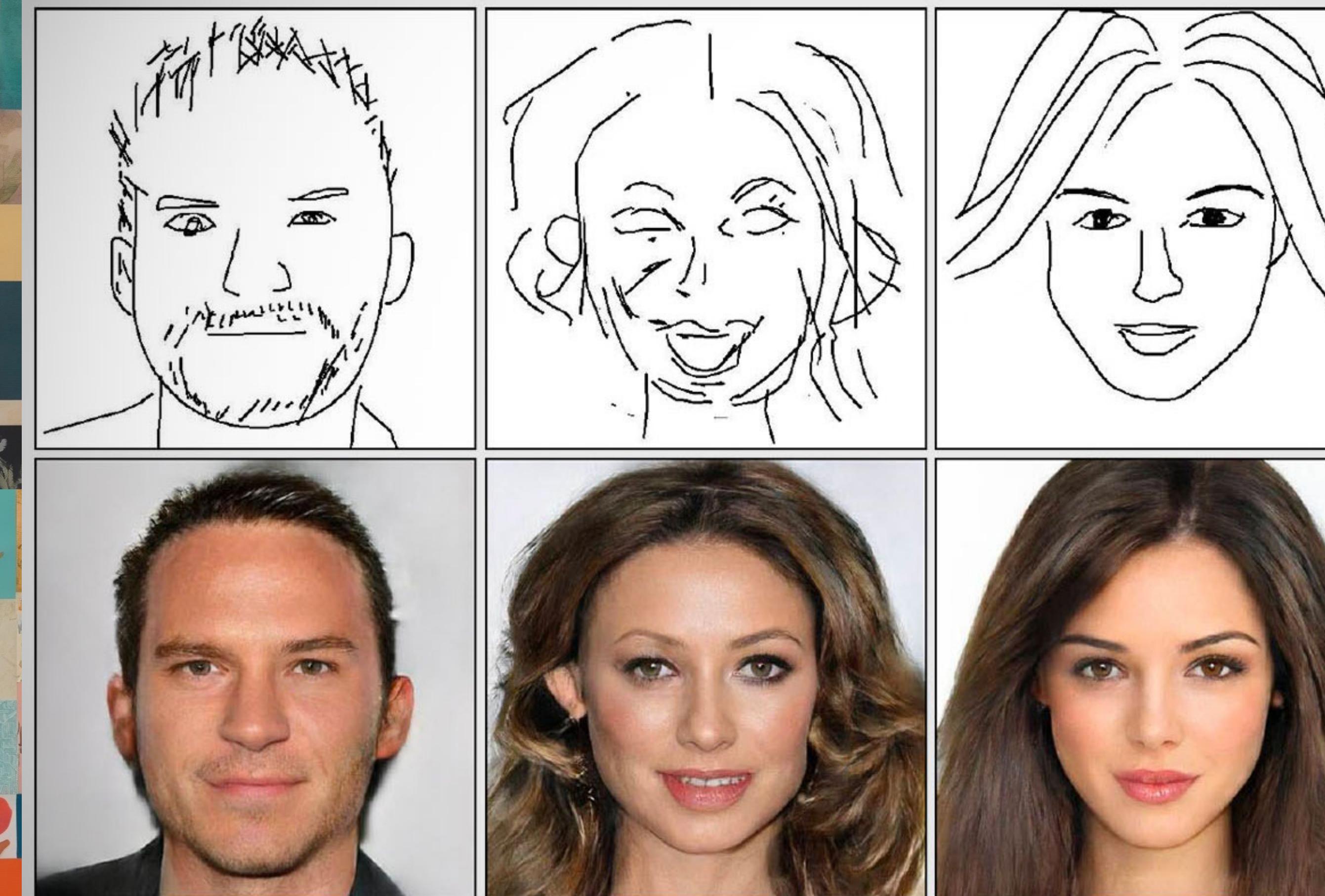


- Research
- Moodboards

