

ImageGen

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

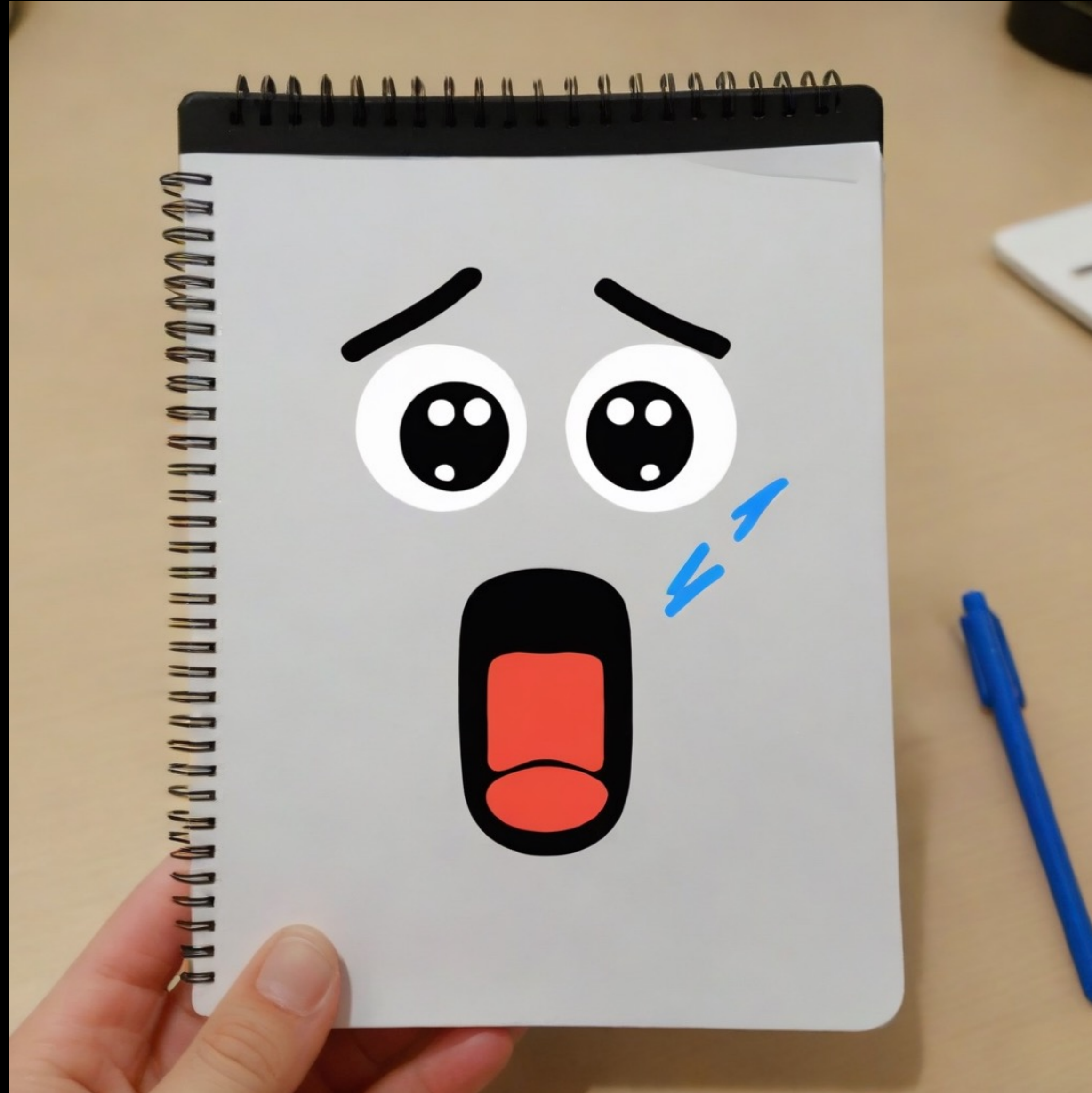
→ Intro

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

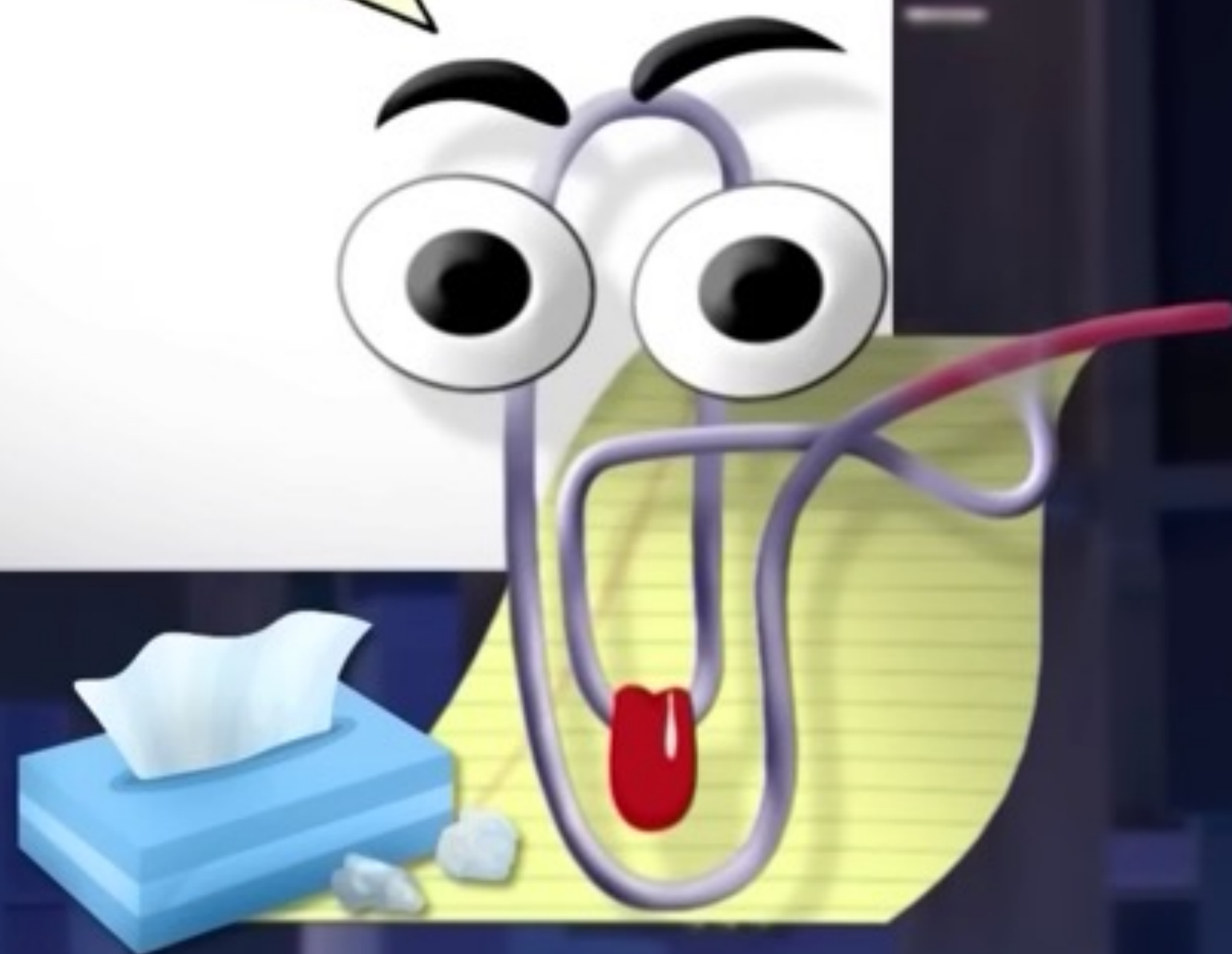
Introduction - The plan



→ “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.







