



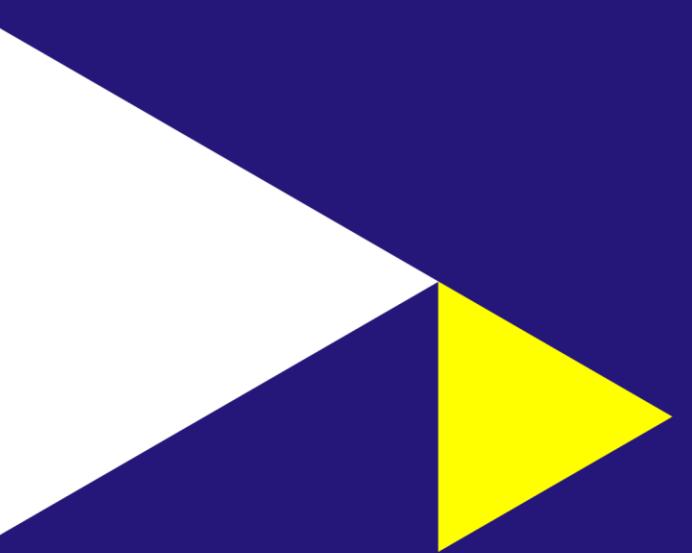
Hogeschool van Amsterdam

minor AAI - Deep Learning Generative AI for images

Michiel Bontenbal & Maarten Post

Woensdag 16 oktober 2024

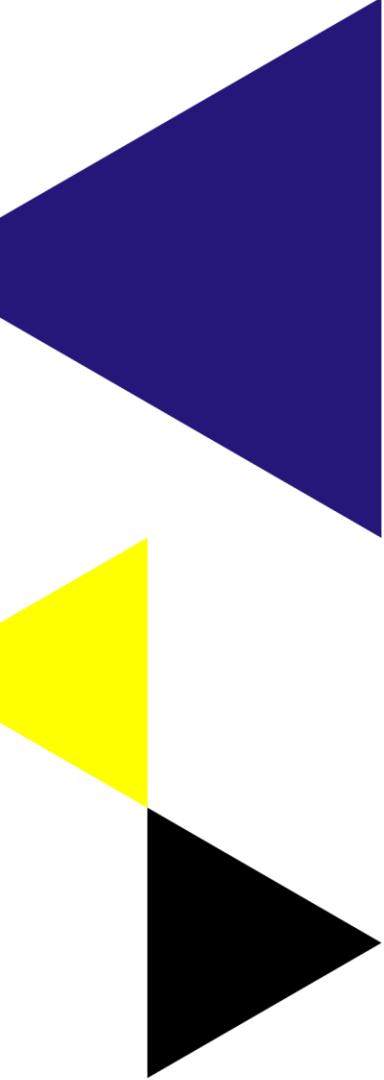
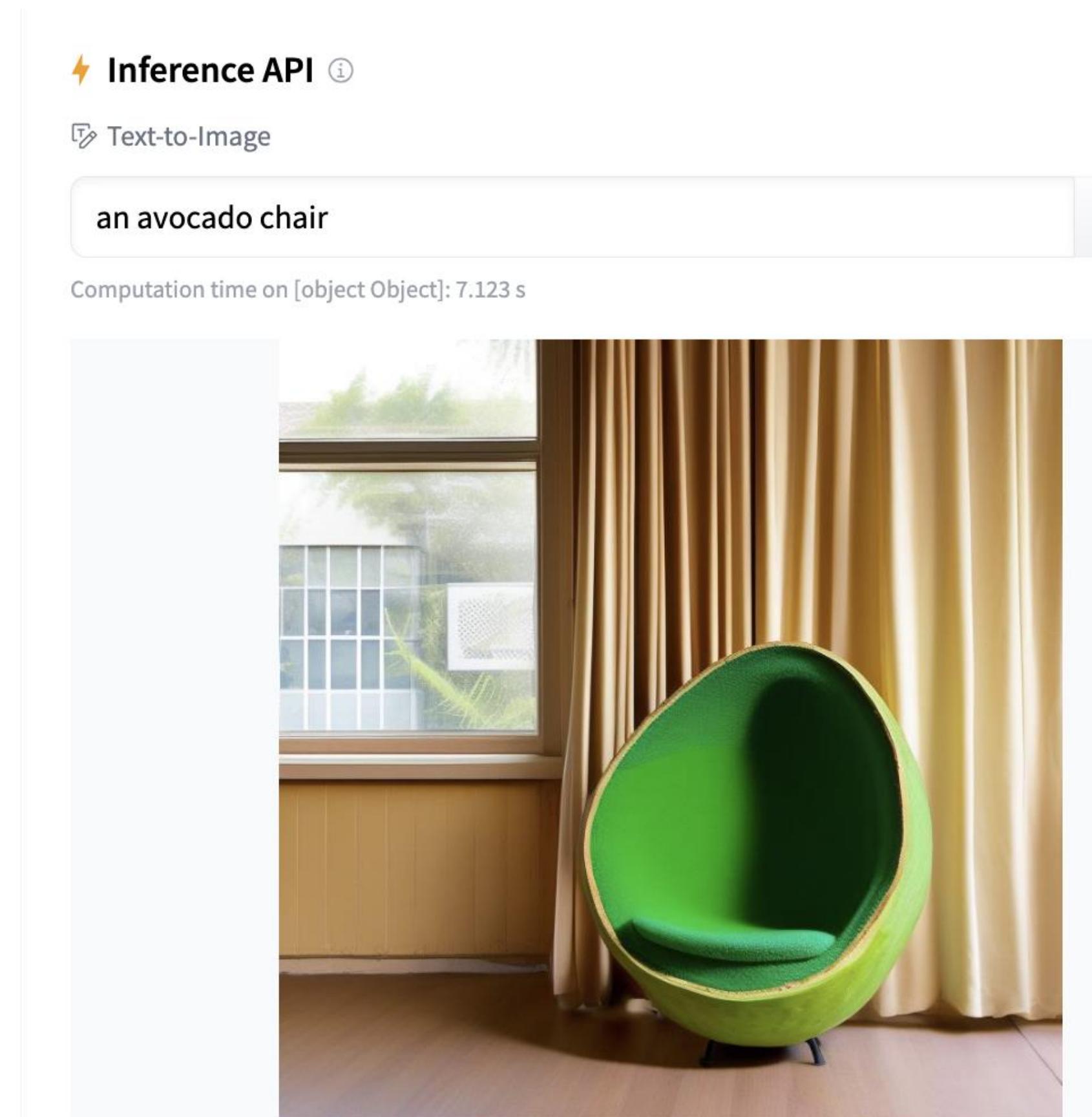
Minor Applied Artificial Intelligence – Hogeschool van Amsterdam



Creating Tomorrow

Computer Vision les 3

- Intro GenAi for Images
- Autoencoders en Variational Autoencoders
 - Notebook
- Problemen met GenAI
- Diffusion algoritme
 - Notebook



Computer Vision sessies

Bootcamp 1: Best of Computer Vision

- MNIST handgeschreven cijfers met NN en CNN's
- Image embeddings & similarity search
- Vision Language Models with ollama

Bootcamp 2: Pre-processing

- OpenCV: Edge Detection en gezichtsherkenning
- Numpy for images

Deep learning 1 - GenAI for images (vandaag)

- **(Variational) Autoencoder**
- **Diffusion models**

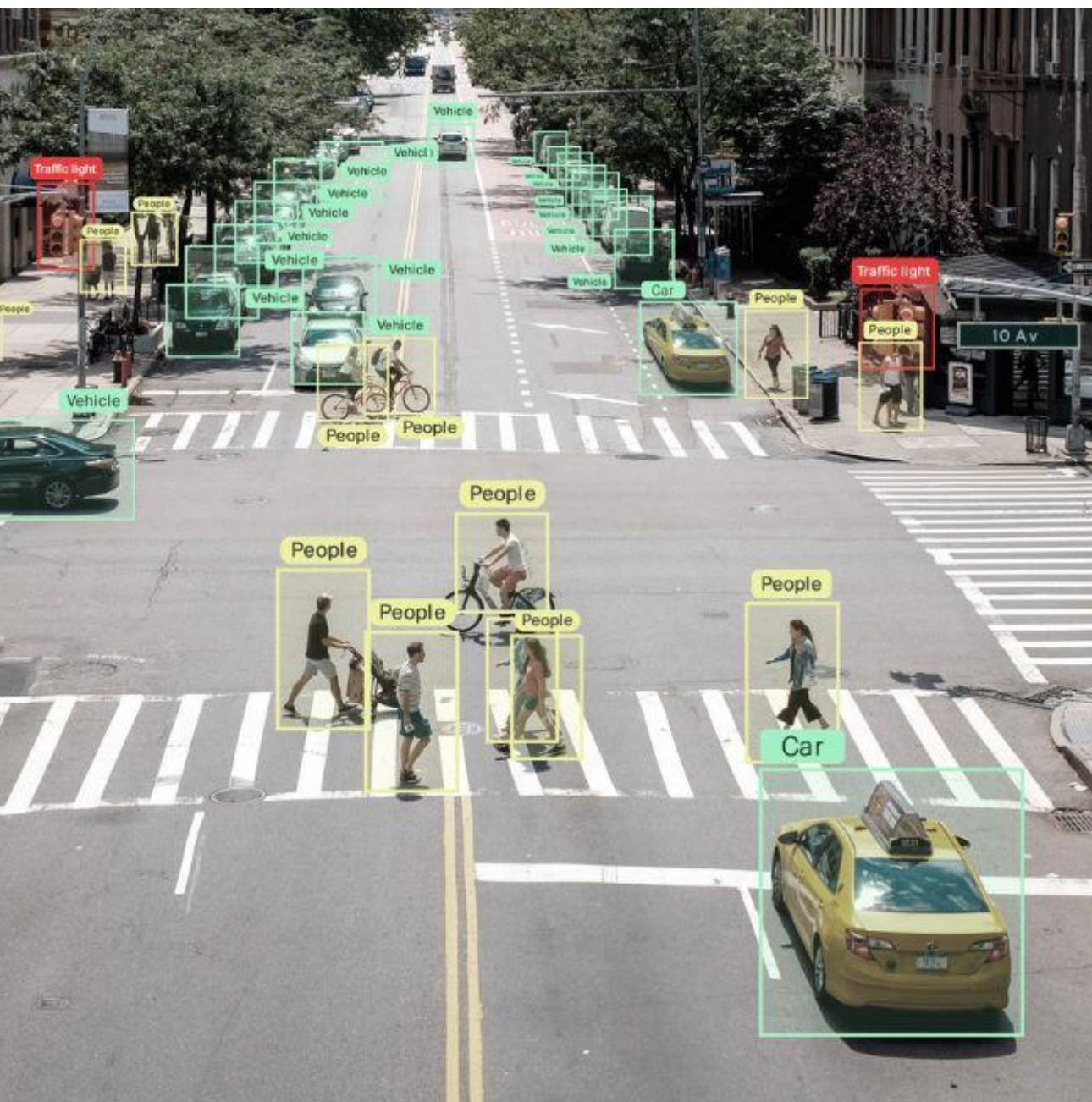
Deep Learning 2 – From DL to Foundation models (23 oktober)

- Segmentatie en object detectie
- Nadelen Convolutional Neural Nets + adversarial neural nets
- Foundation models & self supervised learning

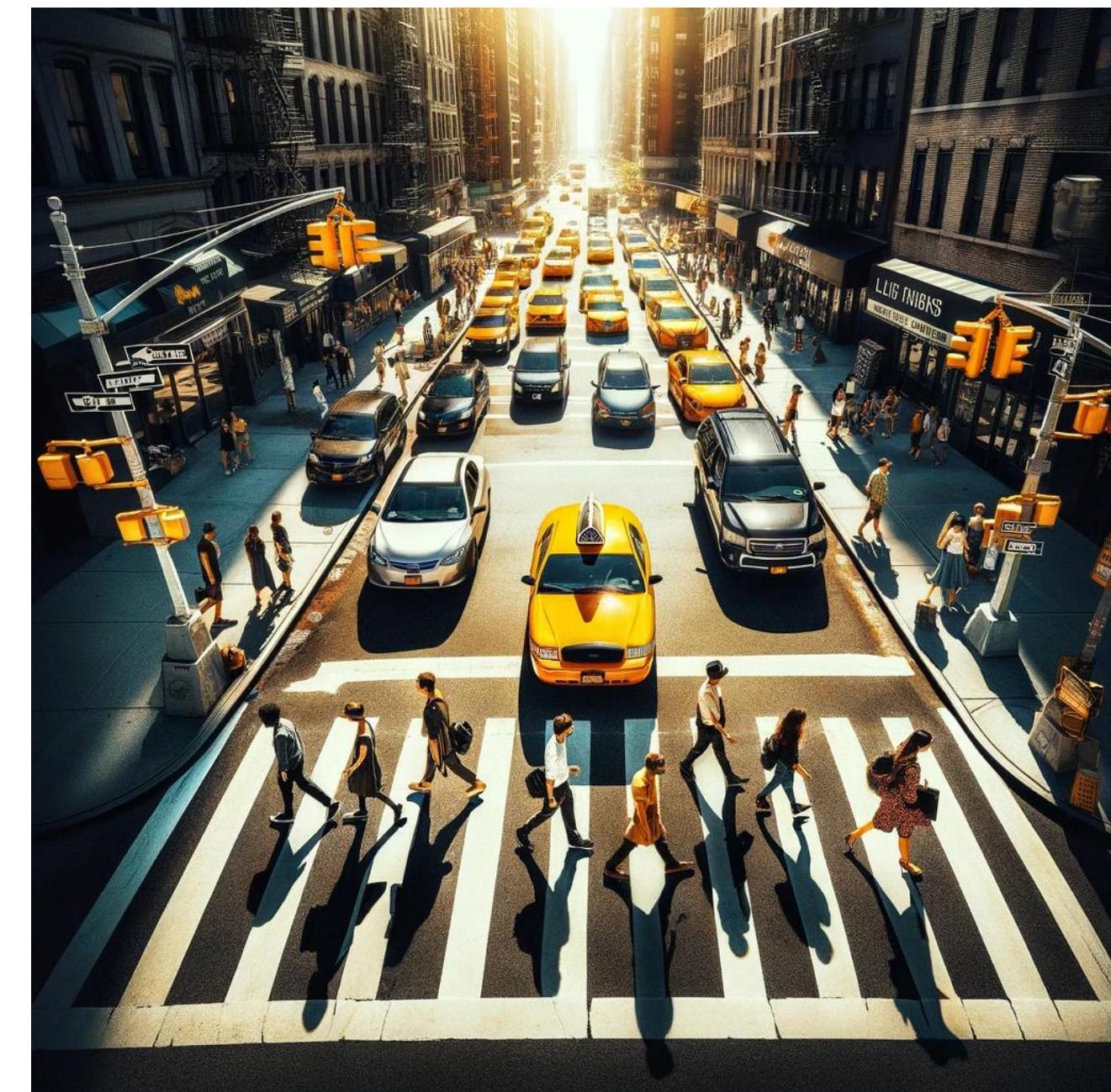
Advanced Topics: Data Centric AI (19 november)