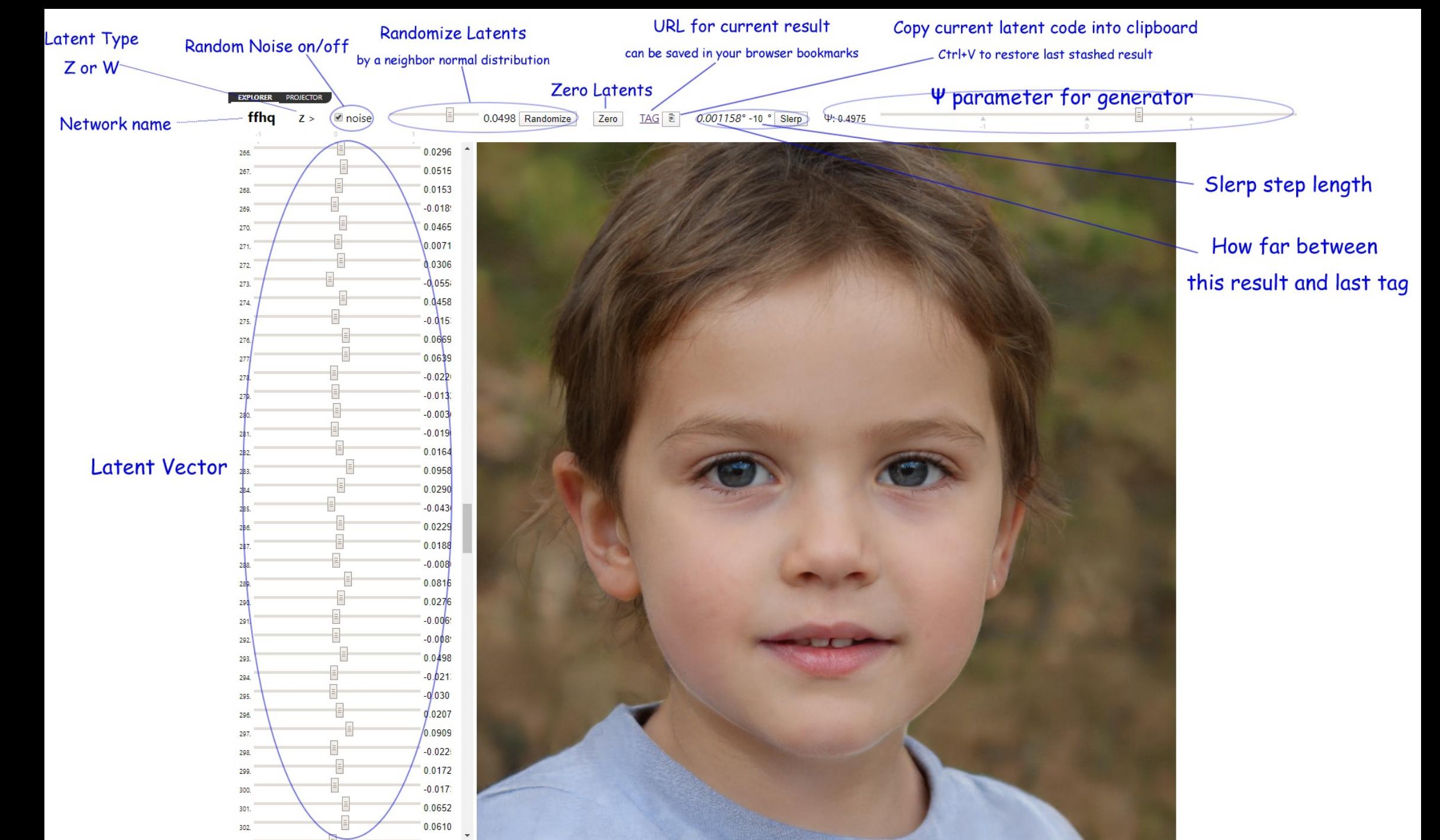
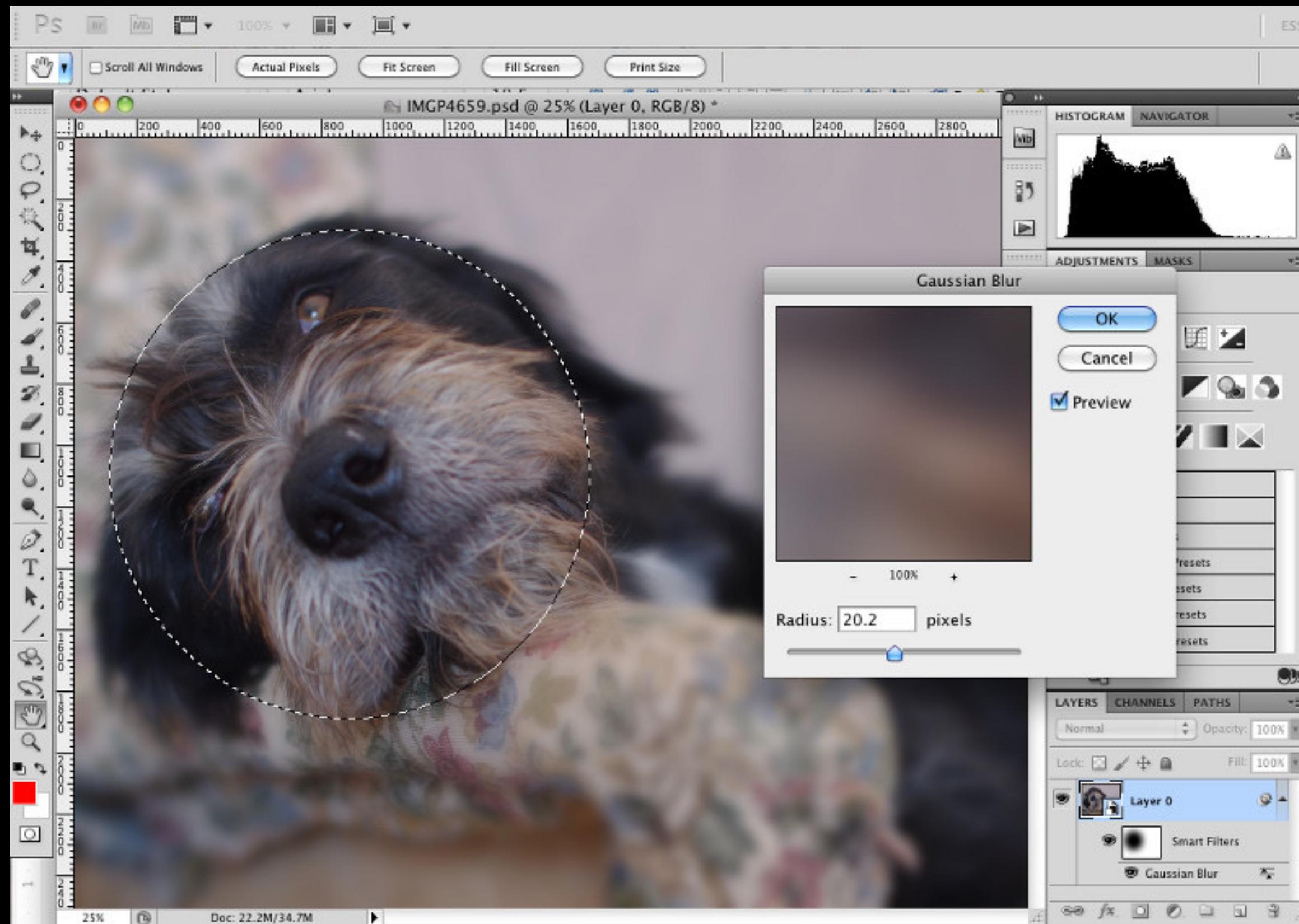