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



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



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







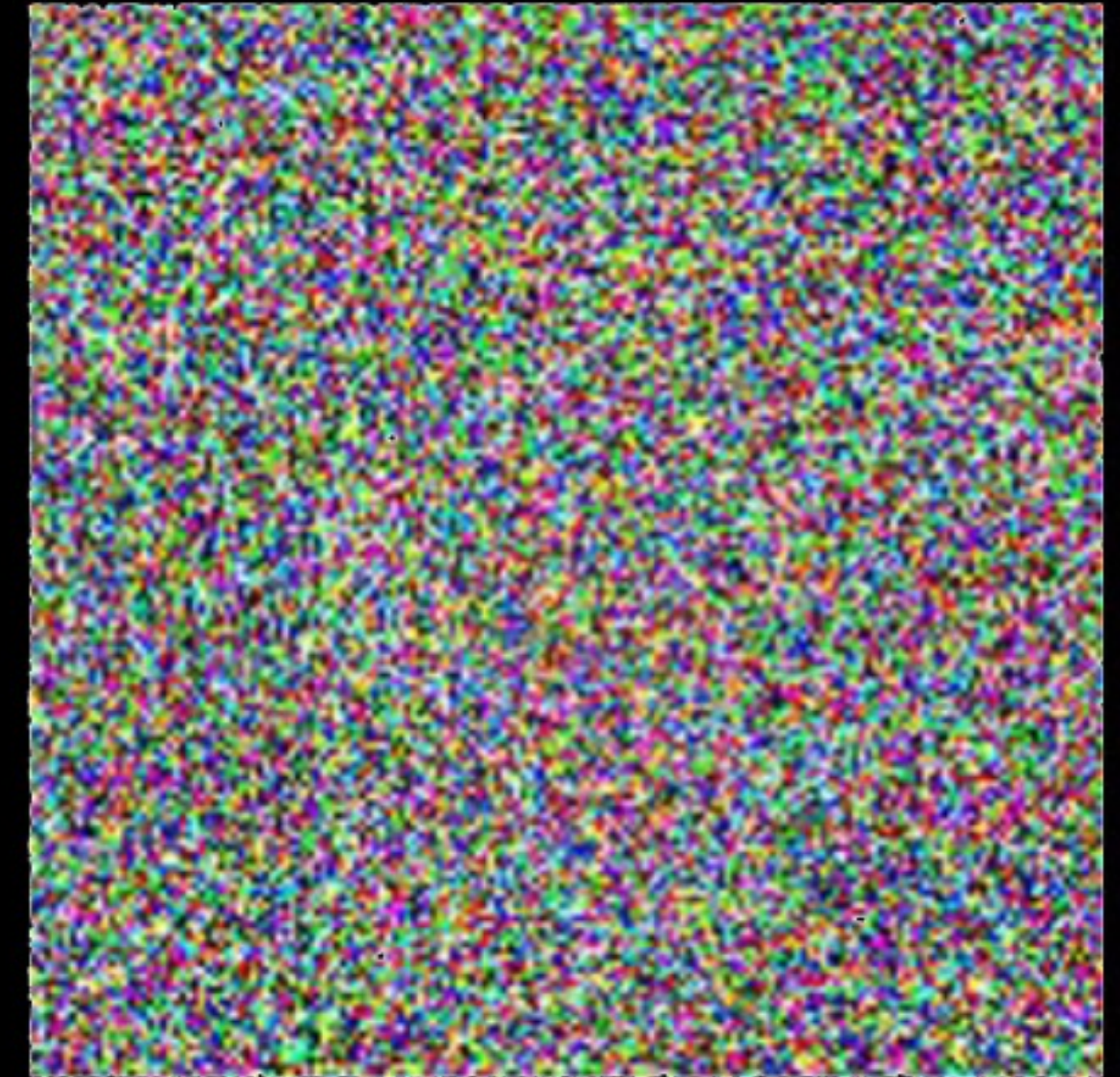
→ Status Quo





Diffusion

Status Quo - Concepts - Diffusion



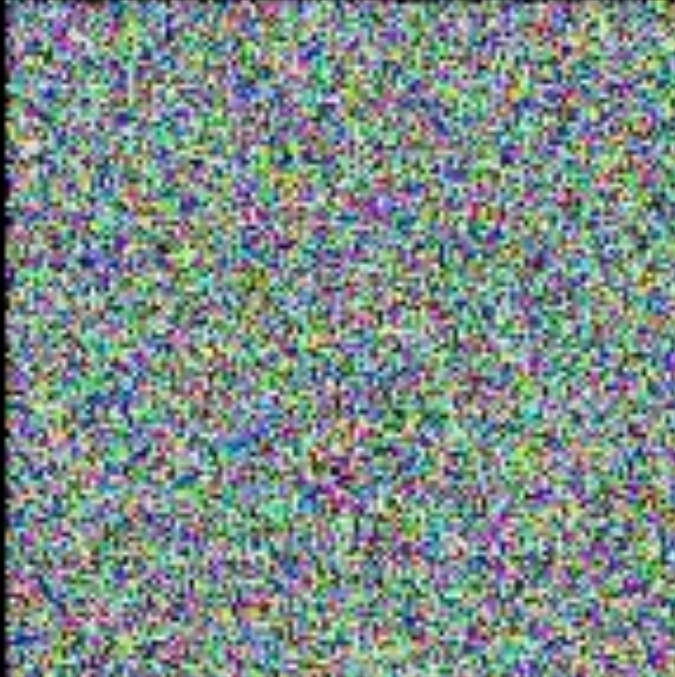
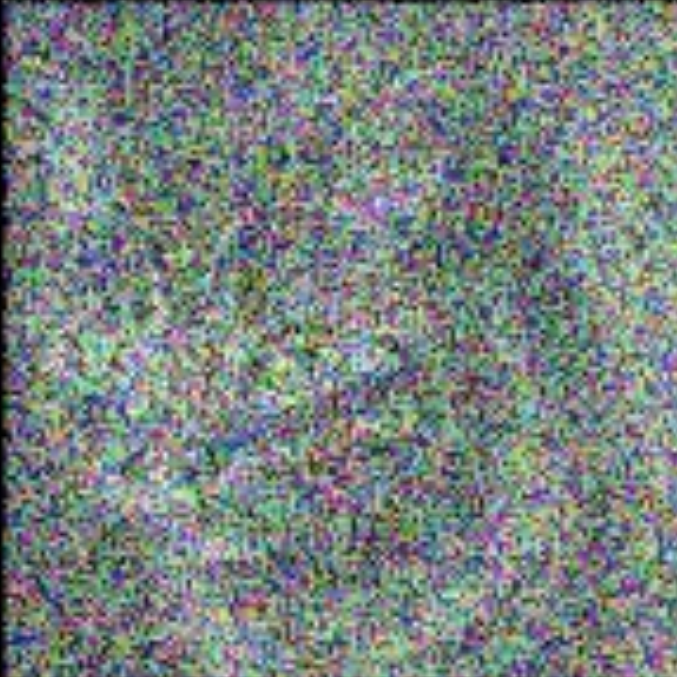
Step