Computer Vision vs Generative AI

CV = Extracting information



Generating images



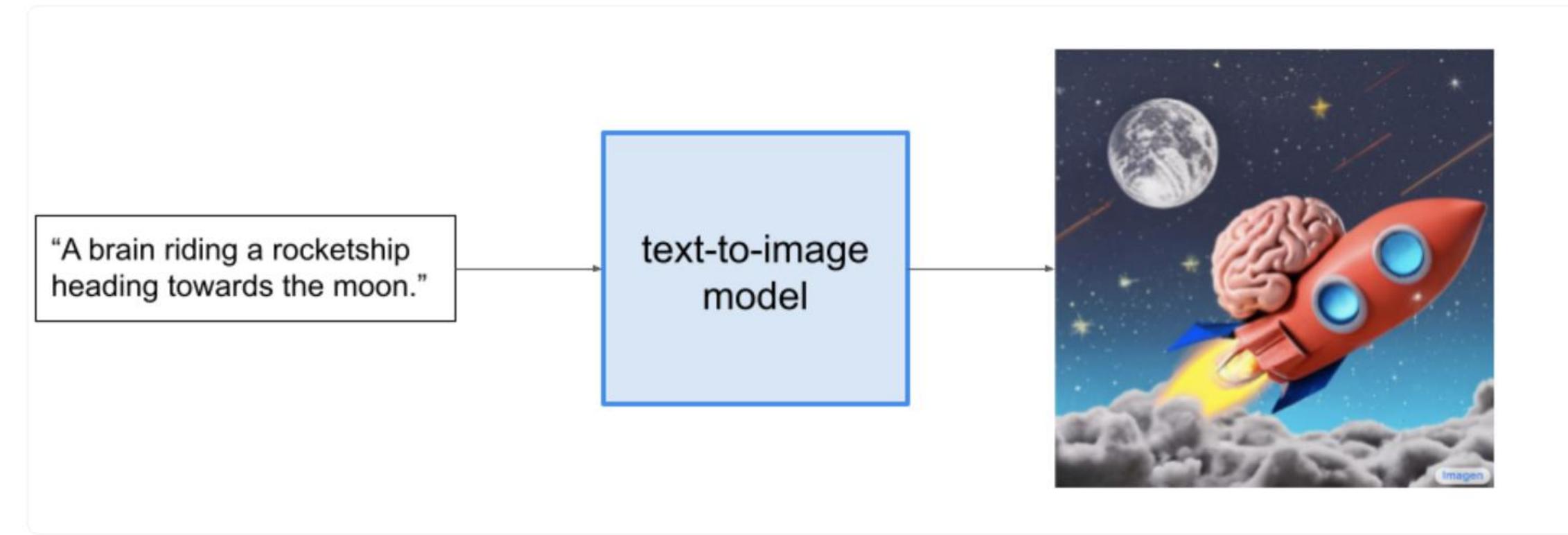
Prompt: *create an image of a sunny street in manhattan with a birdseye view with a yellow cab waiting for a traffic light and 5 people on the pedestrian crossing. Some cars in the background, dall-3, mar 2024*