## How we create - Merging

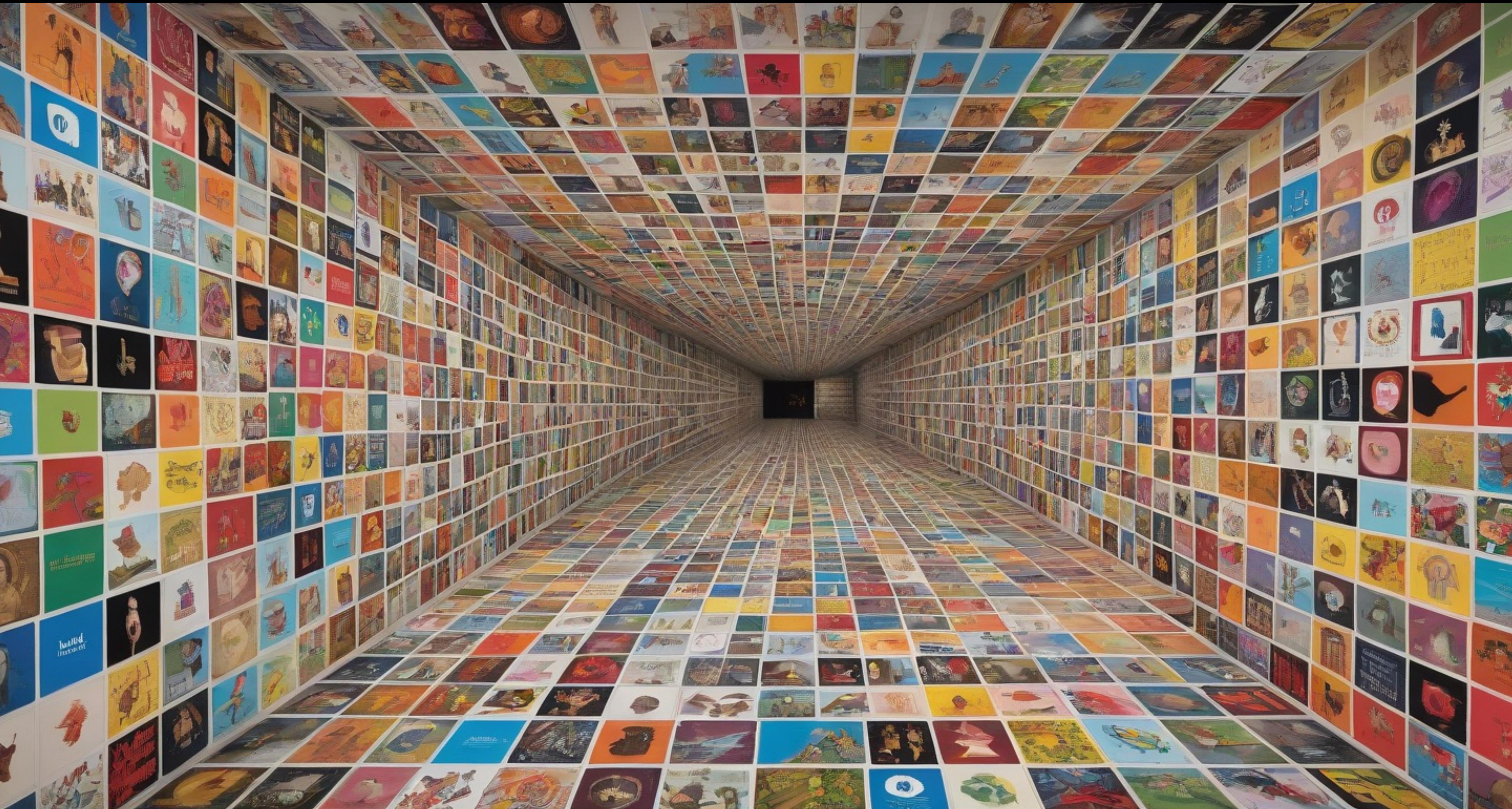
- Drafts
- Sketches
- Lofi concepts
- Hifi concepts
- Etc...