1

2

3

4

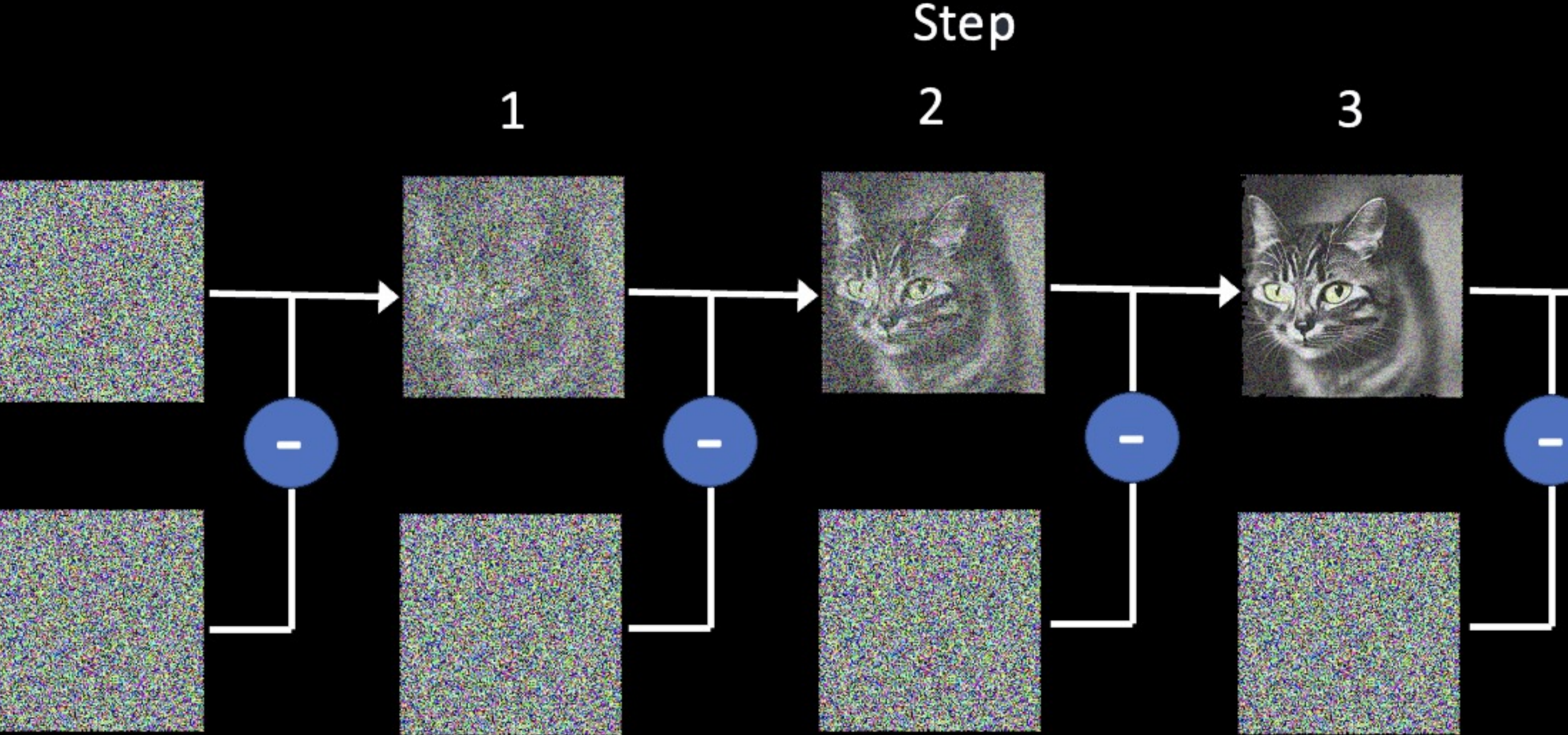


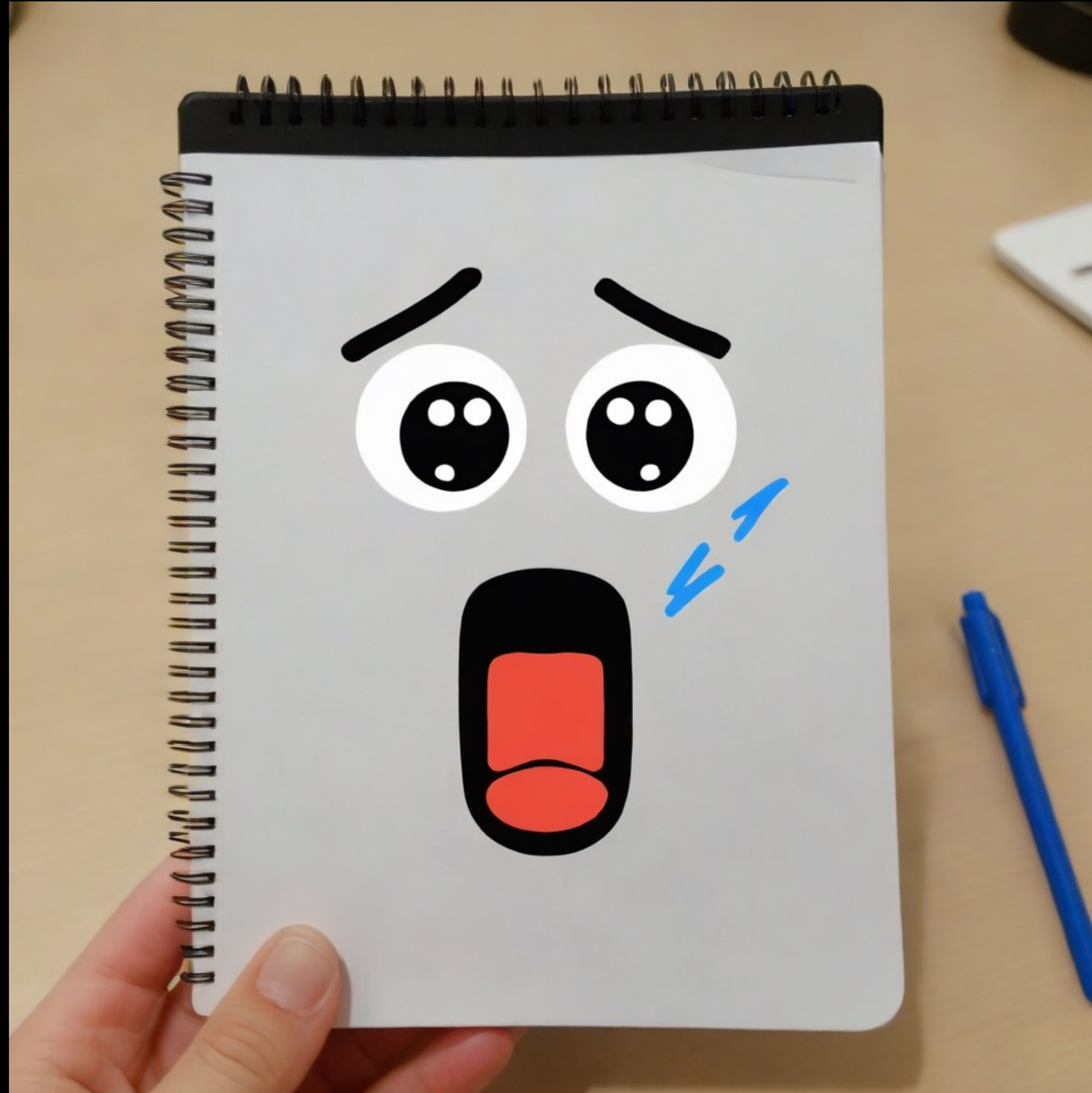
Noise 1

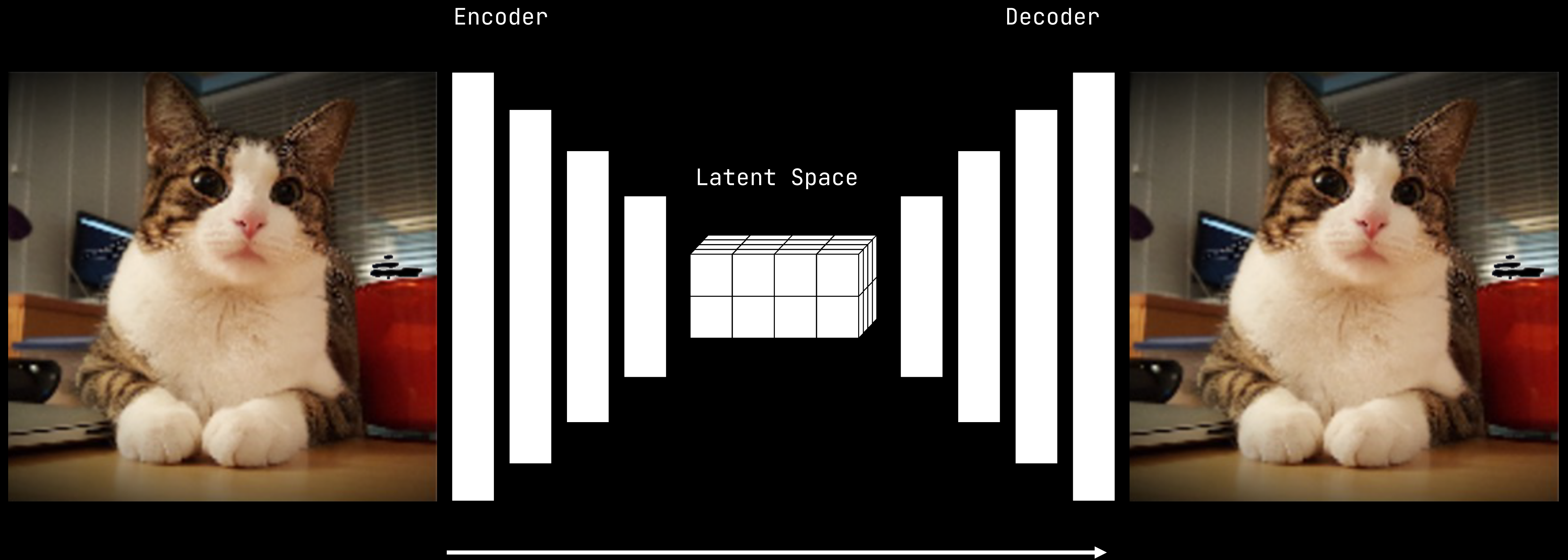
Noise 2

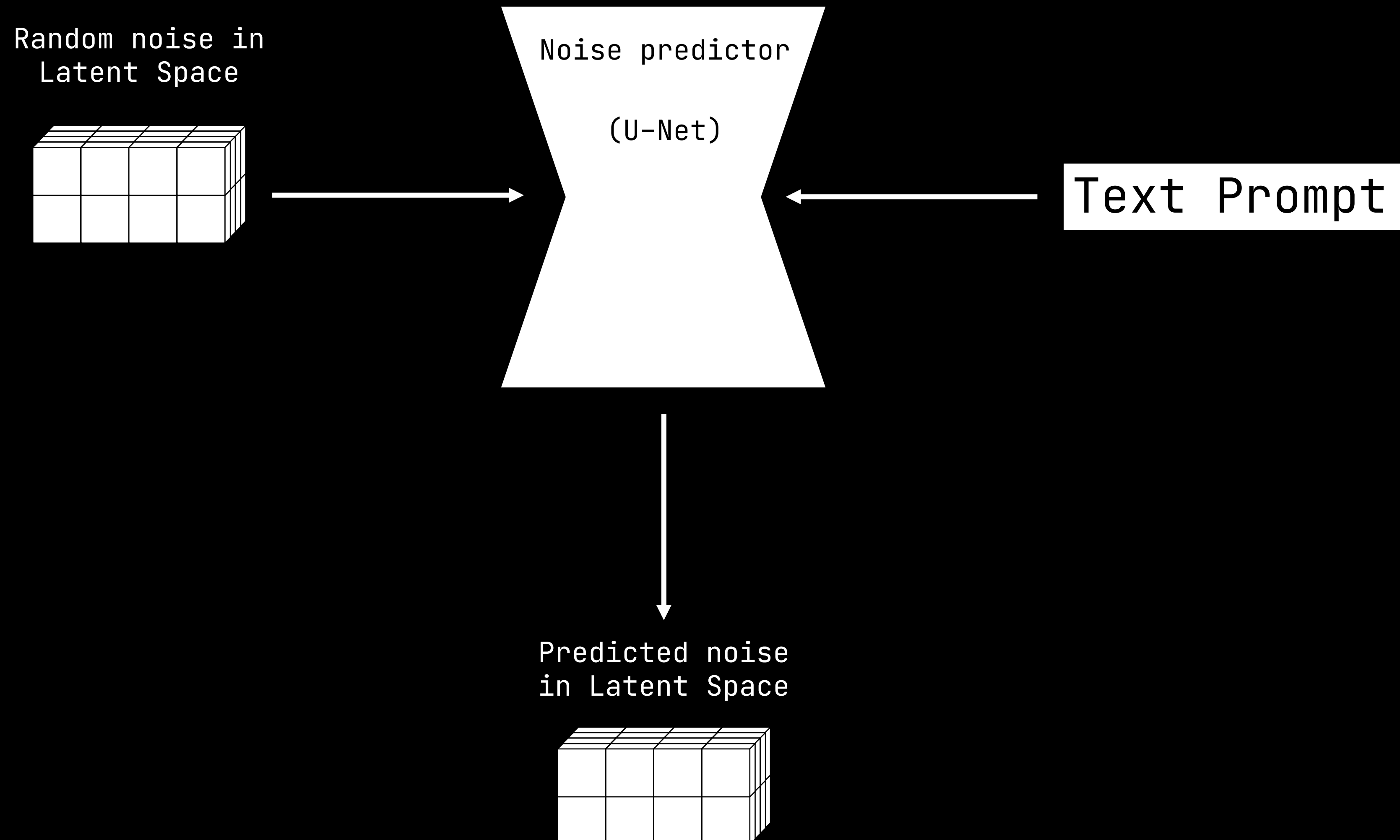
Noise 3

Noise 4



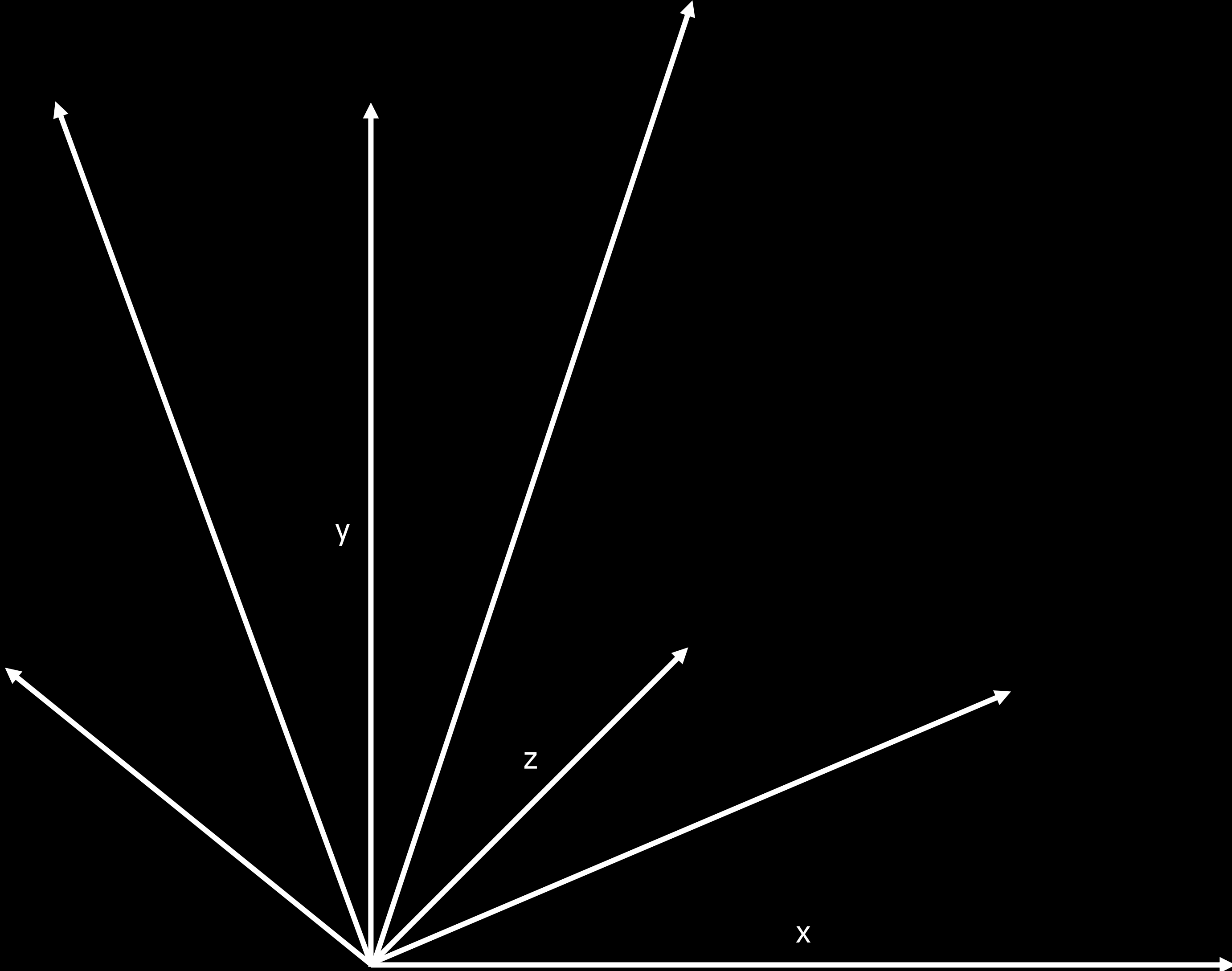








Status Quo - Techniques - input - Latent Space (dimensionality)





→ “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***



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.

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**.

SUN397

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



✓ a photo of a **television studio**.