image-2-image and text-2-image

image-2-image

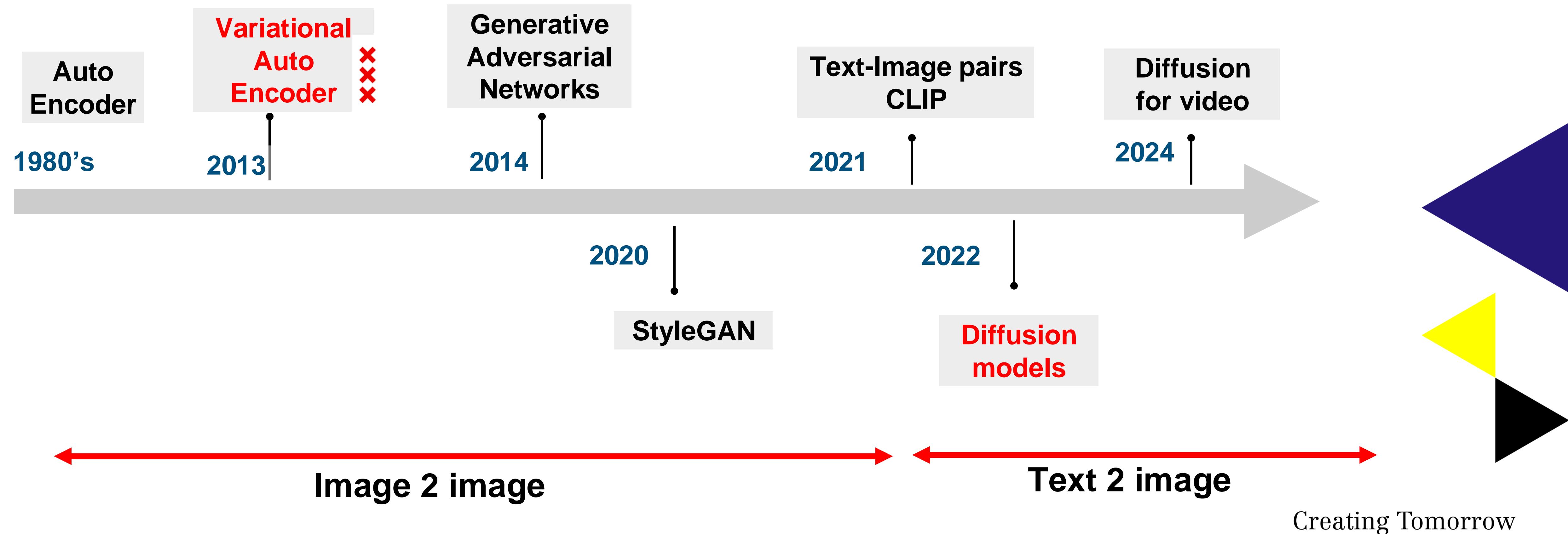


text-2-image

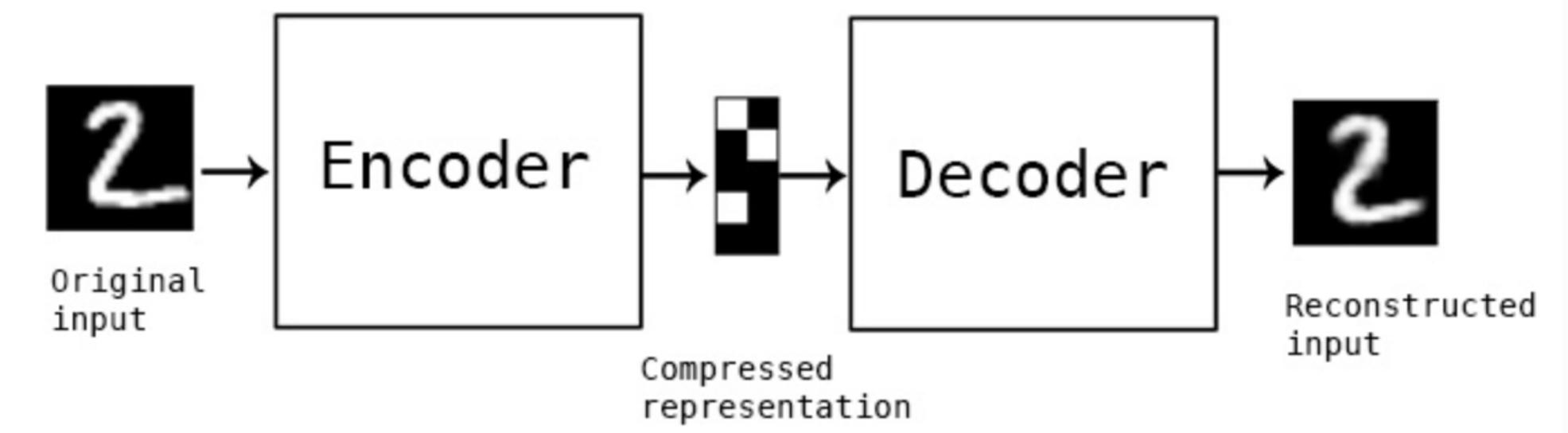


Creating Tomorrow

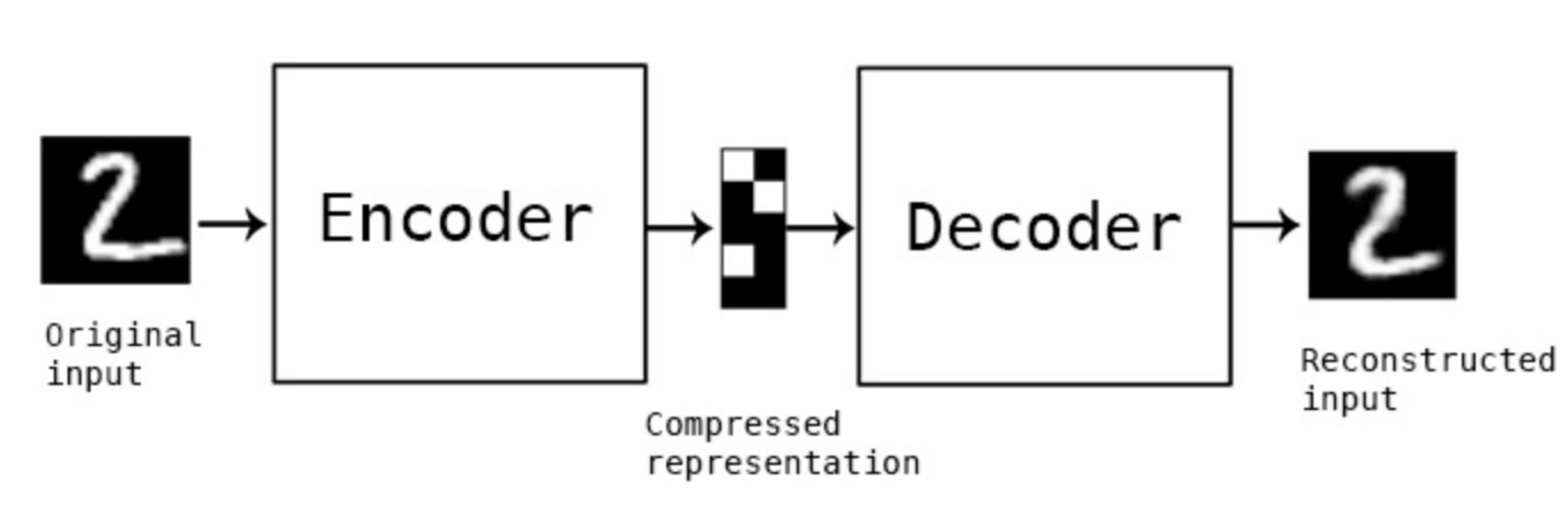
Tijdslijn Generative AI for Vision



Autoencoder en Variational Autoencoder



Met autoencoder maak je nieuw plaatje



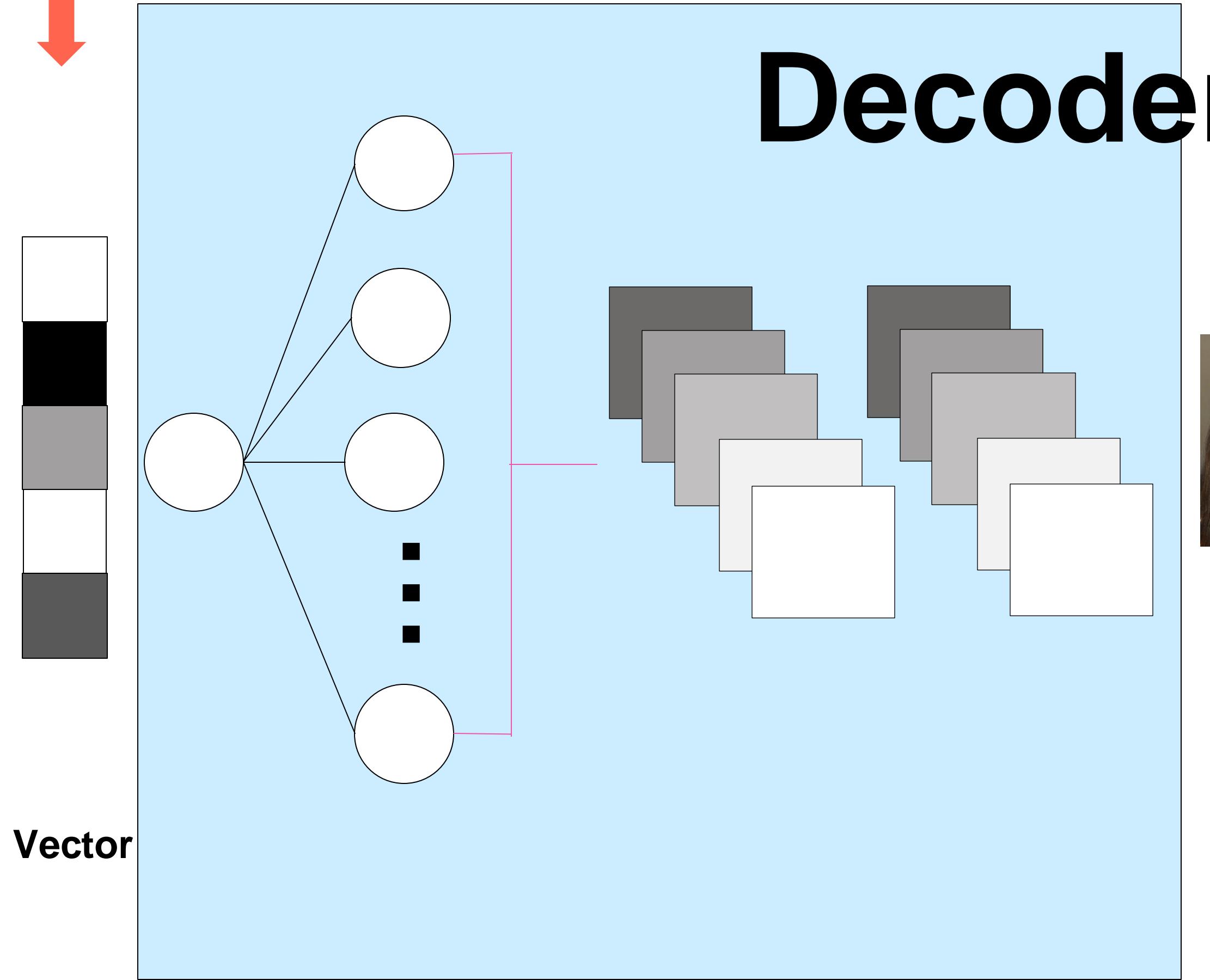
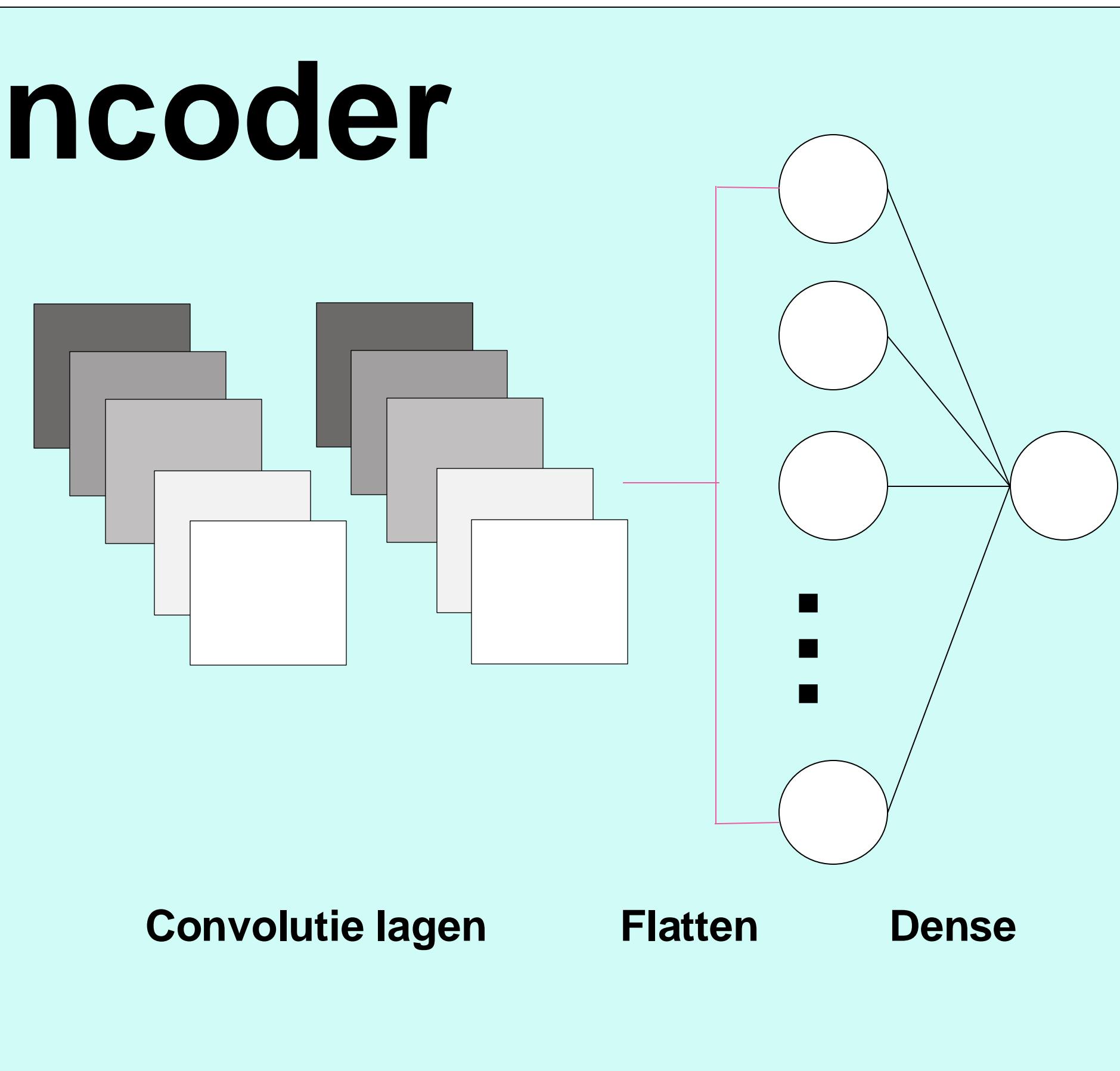
Source: <https://blog.keras.io/building-autoencoders-in-keras.html>

Creating Tomorrow

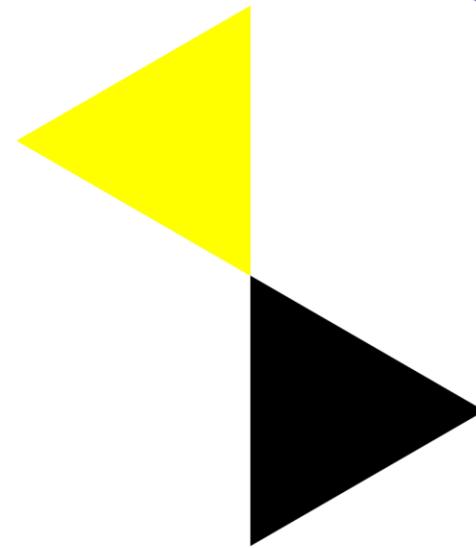
Autoencoder

! Latent Space !