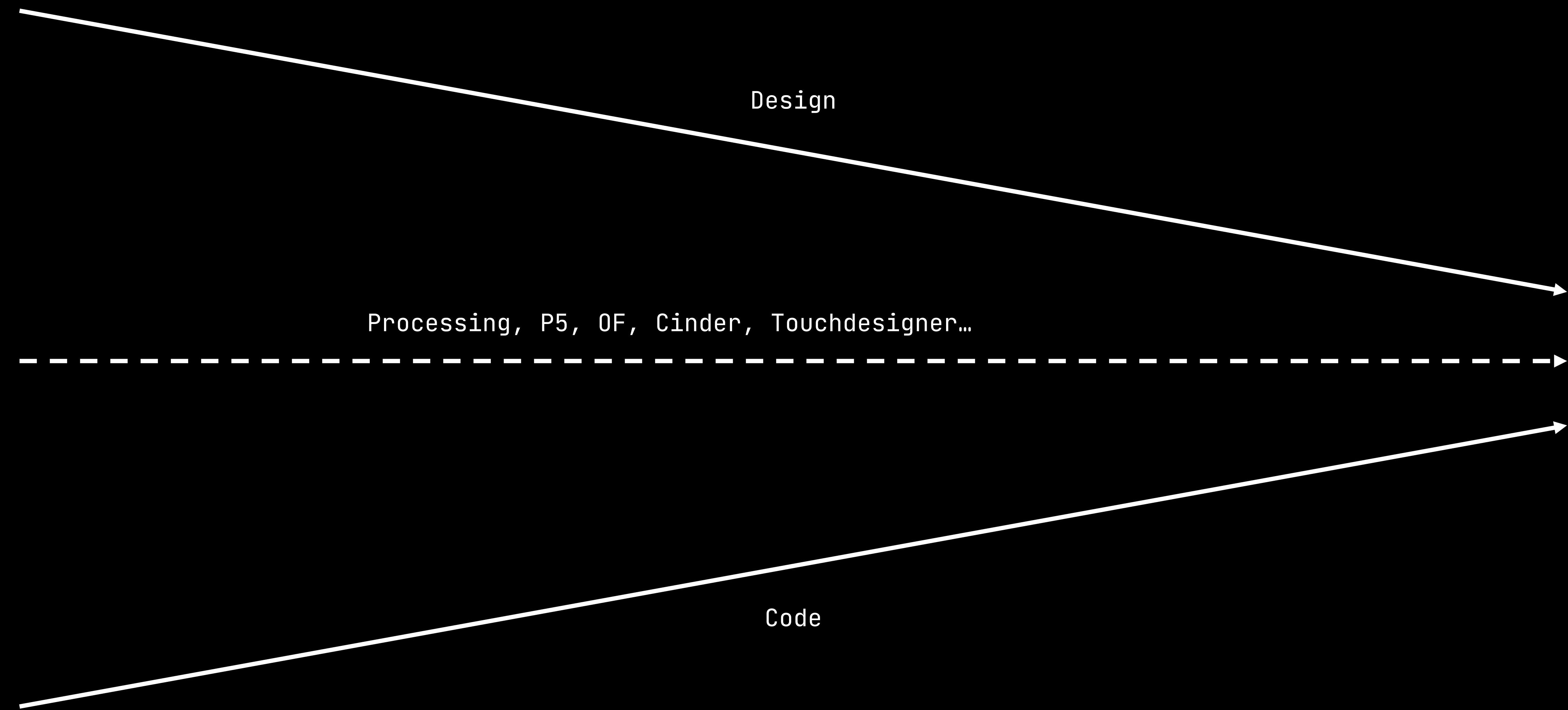
# How we create - Choices - Decision Time 🤔