✗ a photo of a **podium indoor**.

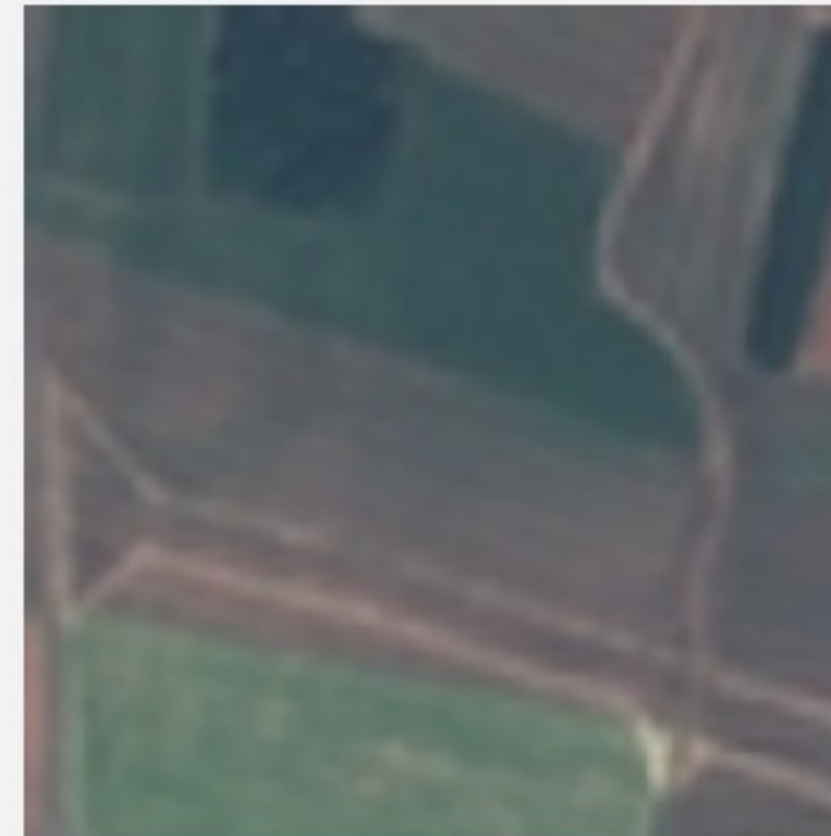
✗ a photo of a **conference room**.

✗ a photo of a **lecture room**.

✗ a photo of a **control room**.

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



- Access
- Tools
- Inspiration
- Education
- Skills?

→ “Creativity is intelligence having fun.”

Albert Einstein

→ How we create

NOBODY
STARTS
OUT
ORIGINAL



- 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.

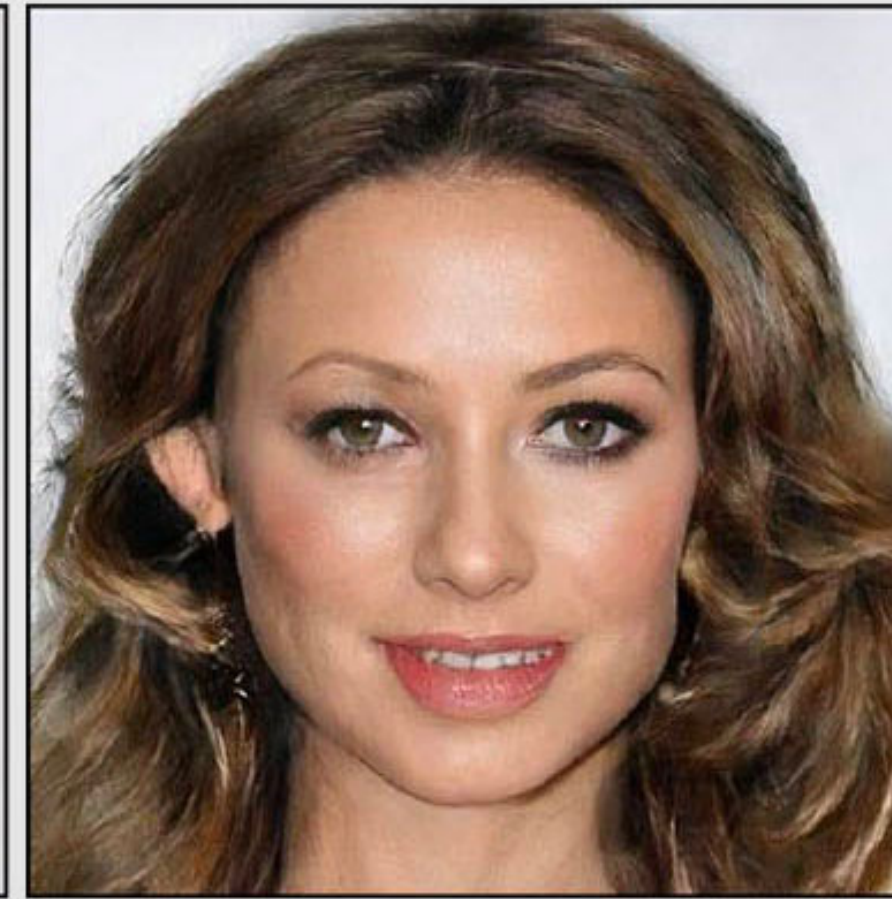
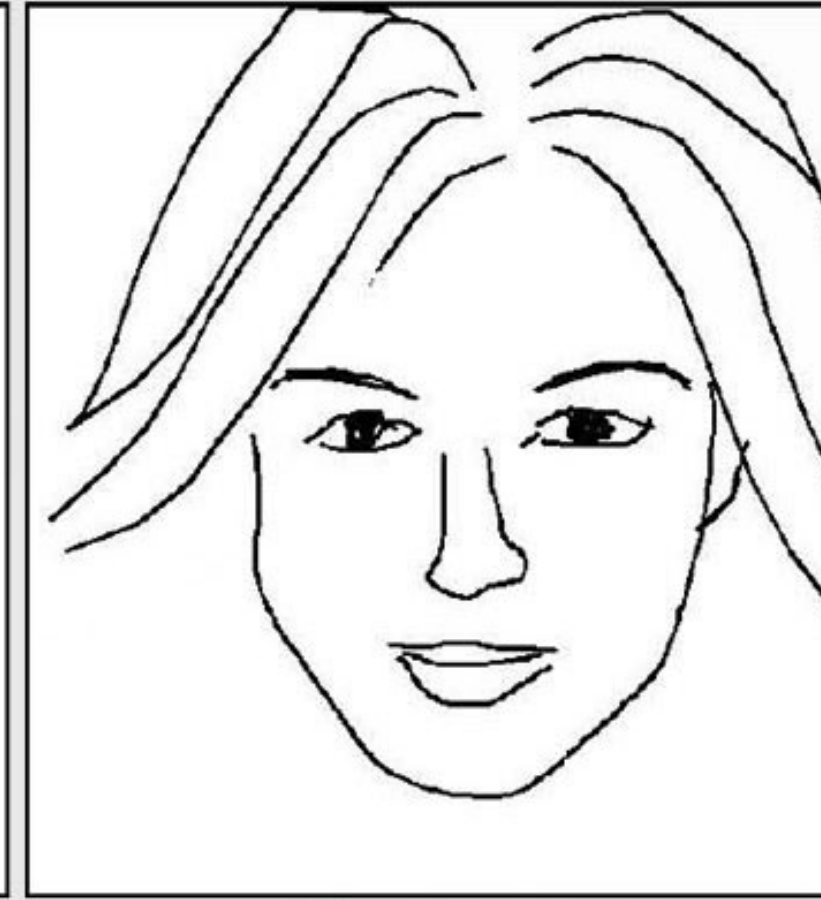
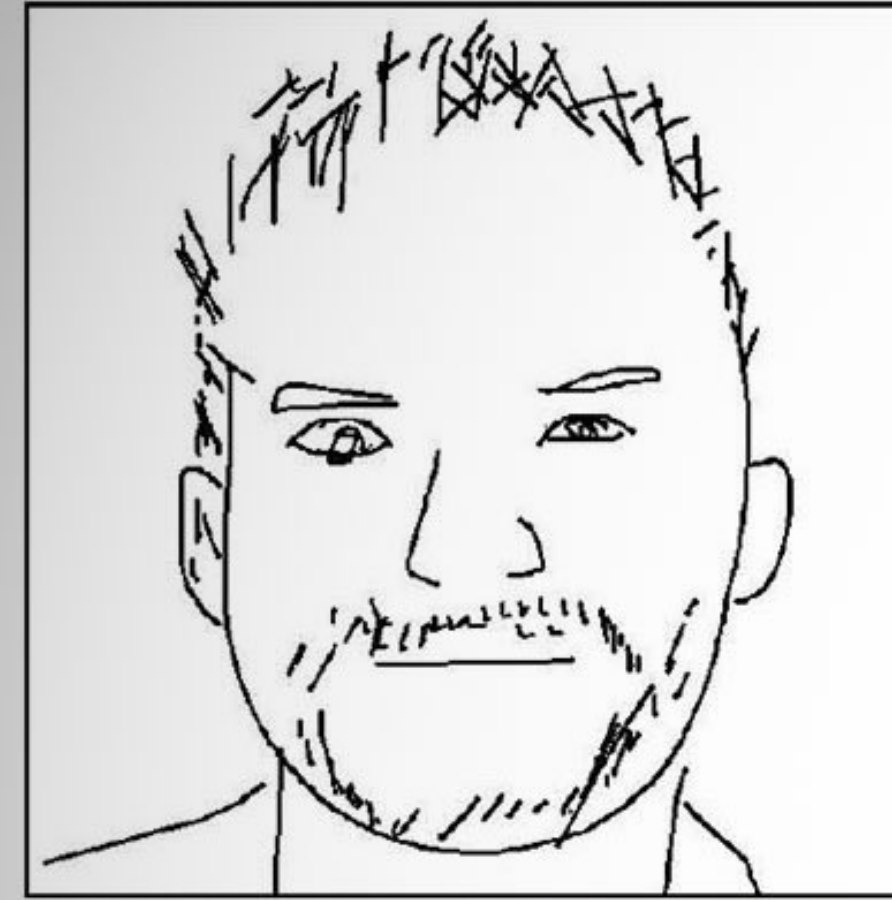