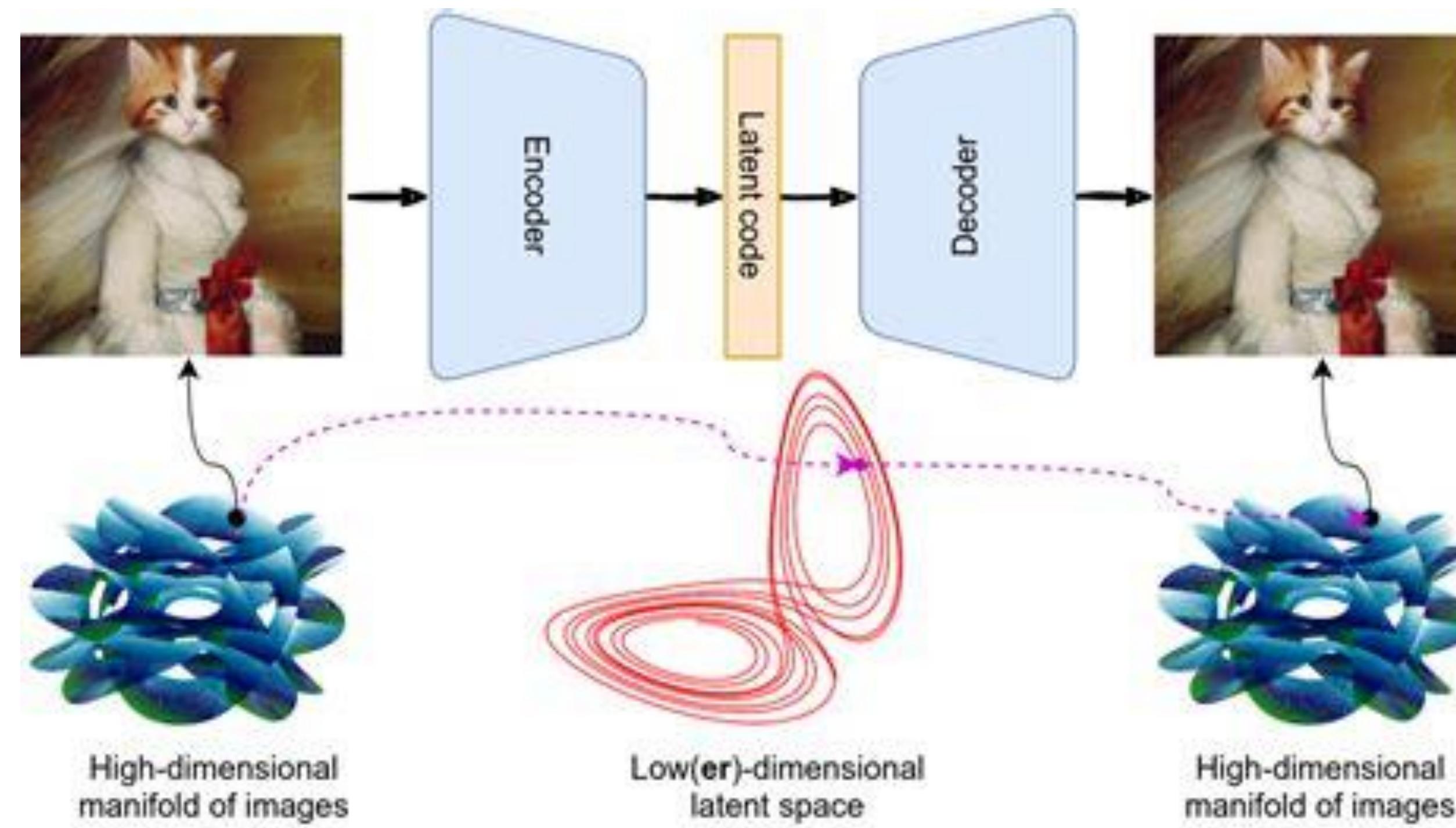
Encoder



Decoder

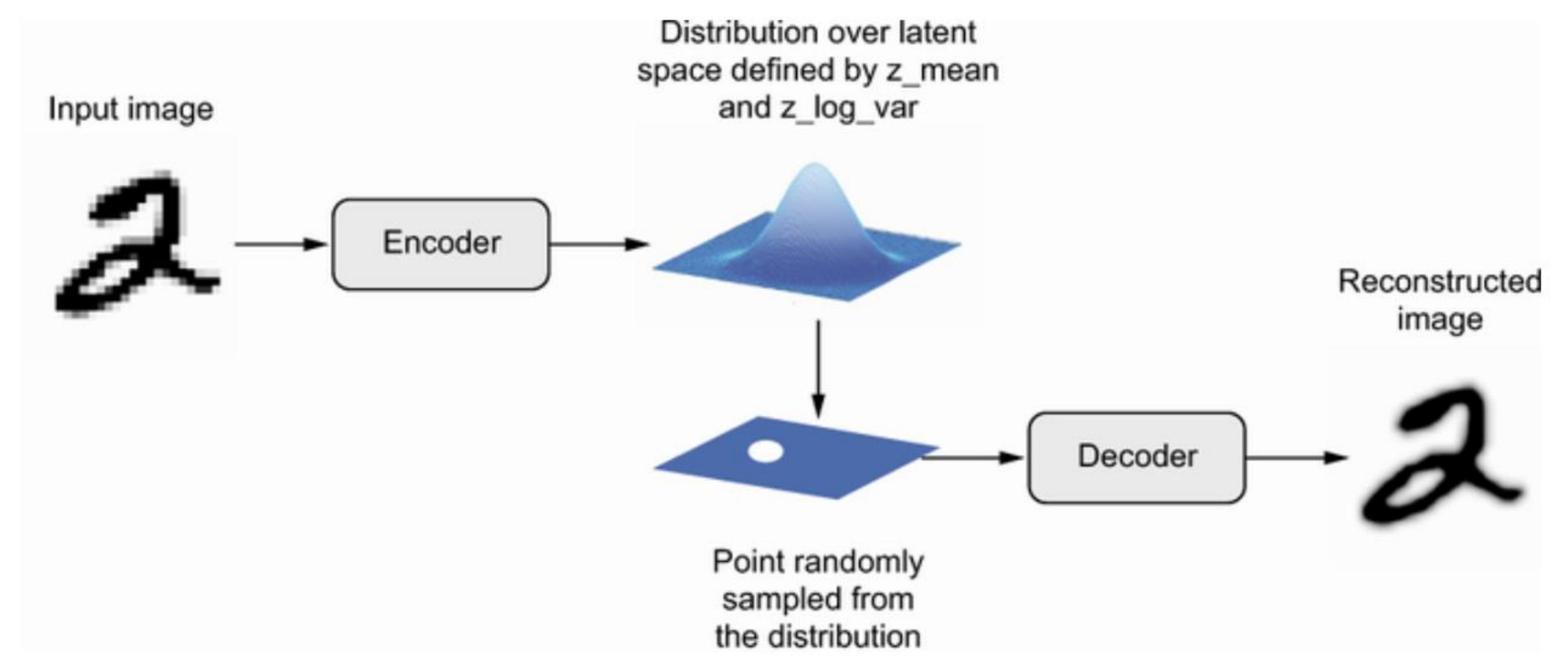


Latent space is where the magic happens



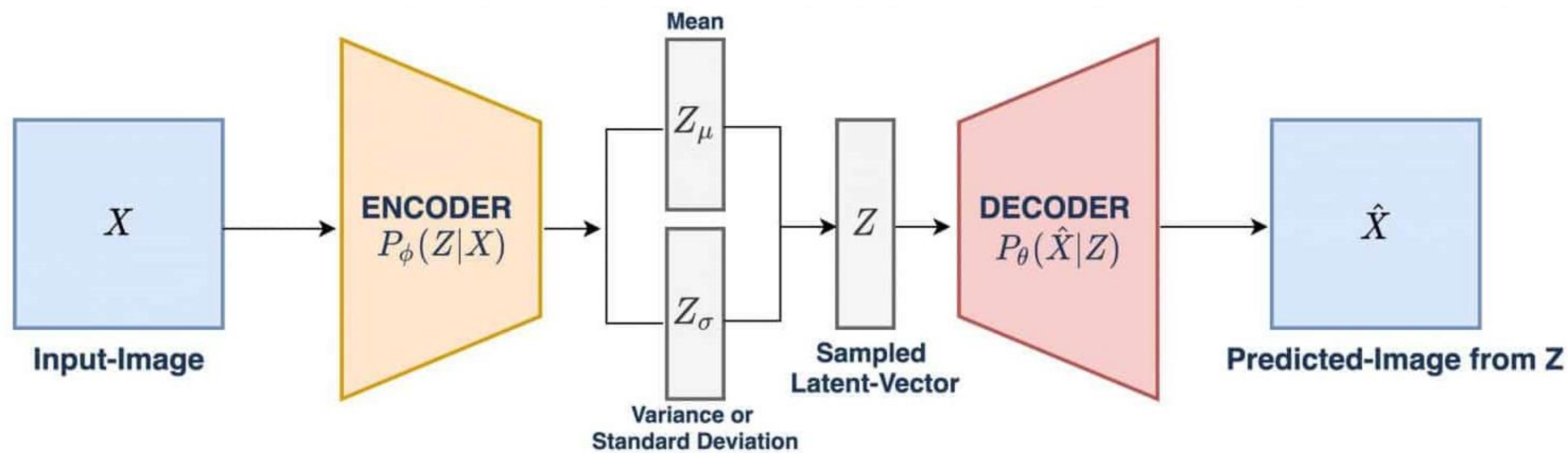
- Source: https://www.perplexity.ai/page/latent-space-101-what-it-is-an-mwXuxYfzS_J4e_uFvOskg

Variational Autoencoder

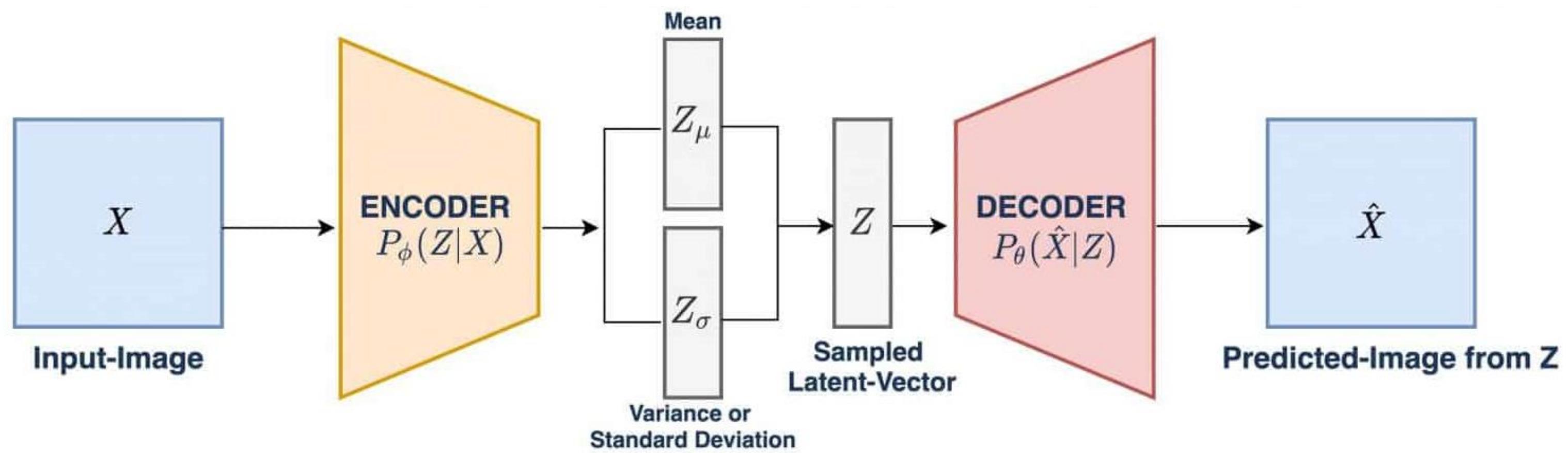


- Source: Deep Learning with Python chapter 12, 2nd edition, Francois Chollet

Variational Autoencoder: encoder – sample layer - decoder



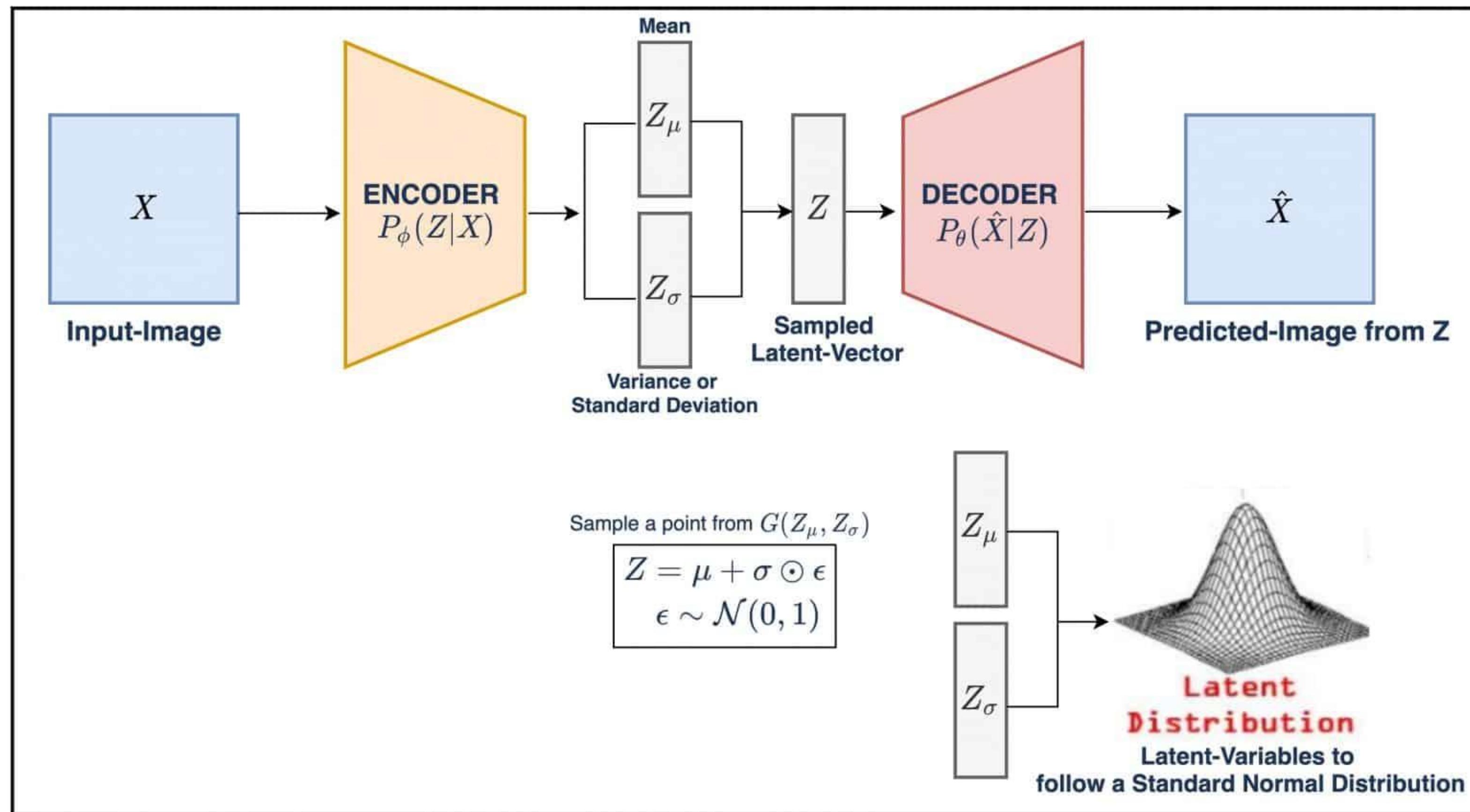
Variational Autoencoder - encoder



De encoder gebruikt dezelfde lagen die we al kennen van CNN: Conv2D, Pooling, Dense

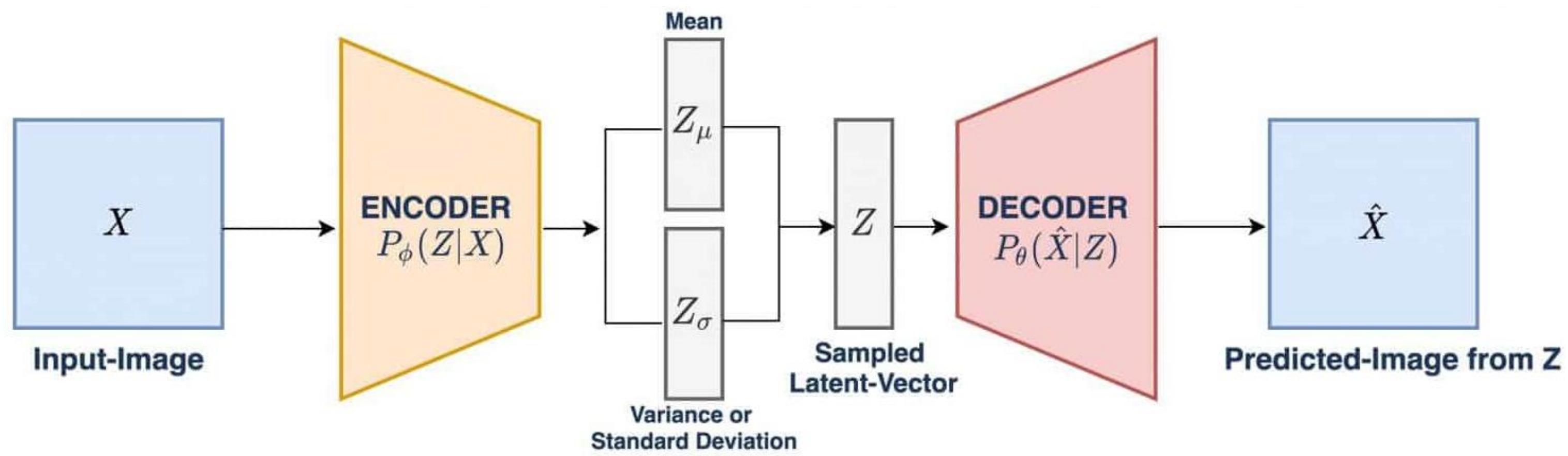
- Source: learnopencv.com

Variational Autoencoder sample layer



- Source: learnopencv.com

Variational Autoencoder - decoder



De decoder gebruikt een Conv2Dtranspose layer

3	2
8	4

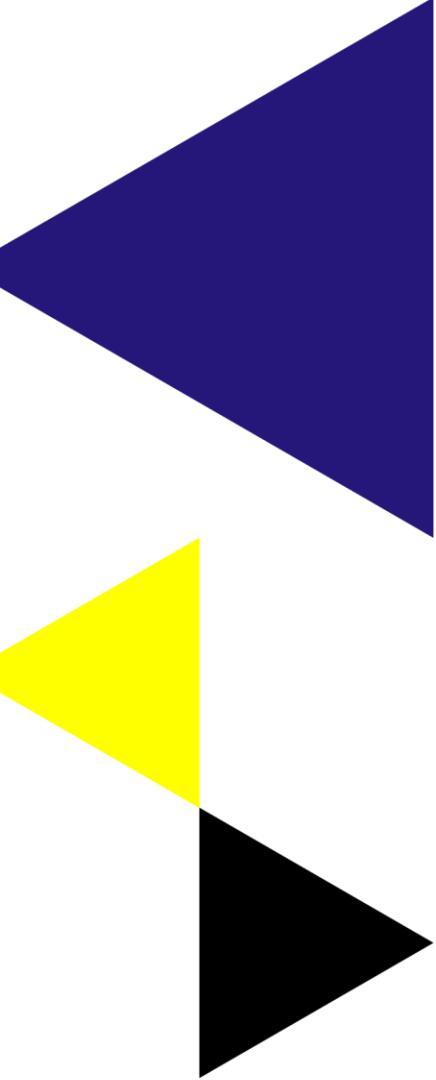
strides = 2

3	0	2	0
0	0	0	0
8	0	4	0
0	0	0	0

- Source: learnopencv.com

Zelf aan de slag

- Vae_opgave.ipynb
- Bespreking om ... uur





Het glas is halfvol Problemen in GenAI



Generative AI snapt het niet



Prompt: 'A glass half full with sparkling water'
Dall-e / ChatGPT, 27 feb 2024



ChatGPT

You're right, and I apologize for the confusion earlier. When we say a glass is "half full," it generally means that the liquid inside the glass fills it up to 50% of its total capacity, not that it's divided left to right or in any other manner. The term "half full" refers to the level of the liquid being at the midpoint from the bottom of the glass to the top, effectively making half of the glass's volume occupied by the liquid and the other half empty. This description disregards any division of the glass into sides or sections; it solely concerns the vertical fill level. I appreciate your patience and the opportunity to correct this misunderstanding.