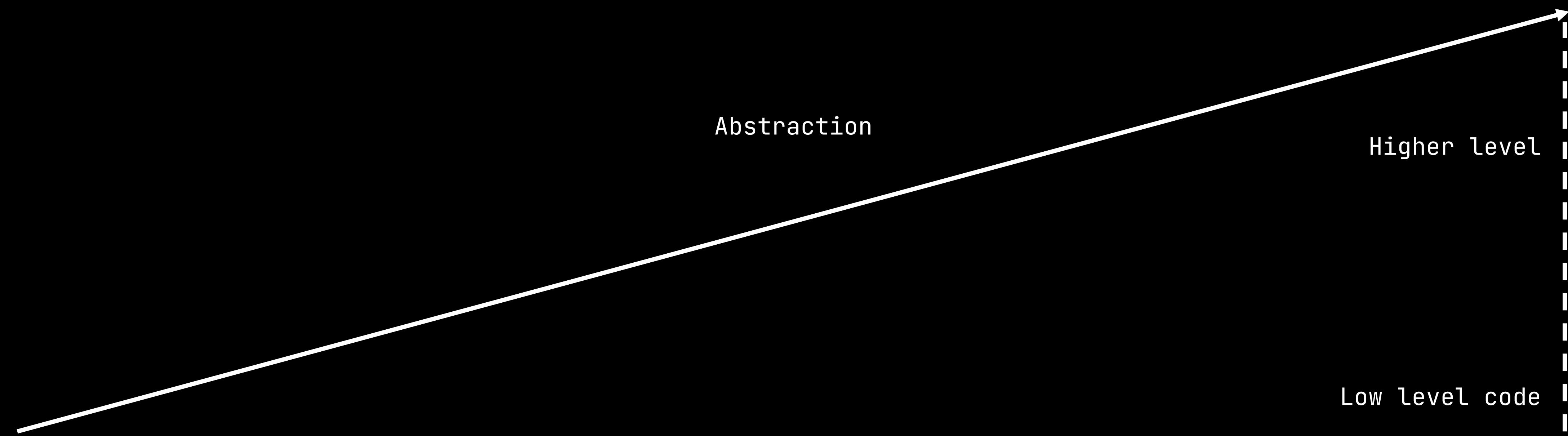
## How we create - Choices - Curation

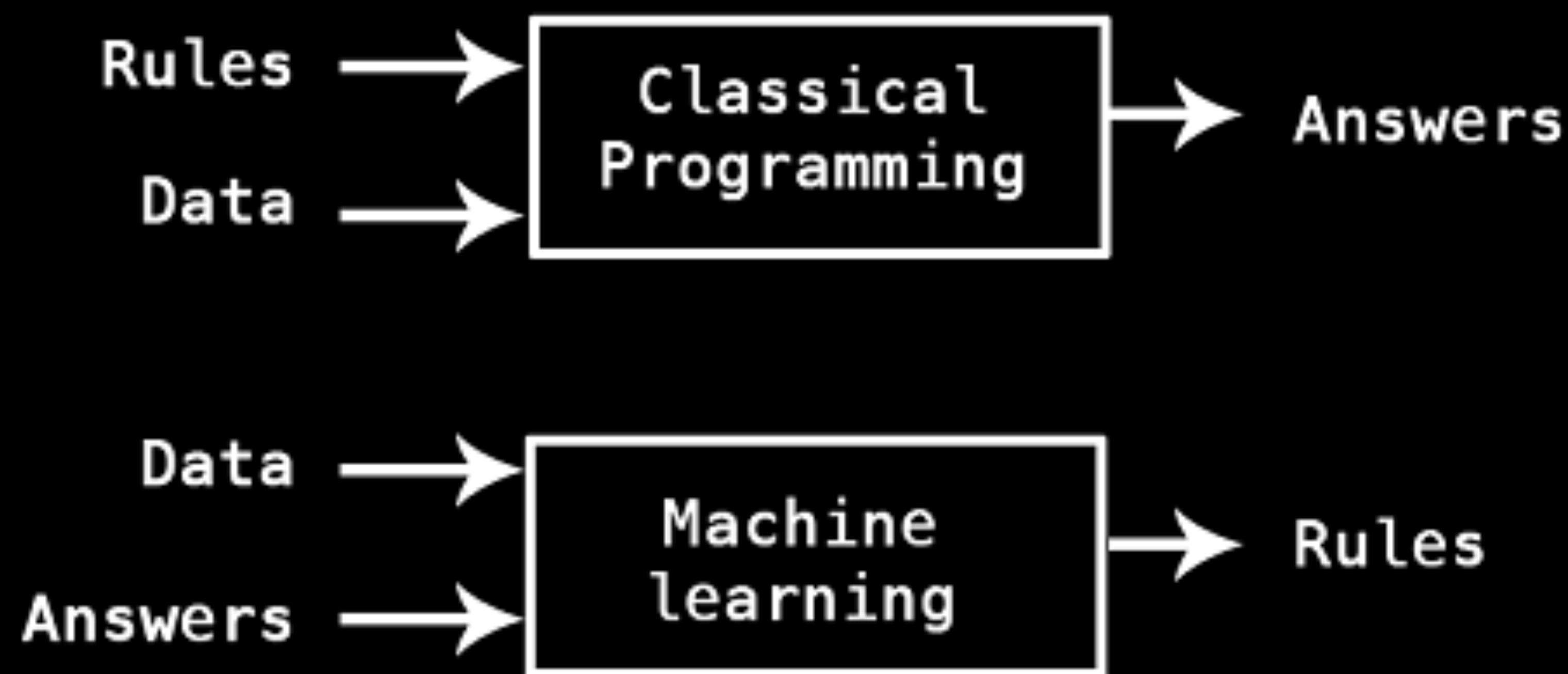


## How we create - Creative tools - Evolution

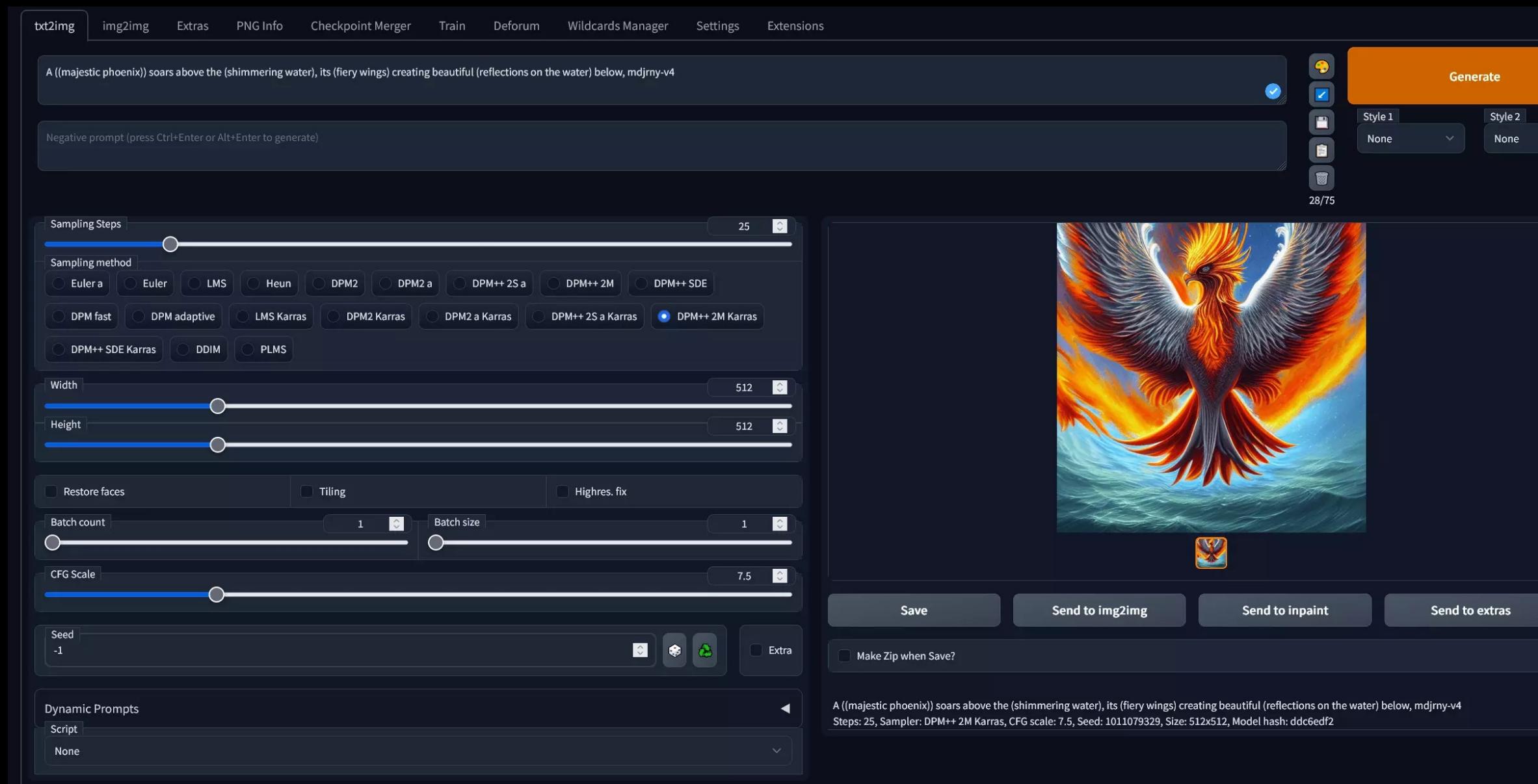


## How we create - Creative tools - Abstraction





# How we create - Creative