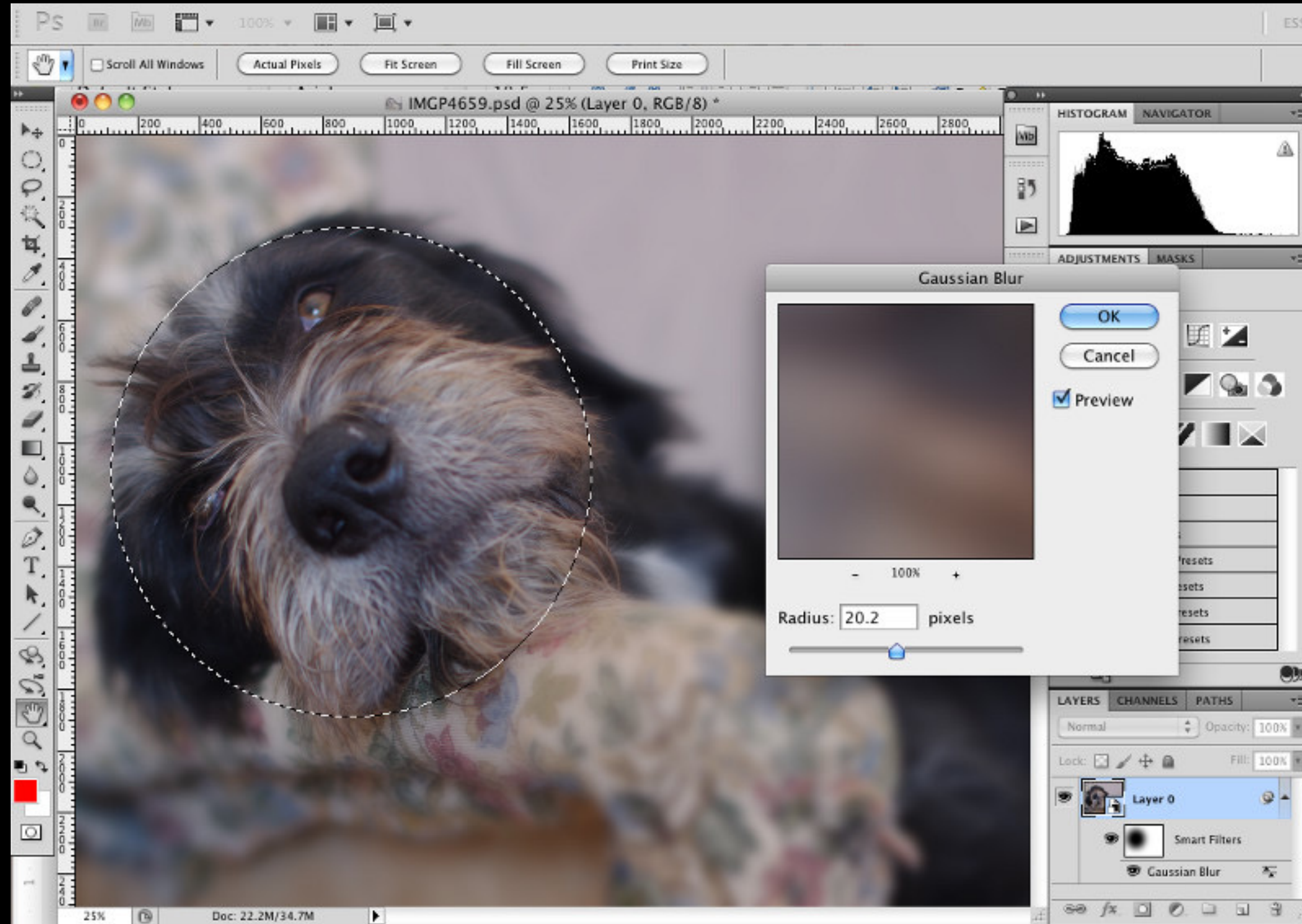
- Research
- Moodboards



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



How we create - Choices - Decision Time 🤔



Latent Type
Z or W

Network name
ffhq

Random Noise on/off
noise

Randomize Latents
by a neighbor normal distribution

Zero Latents
0.0498 Randomize Zero TAG

URL for current result
can be saved in your browser bookmarks
0.001158* -10 * Slerp Ψ : 0.4975