Google Gemini

Prompt: a 17th century British king eating a watermelon.

Business / Tech

Google halts AI tool's ability to produce images of people after backlash

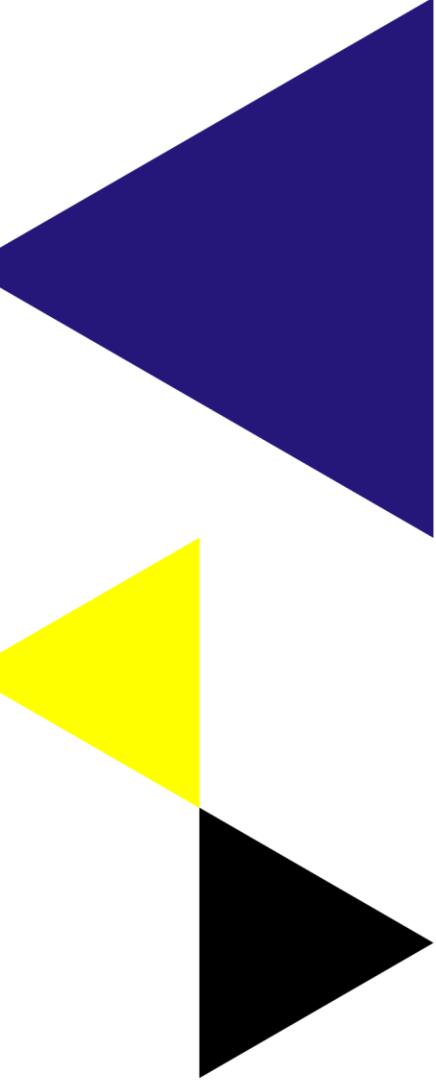
By Catherine Thorbecke and Clare Duffy, CNN
🕒 3 minute read · Updated 2:28 PM EST, Thu February 22, 2024



March 2024

ating Tomorrow

OpenAI Sora

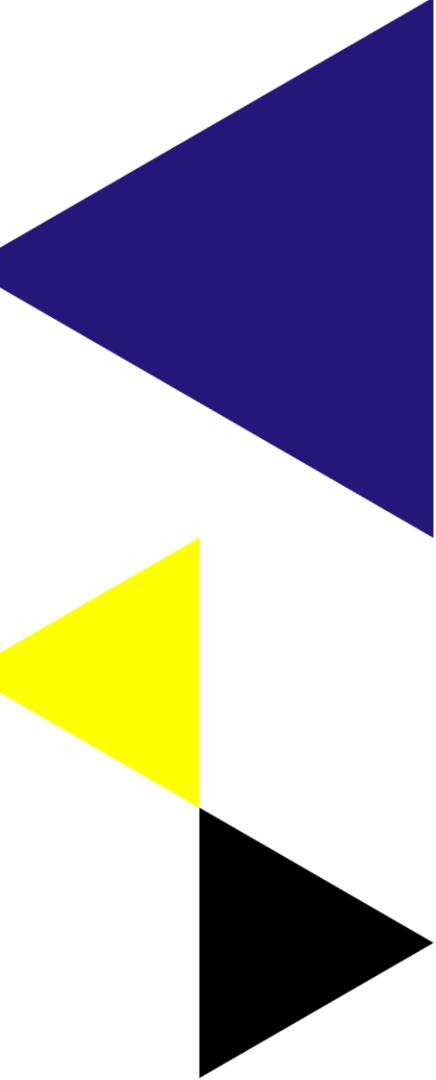


<https://www.youtube.com/watch?v=jspYKxFY7Sc>, feb 2024

Creating Tomorrow

Welke problemen met GenAI?

1. Niet logisch of realistisch
 - Gebrek aan ‘world model’
2. Copyright
3. Deep fakes

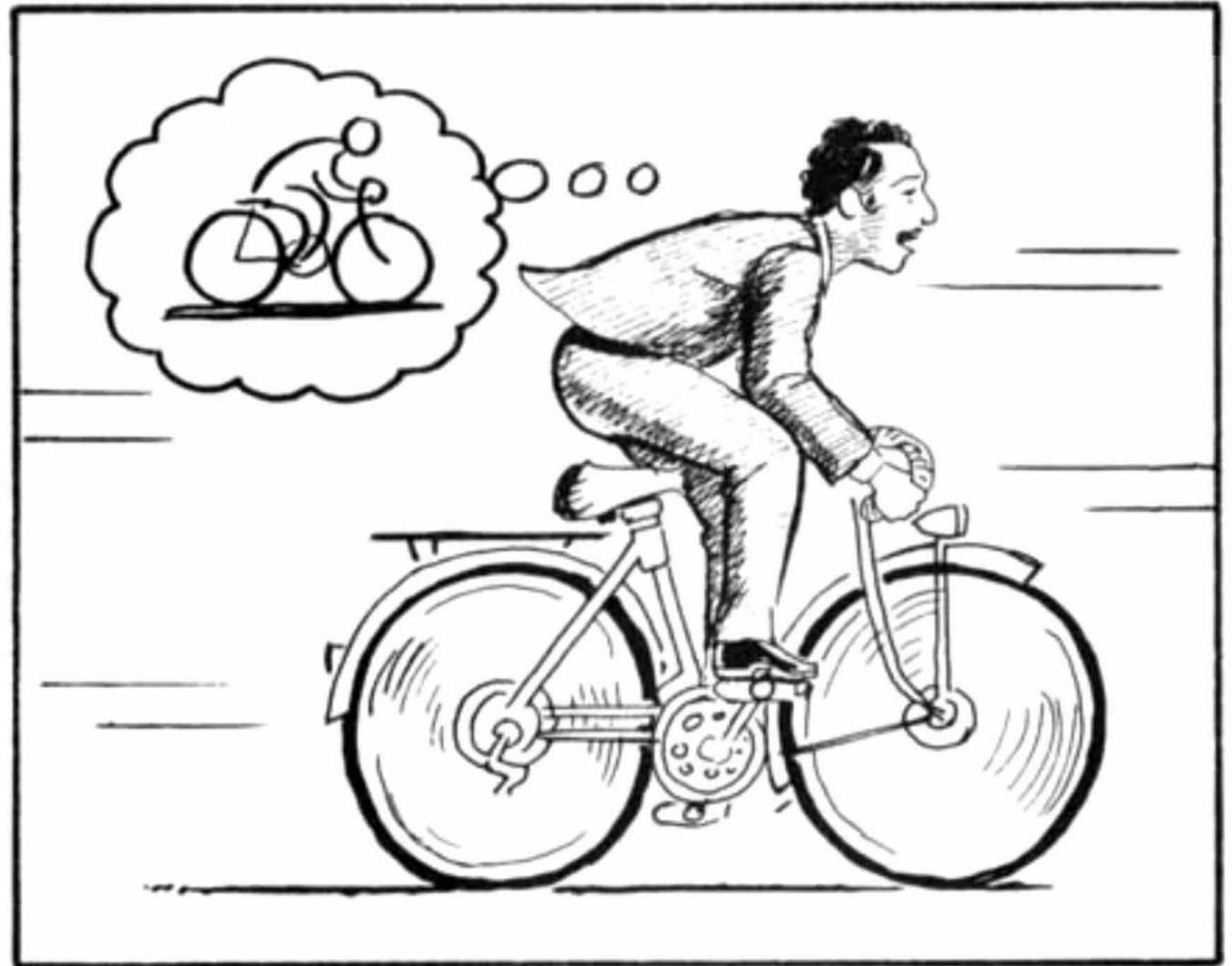


World model

Heeft een AI-model *begrip* van de wereld?

Is het AI-model zich bewust van:

- Fysieke wereld (natuurkunde / biologie)
- Oorzaak – gevolg



Een AI-model dat begrip heeft van de wereld is cruciaal voor betere GenAI modellen.

En ook voor robots.

Source: <https://worldmodels.github.io/>

World model voorbeelden

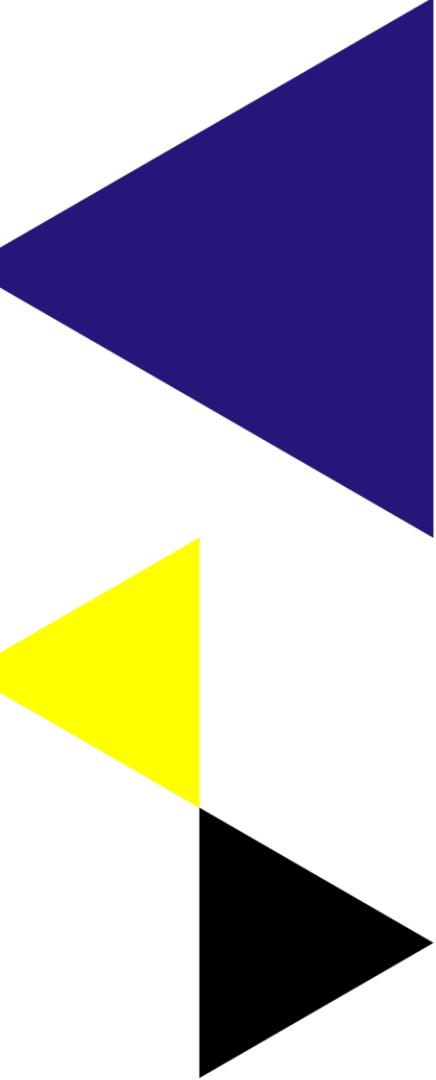
Geen begrip van natuurkunde of biologie



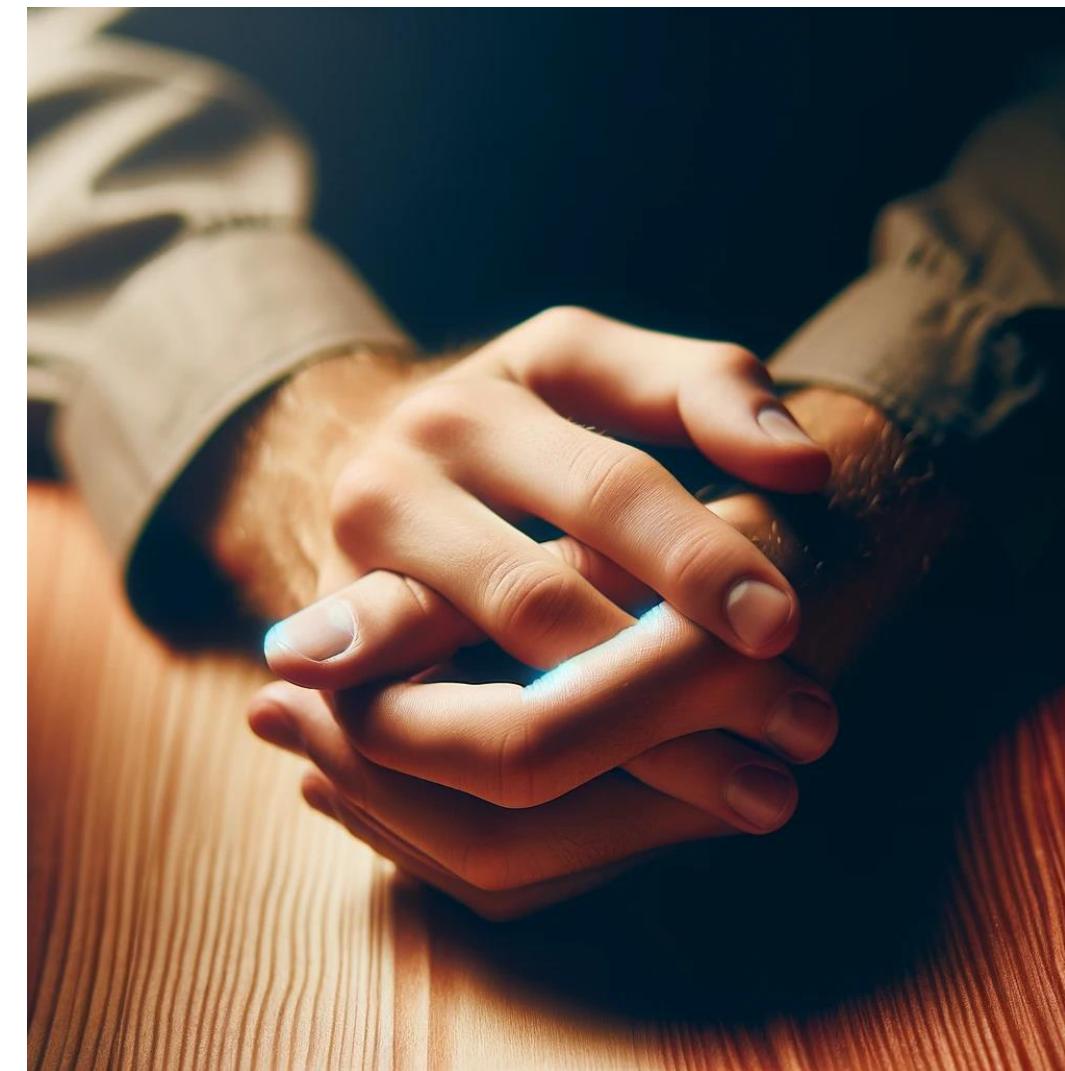
“A glass half full..”