tools



The screenshot shows a Jupyter Notebook interface. In the top right, it says "Trusted" and "Jupyter Server". The notebook has two open files: "plot\_venturelli2020\_fig1.ipynb" and "test.ipynb". The "test.ipynb" file is active. It contains the following code:

```

import os
import glob
#test
import sys
print(sys.version)
import pygmt
import pandas as pd
import numpy as np

```

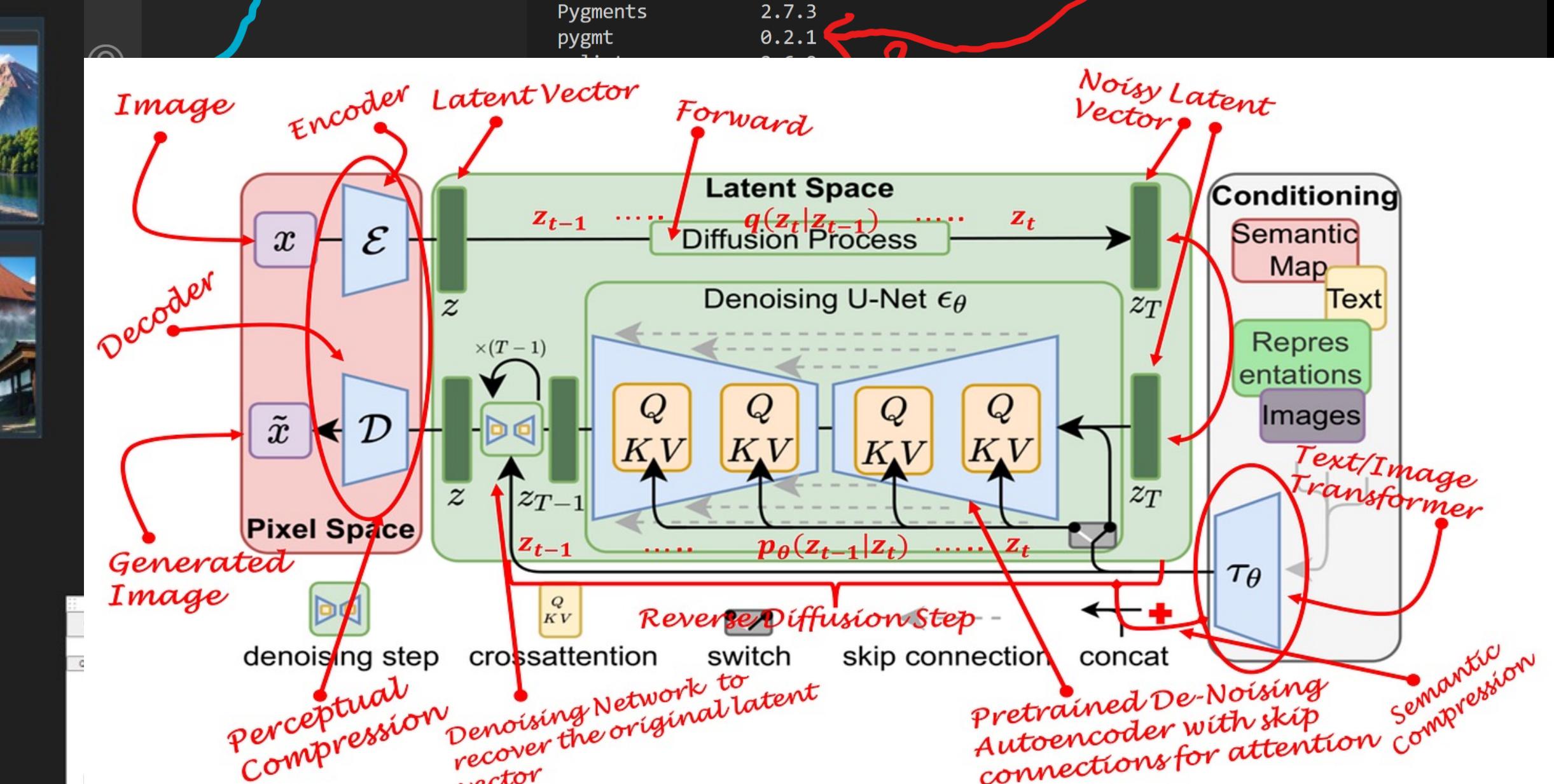
Below the code, the output shows:

```

3.8.5 (default, Sep 3 2020, 21:29:08) [MSC v.1916 64 bit (AMD64)]
ModuleNotFoundError: No module named 'pygmt'