Copy current latent code into clipboard
Ctrl+V to restore last stashed result

Ψ parameter for generator

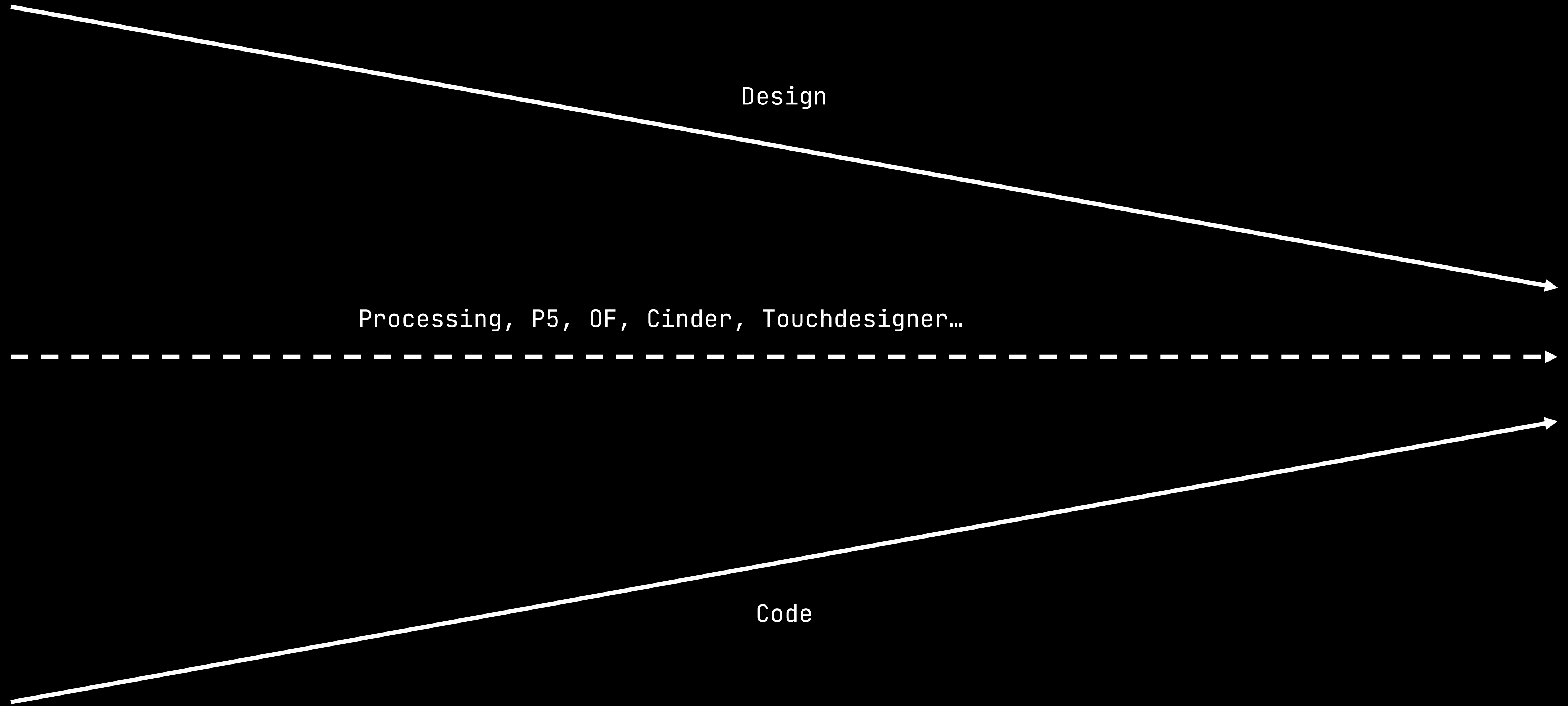
Slerp step length

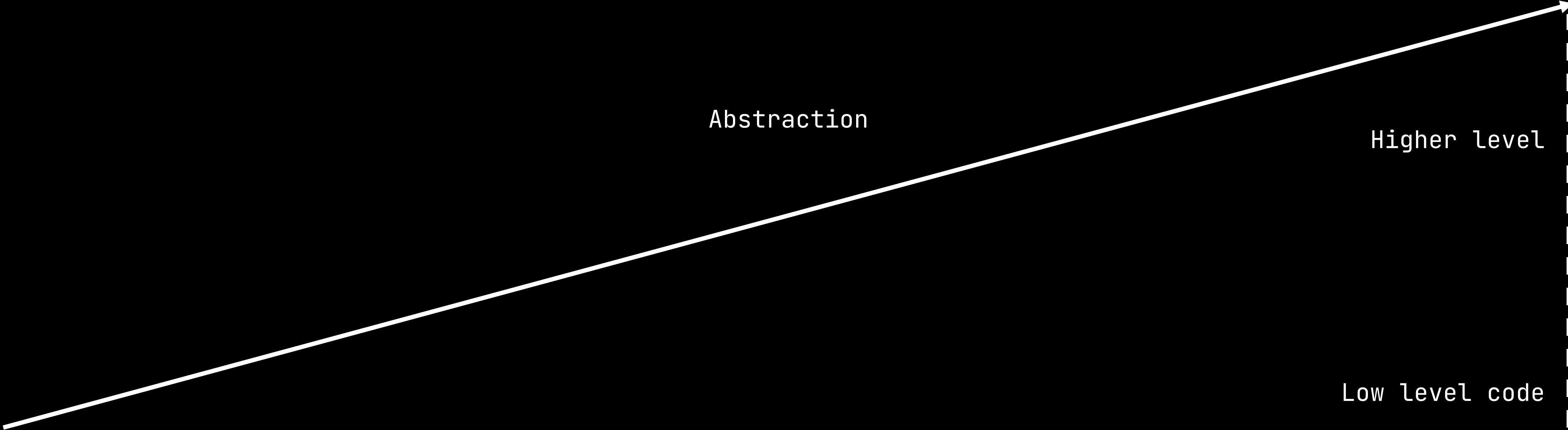
How far between this result and last tag

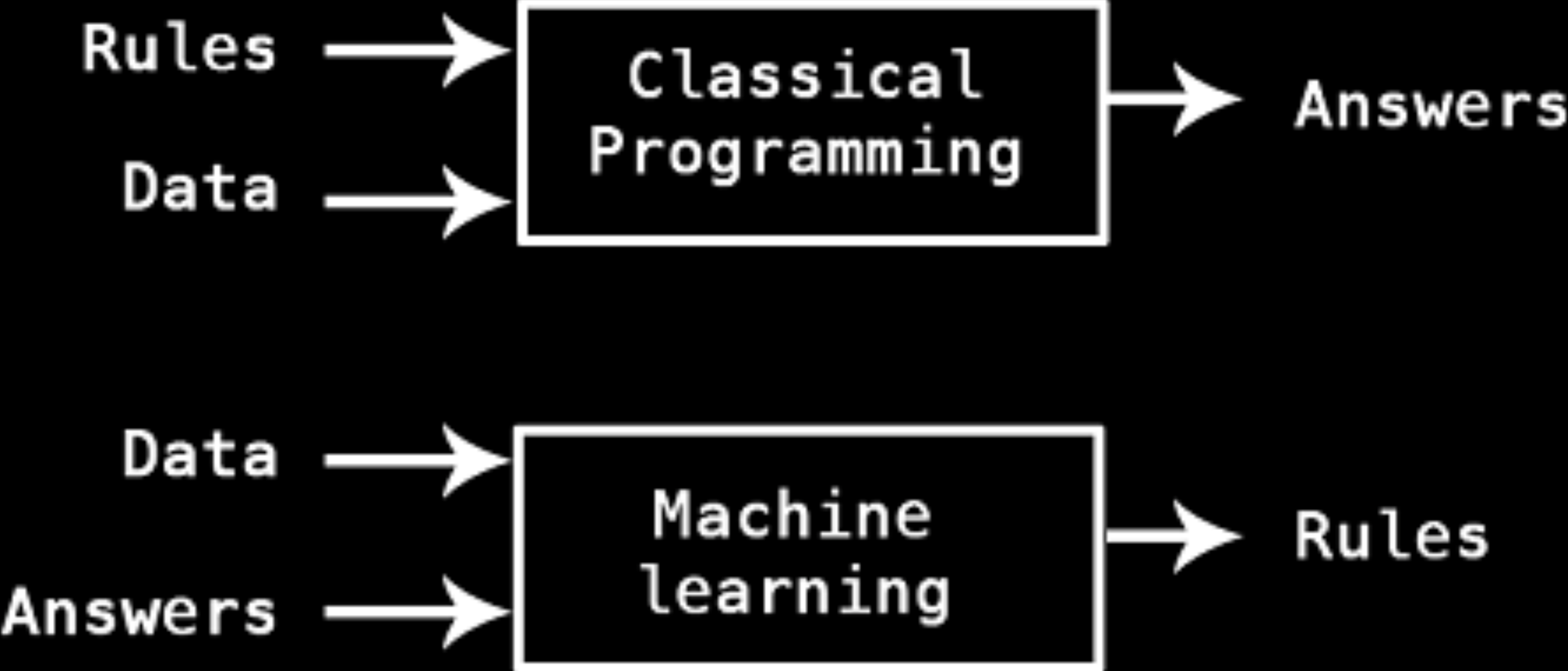
Latent Vector

The image shows a web interface for a generative model. On the left, a 'Latent Vector' is displayed as a vertical list of 32 numerical values. On the right, a generated image of a young child's face is shown. The interface includes various controls and labels as described in the text above.