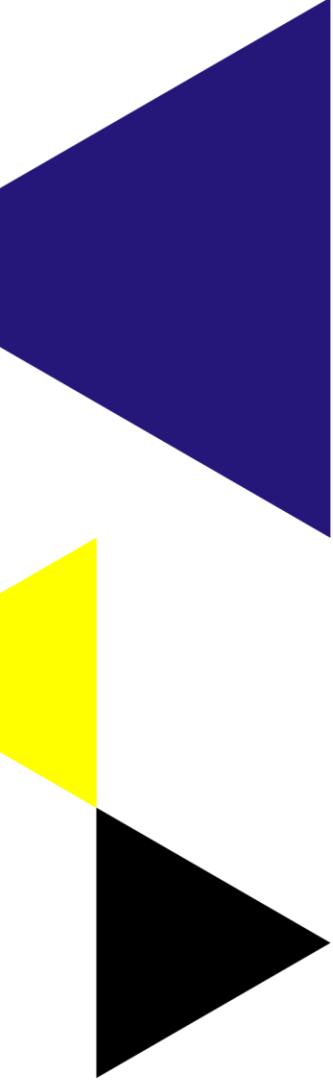
Mier met 4 poten



World model: fingers



“Dear God, please make my supporters not notice this AI-generated photo of me praying in church has 6 fingers. Amen”



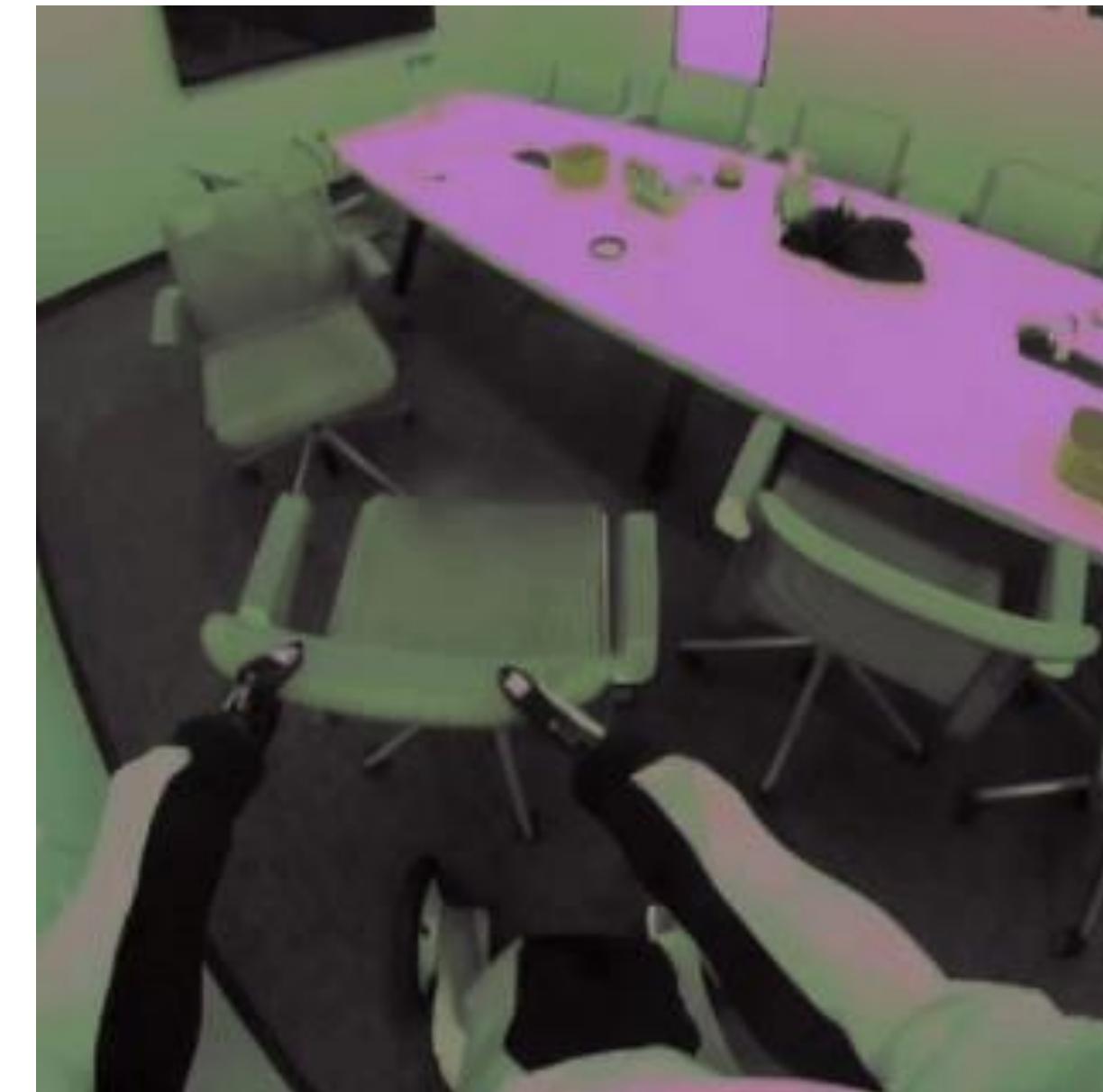
Soms is het gewoon gek...

Flux.Schnell 'A canal in Amsterdam', oktober 2024



row

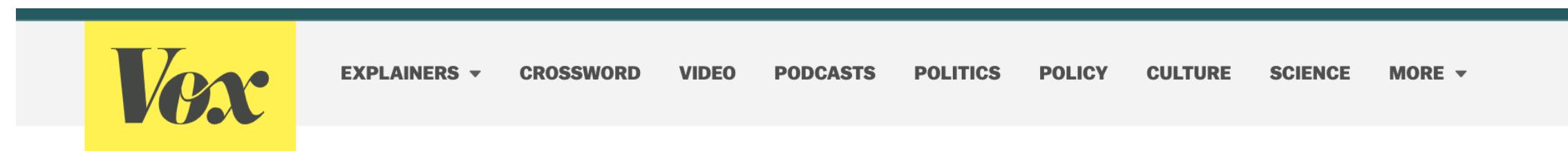
Voor robots cruciaal om de wereld te begrijpen en te voorspellen



- Om dit te voorspellen gebruiken robots ook GenAI
- Source: <https://www.1x.tech/discover/1x-world-model>

Copyright wordt massaal geschonden

- De datasets met images van OpenAI bevatten copyrighted afbeeldingen
 - OpenAI is nu verwikkeld in meerdere rechtszaken.
- Dit geldt ook voor datasets van Meta of Google



How copyright lawsuits could kill OpenAI

The New York Times v. OpenAI, explained.

By Adam Clark Estes | @adamclarkestes | ace@vox.com | Jan 18, 2024, 7:00am EST

- N.B. Het Duitse onderzoekscollectief LAION heeft recent rechtszaak gewonnen. Omdat zij alleen onderzoek doen mogen ze wel foto's met copyright gebruiken. Is eventueel een alternatief voor jou?

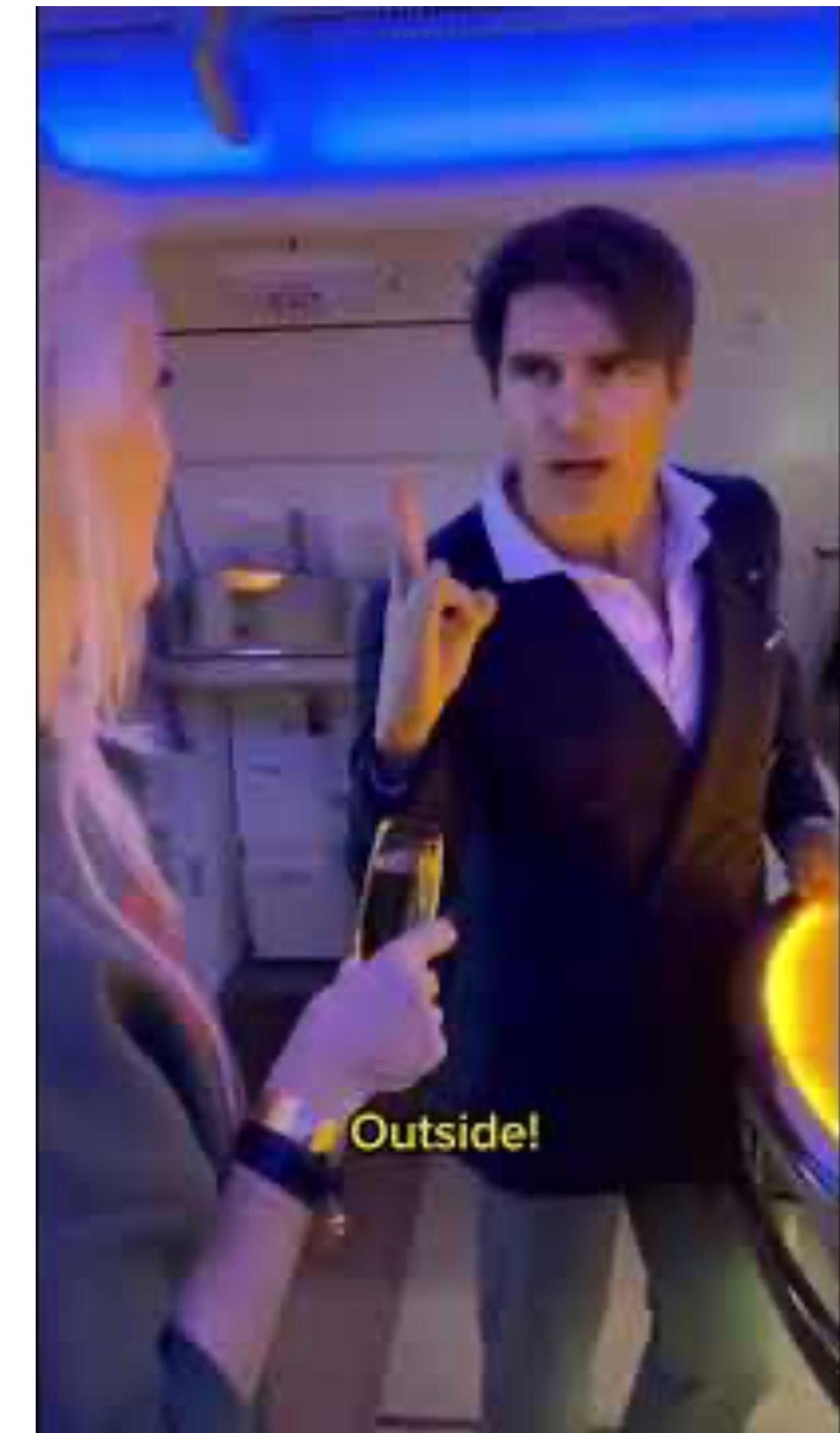


Home > Tech

More news organizations are suing OpenAI for copyright infringement

OpenAI is in its lawsuit era.

Deep fakes



tomorrow

Deep fakes are polluting the internet

Google X |

Alle Afbeeldingen Video's Nieuws Web Boeken Financieel Tools

White Images real Called Beautiful For sale Male Images free Pictures cute

Tip: Toon resultaten in het **Nederlands**. Je kunt ook meer informatie doornemen over [hoe je filtert op taal](#).

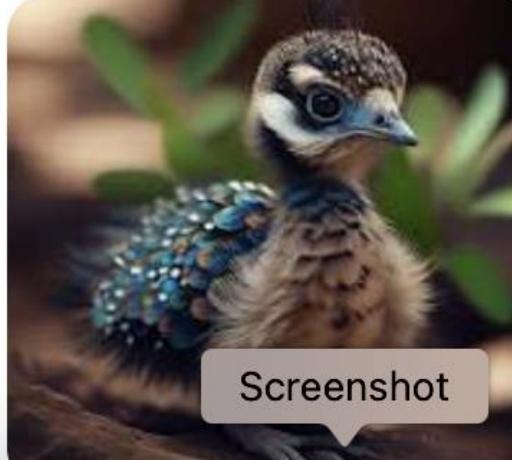
Afbeeldingen :