```

A red circle highlights the error message "ModuleNotFoundError: No module named 'pygmt'". Red arrows point from this error to the "PROBLEMS" tab in the bottom right, which lists "Pygments" and "pygmt" with versions "2.7.3" and "0.2.1" respectively.



How we create - To the metal



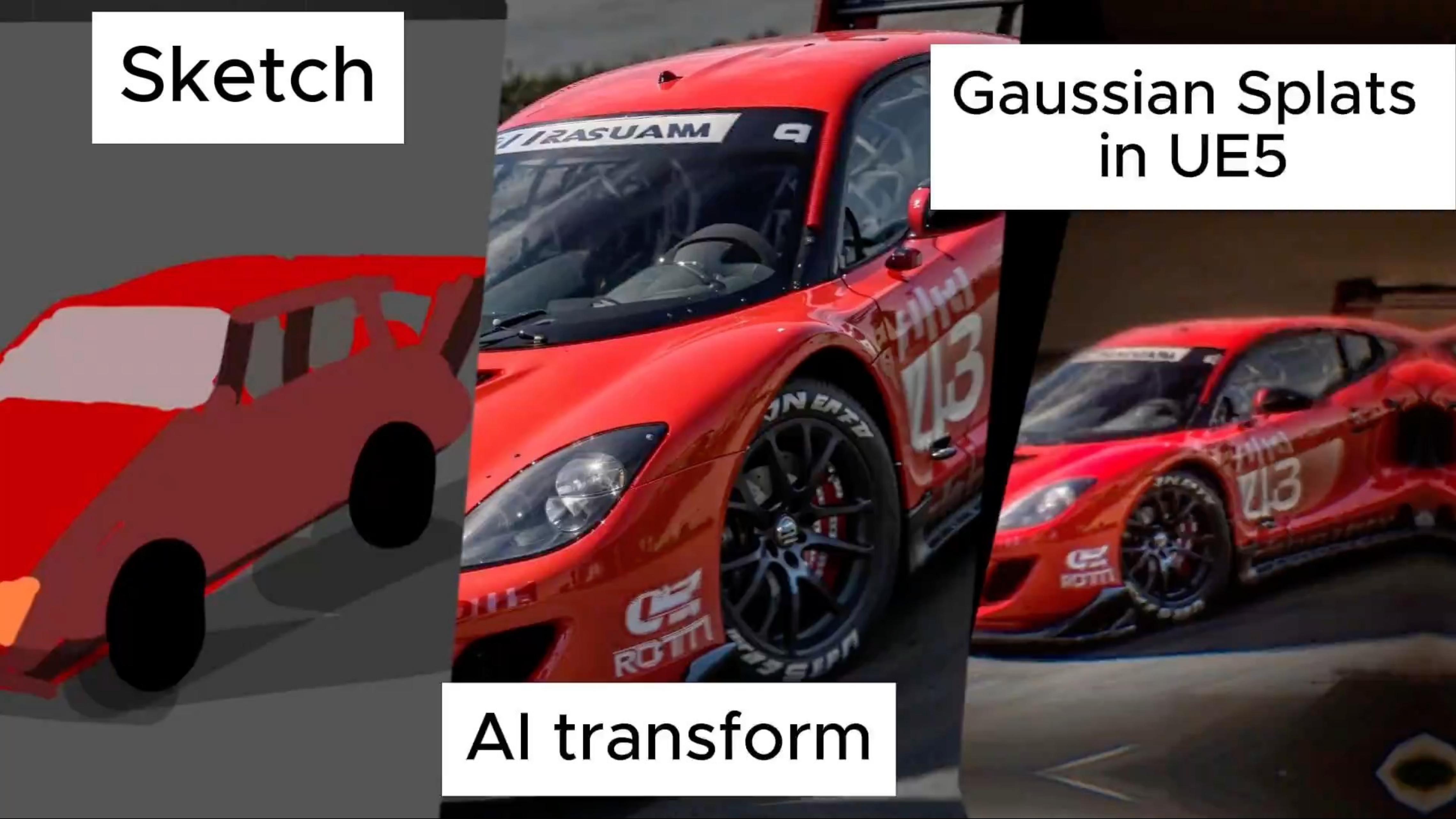
How we create

We are not  
Only  
A brain...



## How we create - Workflow





Sketch

Gaussian Splats  
in UE5

AI transform

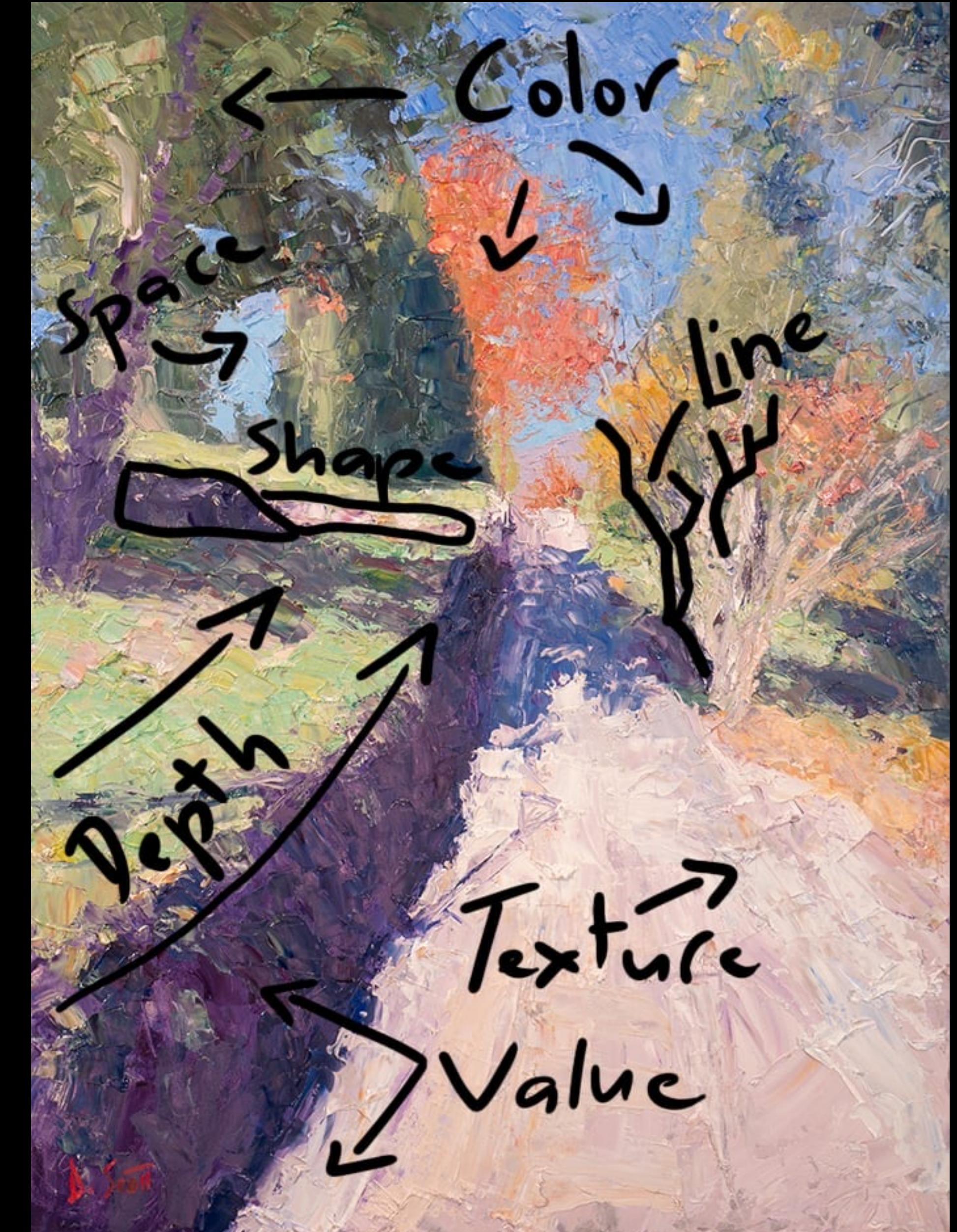
What do I want to say?

How am I going to say it?

## How we create - Rules - Composition



- Line
- Shape
- Color
- Texture
- Value
- Space
- Depth



→ “Learn the rules like a pro  
So you can break them like  
an artist!”

*Pablo Picasso*



→ The bigger picture

The bigger picture

- Democratised
- Techtopia



The bigger picture

Automation