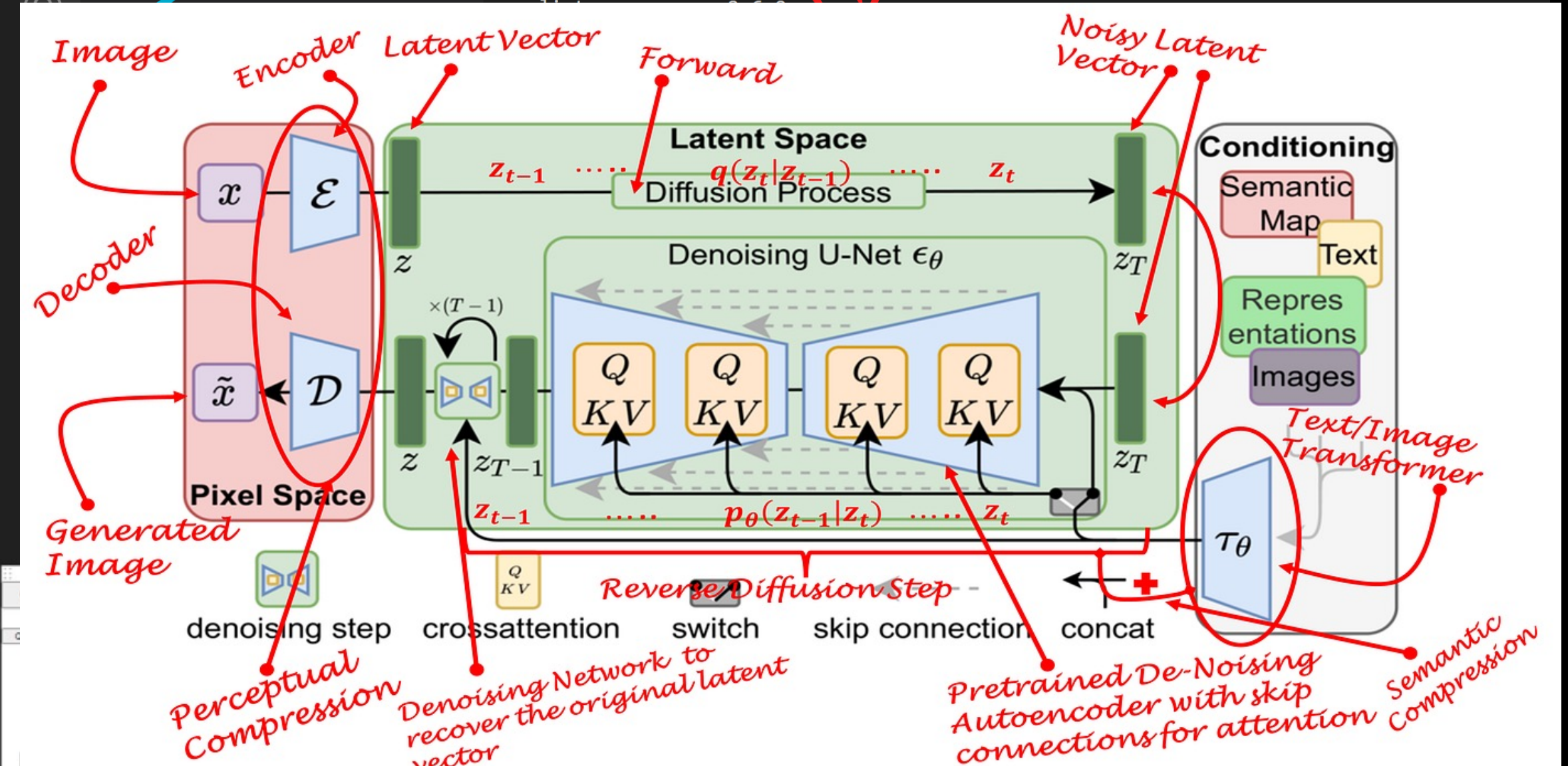
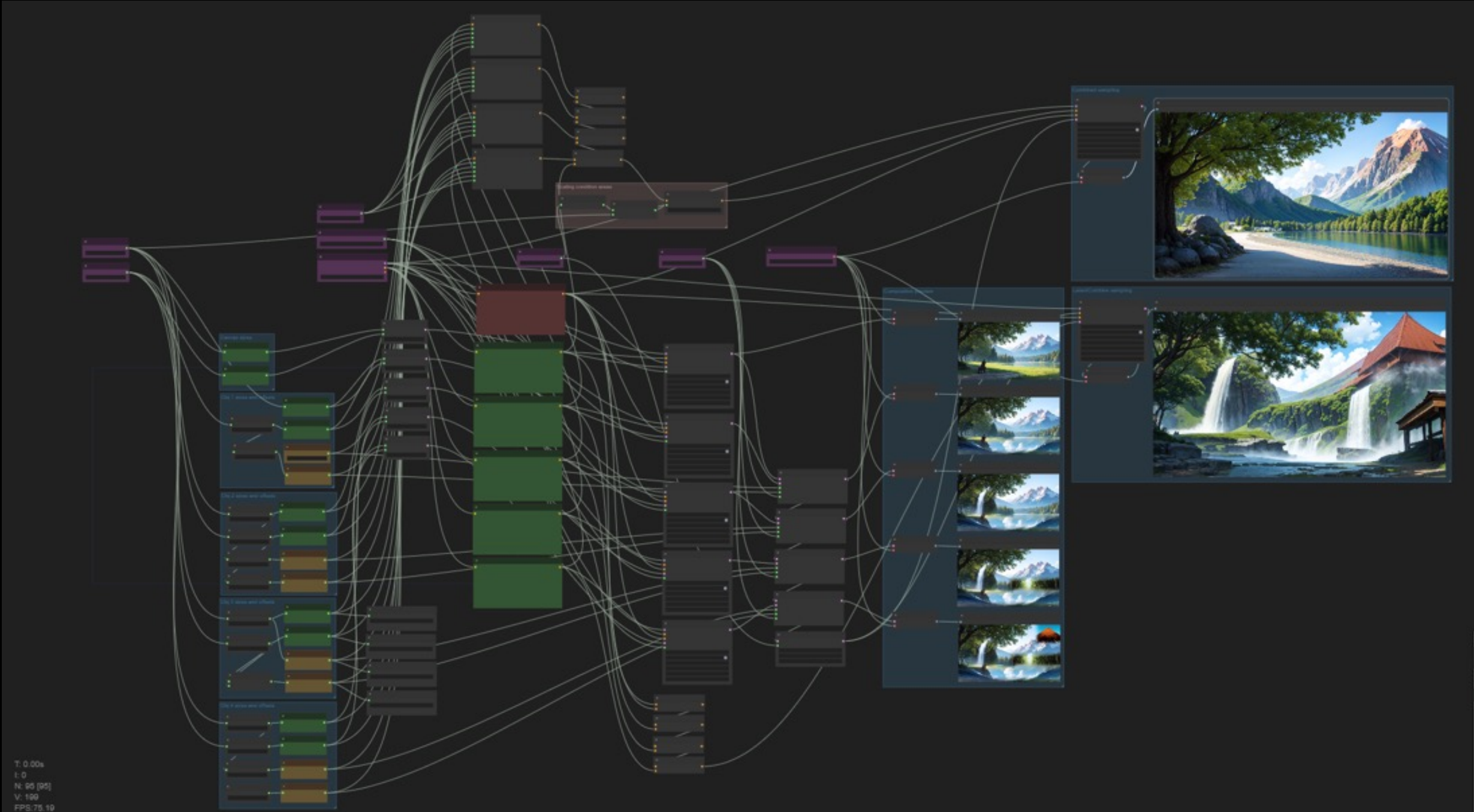
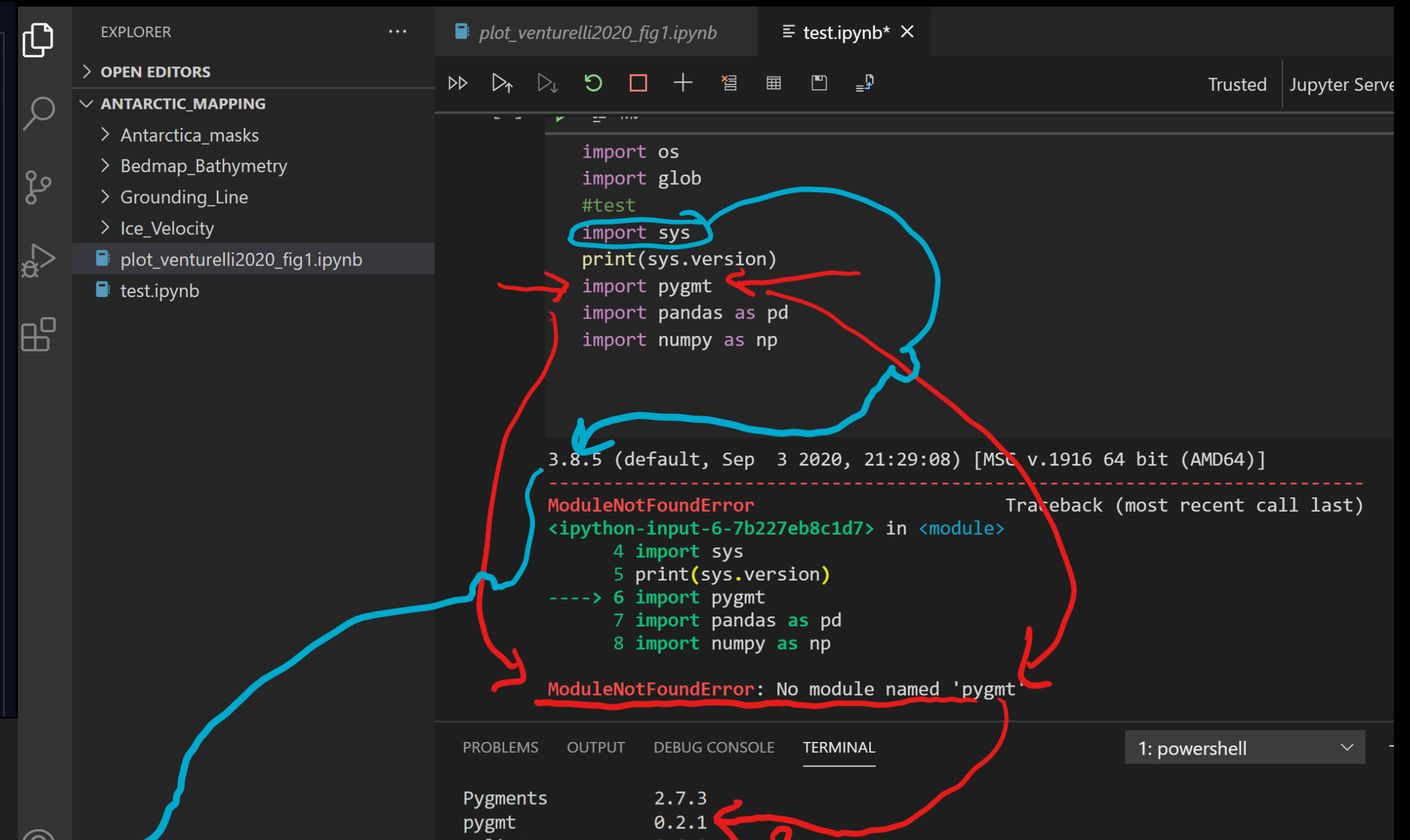
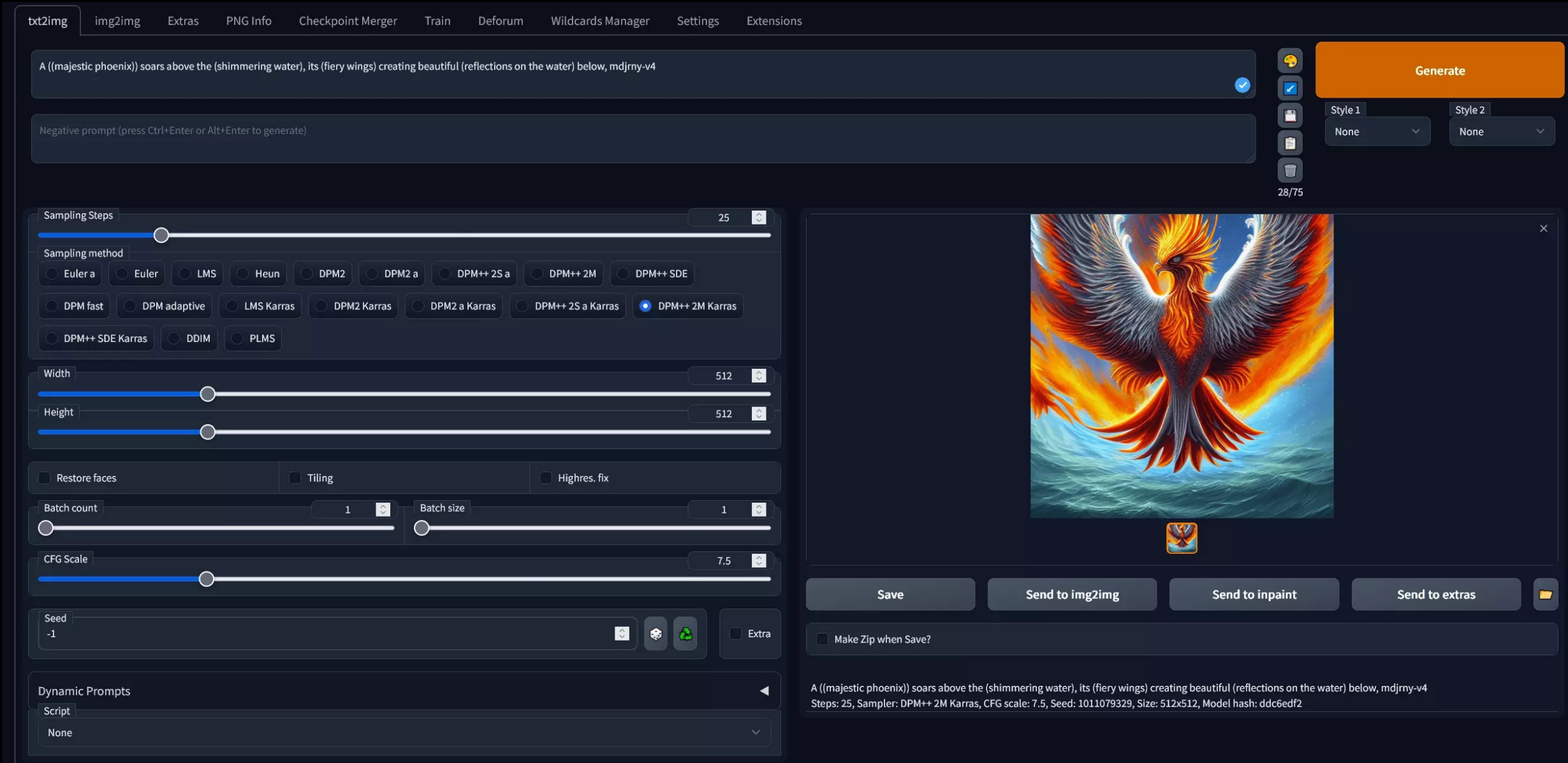