Baby Peacocks: All You Need ...
Birdfact

This is what a baby peacock l...
Reddit

Rare White Peacock - Baby P...
TikTok

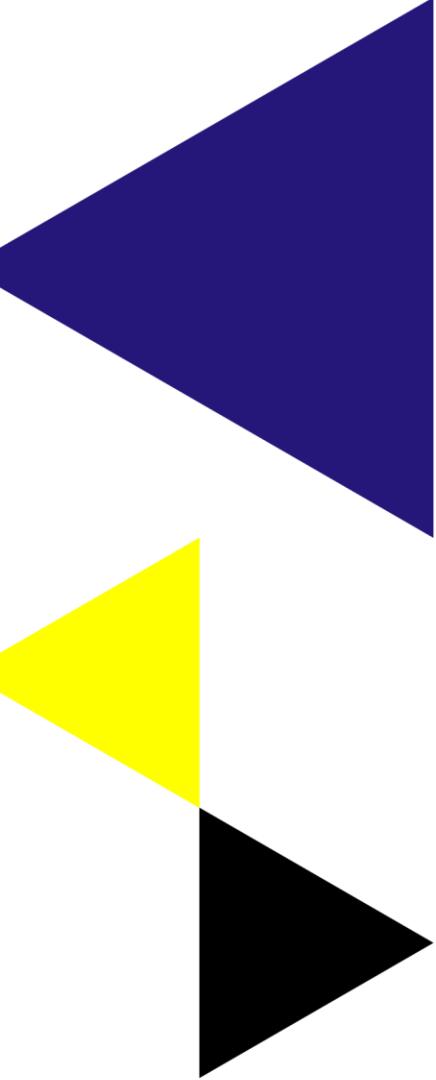


Screenshot

Creating Tomorrow

Samengevat

1. Niet logisch of realistisch
 - Gebrek aan ‘world model’
2. Copyright
3. Deep fakes

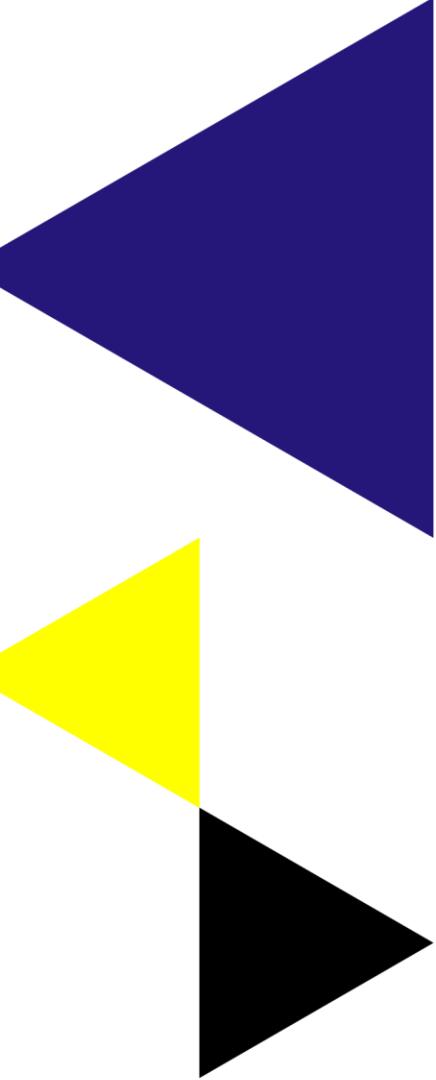




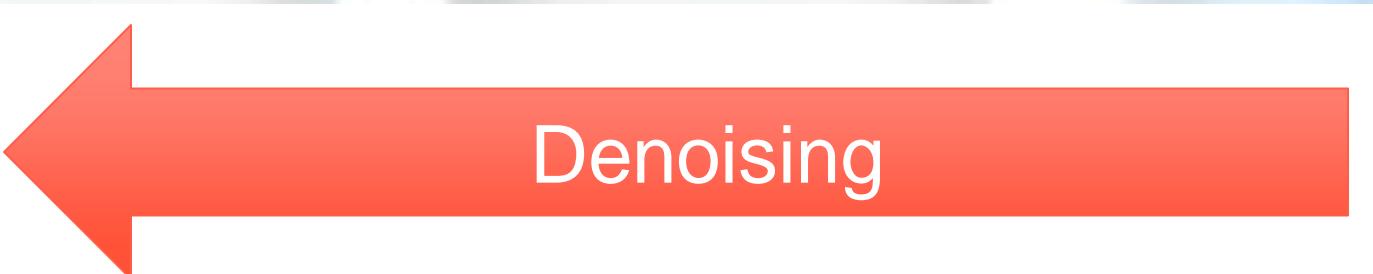
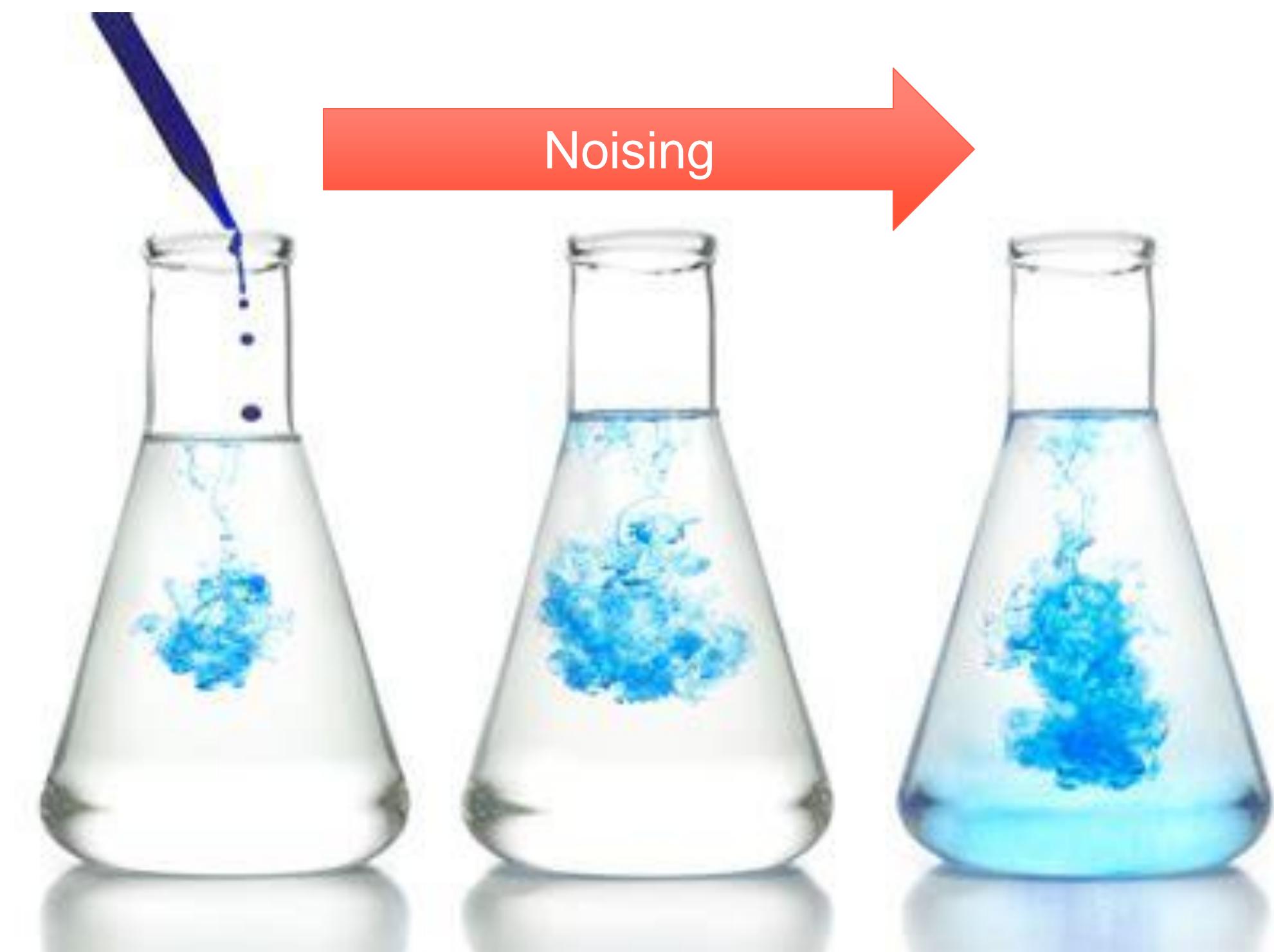
Diffusion models



Diffusie

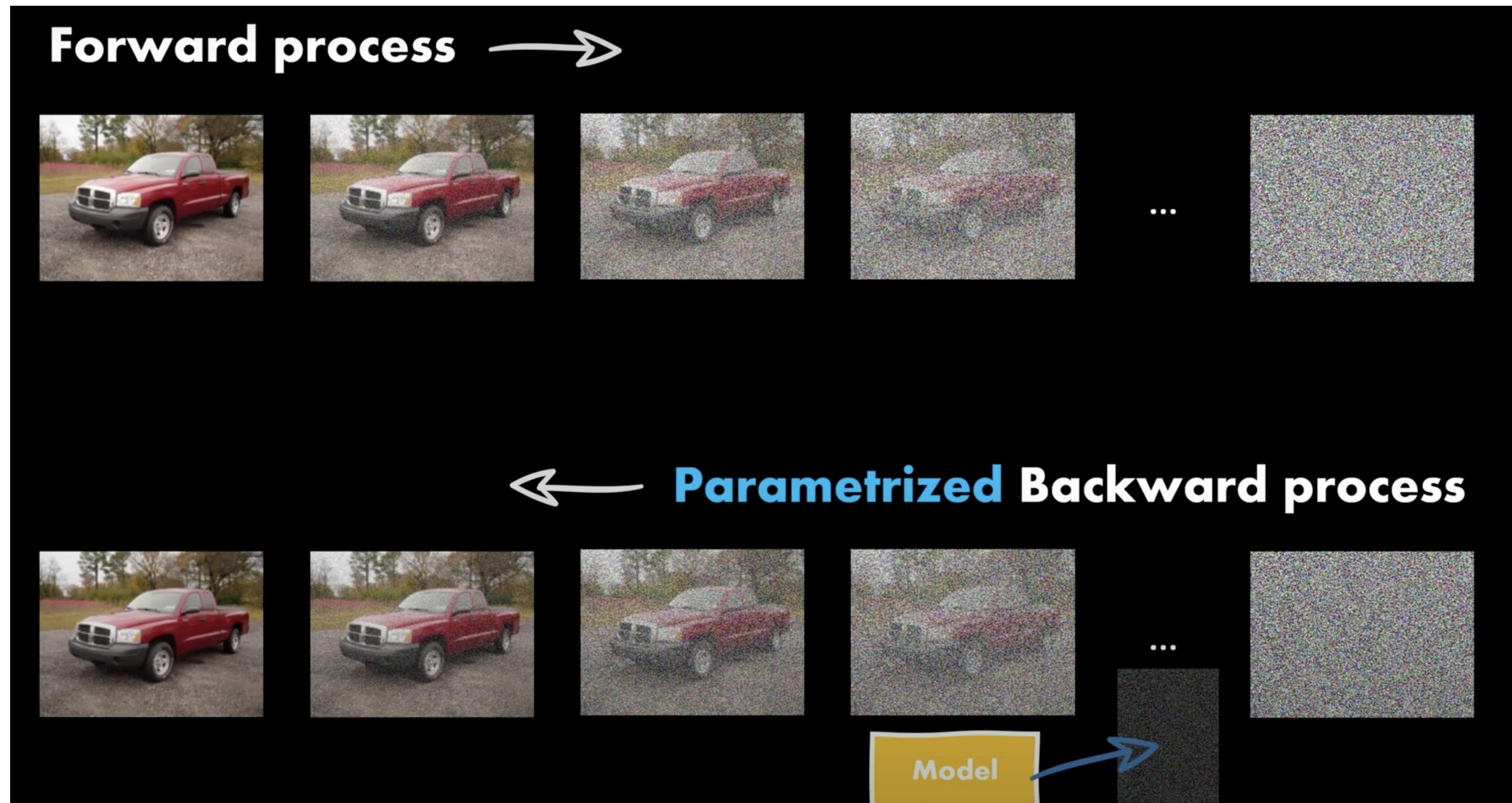


Diffusion model in AI



Creating Tomorrow

Noising and denoising



Creating Tomorrow

Source: <https://www.youtube.com/watch?v=a4Yfz2FxXi>

Denoising an image

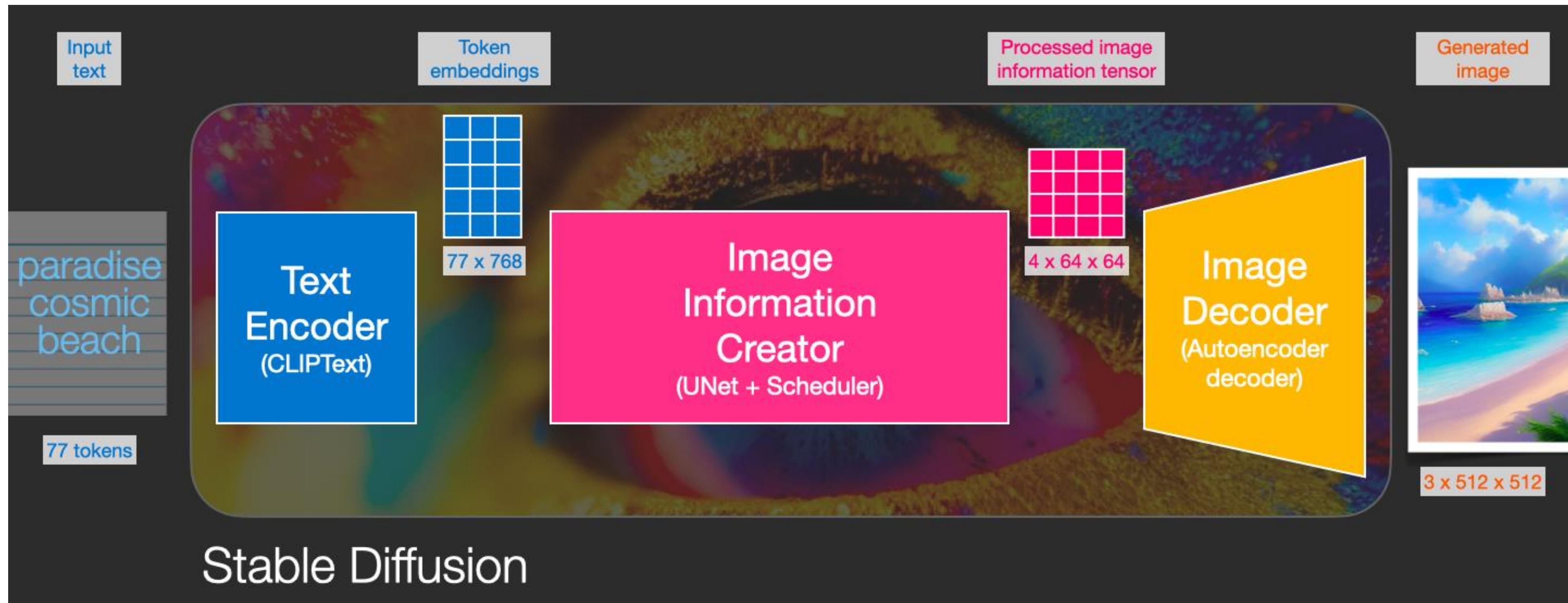


This process of denoising a signal is the secret sauce in diffusion models.

Image gemaakt met stable diffusion.