How we create - Creative tools





We are not
Only
A brain...

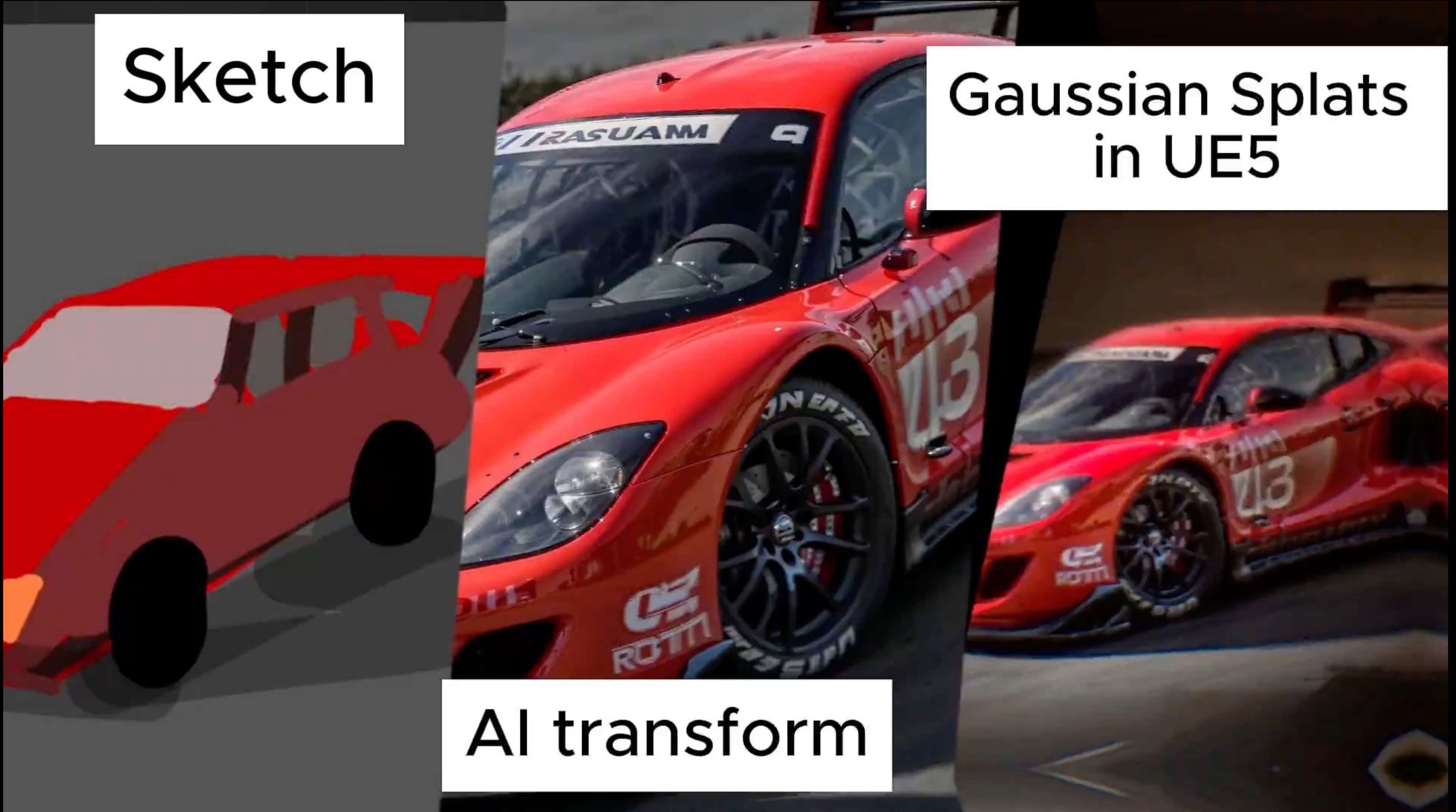




Sketch

Gaussian Splats
in UE5

AI transform



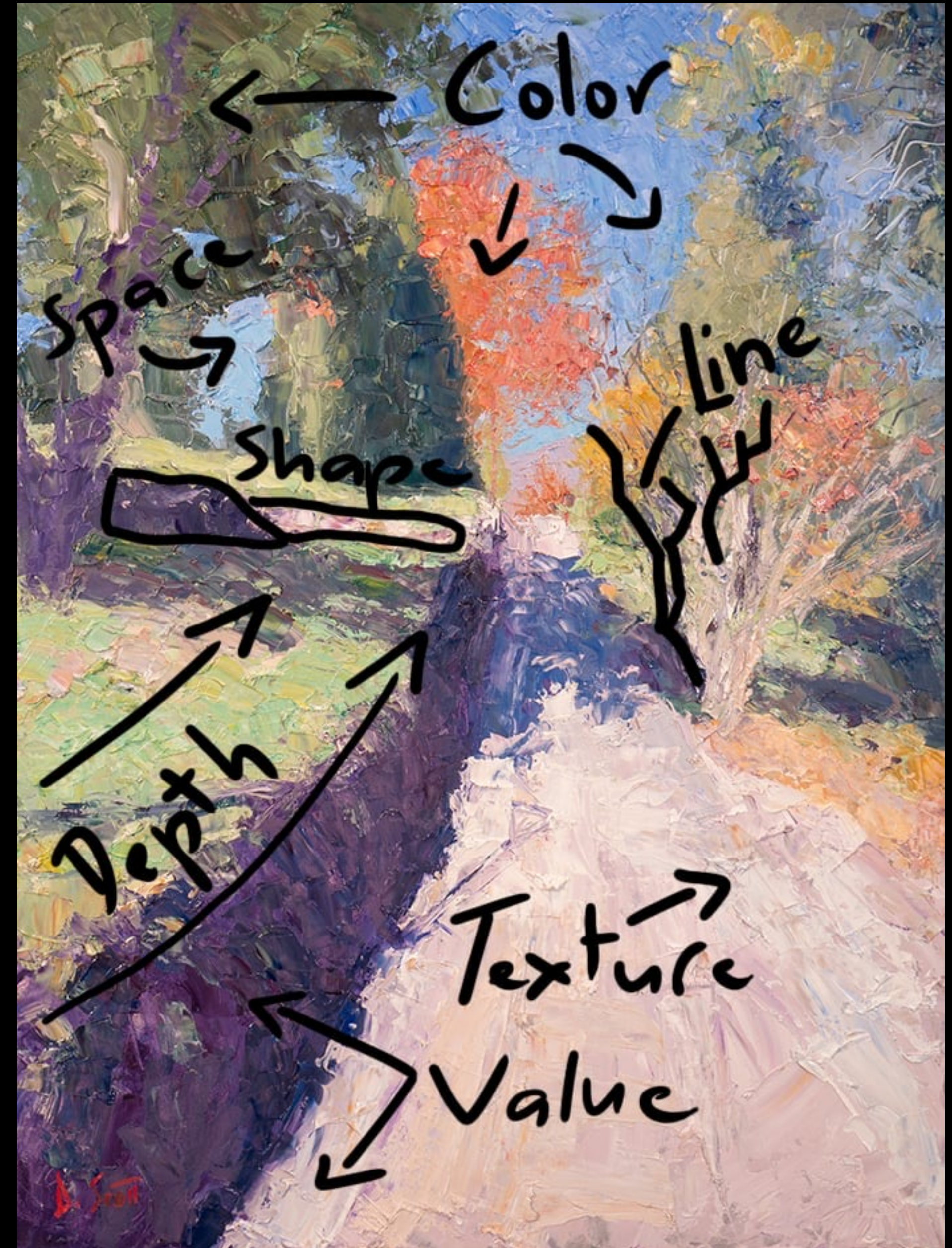
What do I want to say?

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

- Democratised
- Techtopia

The bigger picture

Automation