Source: <https://towardsdatascience.com/audio-diffusion-generative-musics-secret-sauce-f625d0aca800>

Diffusion algoritme

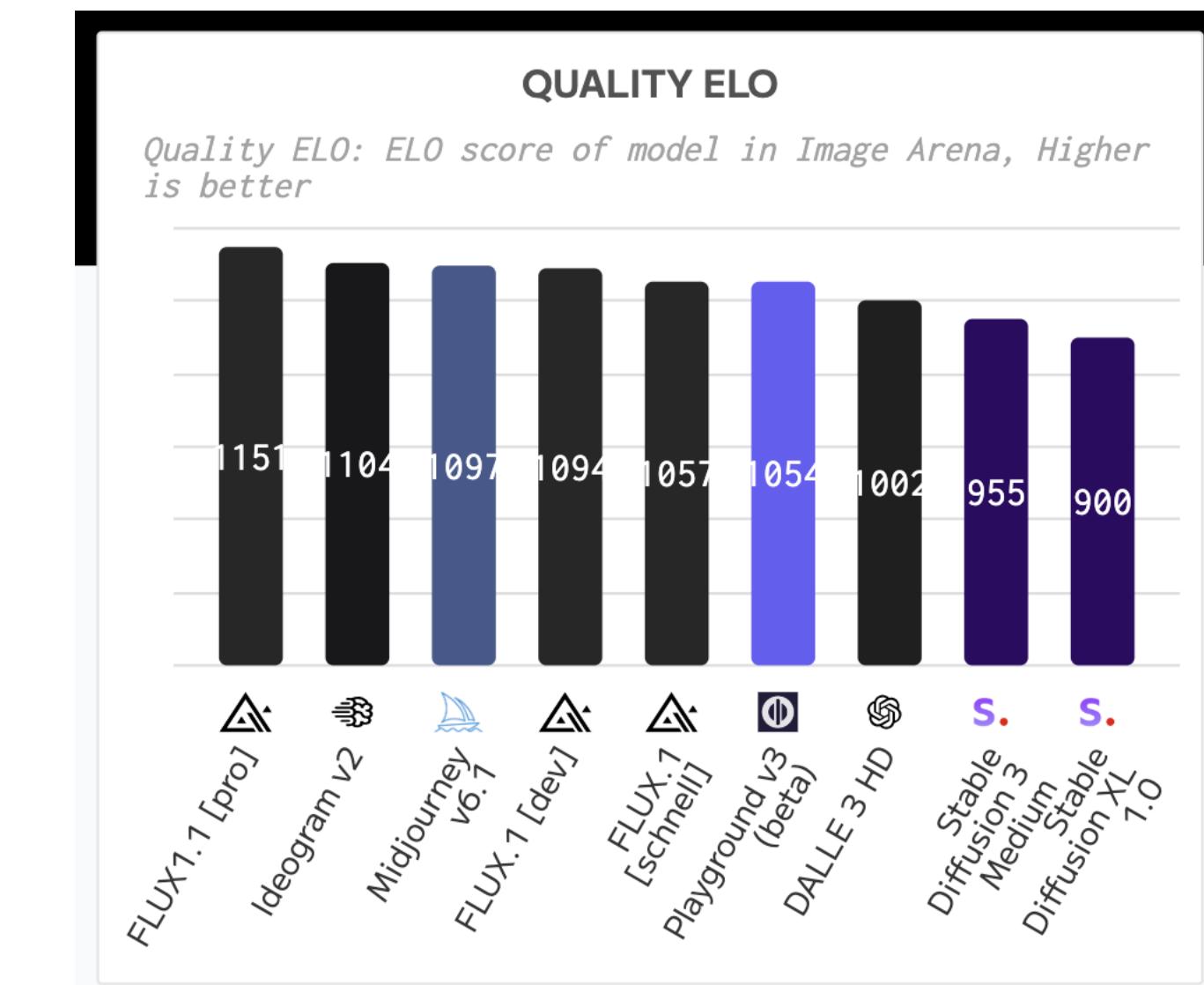


source: <https://jalammar.github.io/illustrated-stable-diffusion/>

Creating Tomorrow

Beste Image Generators (okt 2024)

1. Flux EU
2. Ideogram
3. Midjourney



Source: <https://artificialanalysis.ai/text-to-image>

Read more: <https://www.deeplearning.ai/the-batch/text-to-image-generators-face-off-in-arena-leaderboard-by-artificial-analysis/>

Zelf maken via Huggingface



Maak een image met FLUX op Huggingface:
En bestudeer even Huggingface.

- <https://huggingface.co/black-forest-labs/FLUX.1-schnell>

Some tasks for Diffusion Models



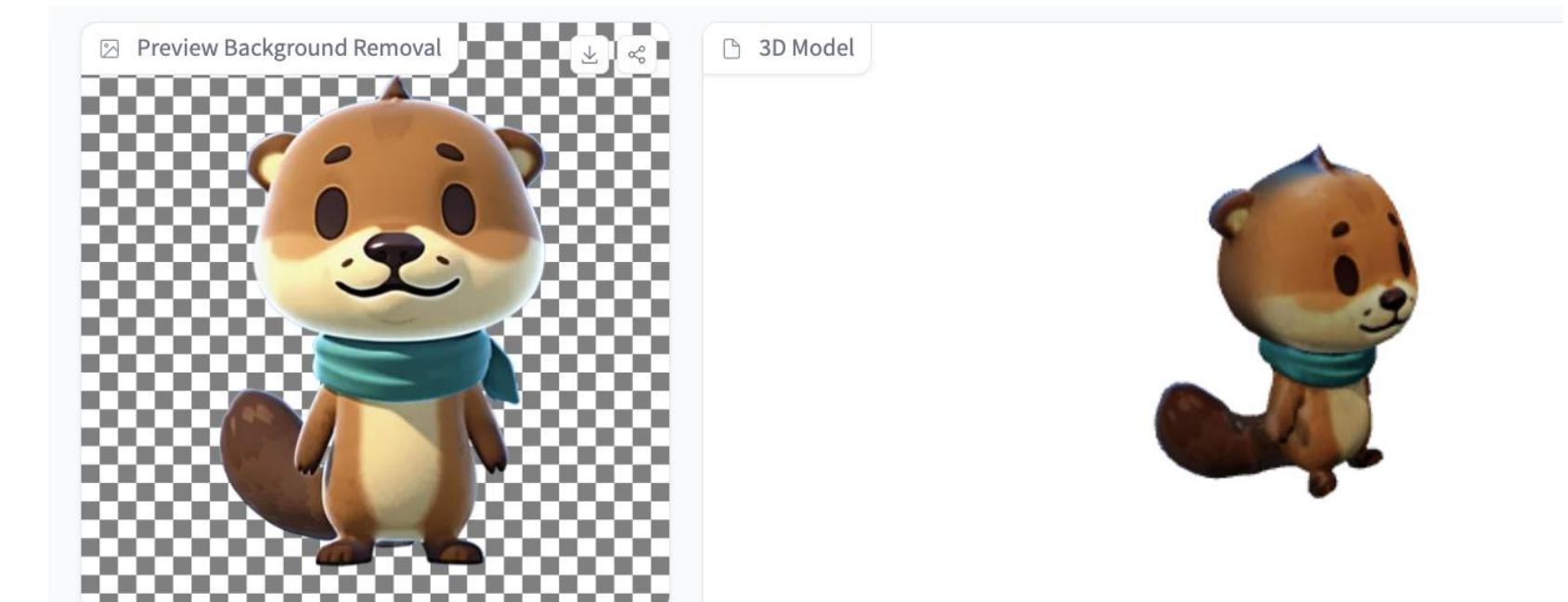
Image generation



Inpainting



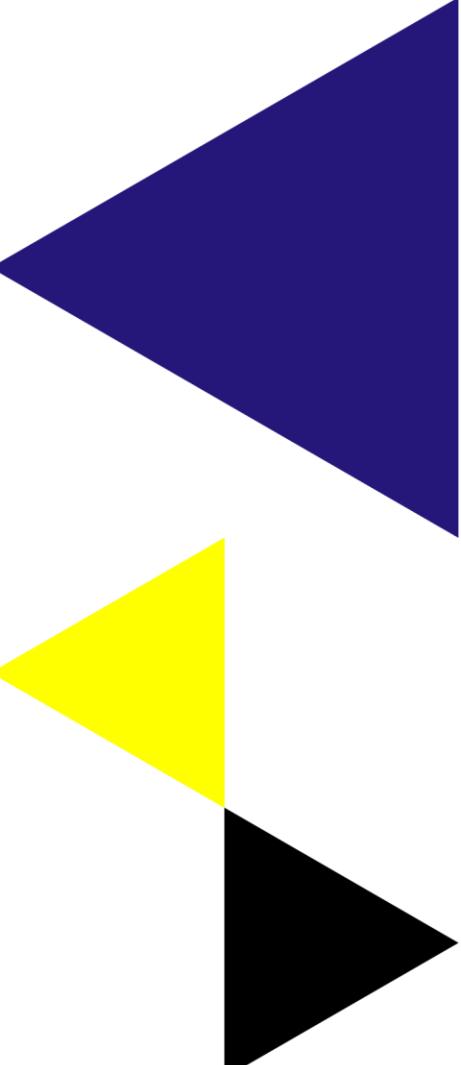
Outpainting



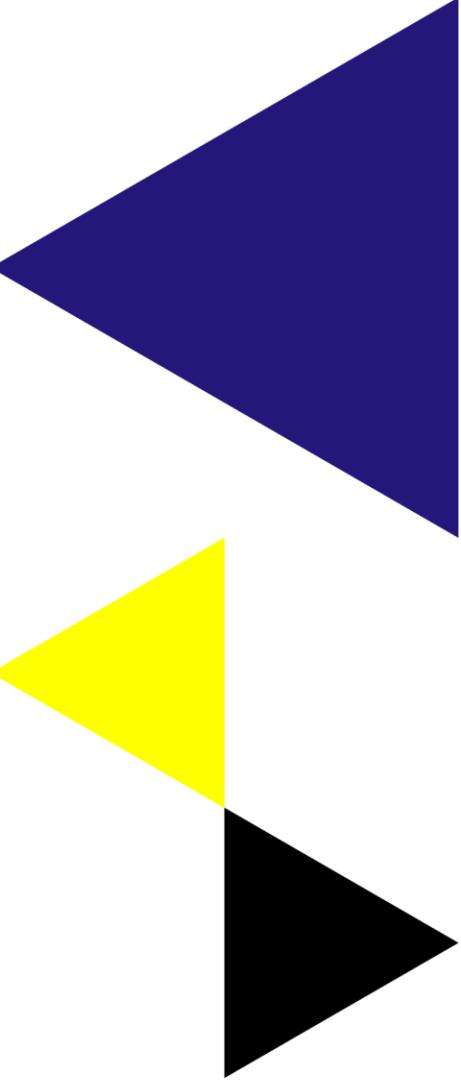
2D => 3D



Creating Tomorrow



Diffusion models nu ook voor video



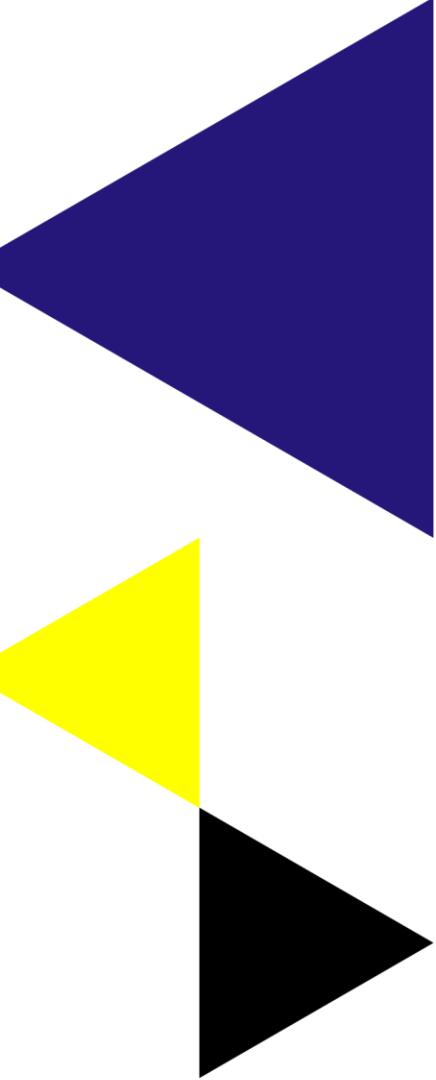
Creating Tomorrow

Text2video

Deze algoritmes zijn heel zwaar, kan alleen in de cloud. En dan alleen korte clips.

Enkele start-ups die dit doen:

- Klingai.com
- Pika.art
- Luna Labs Dream Machine
- Runway ML



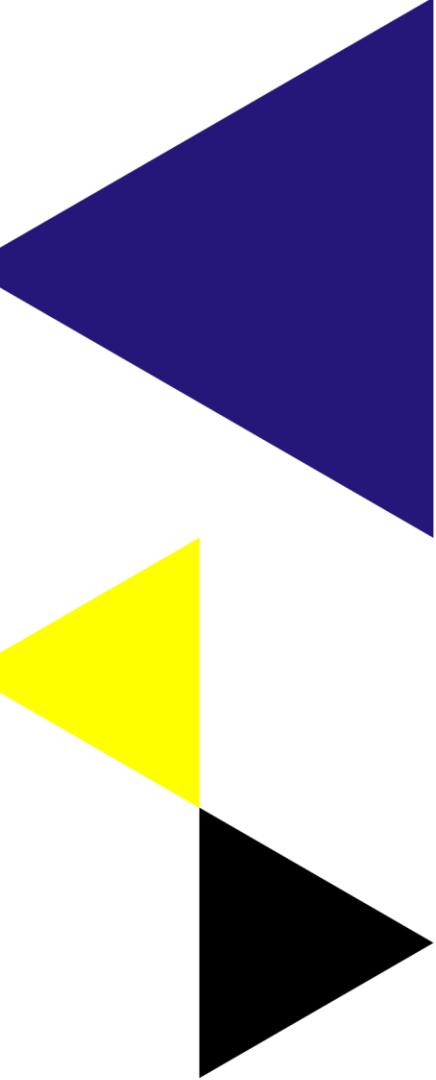
Diffusers Library Huggingface



Huggingface 😊 heeft een library gepubliceerd ‘Diffusers’

Maar let op! Deze modellen zijn vaak erg groot:

- Flux1.Schnell = ± 30 Gb
- Stable Diffusion Turbo = ± 14 Gb
- Daarom gebruiken we in de klas “google/ddpm-cat-256”, wat maar 450 Mb groot is. Maar de kwaliteit is dan ook matig, zoals je begrijpt.



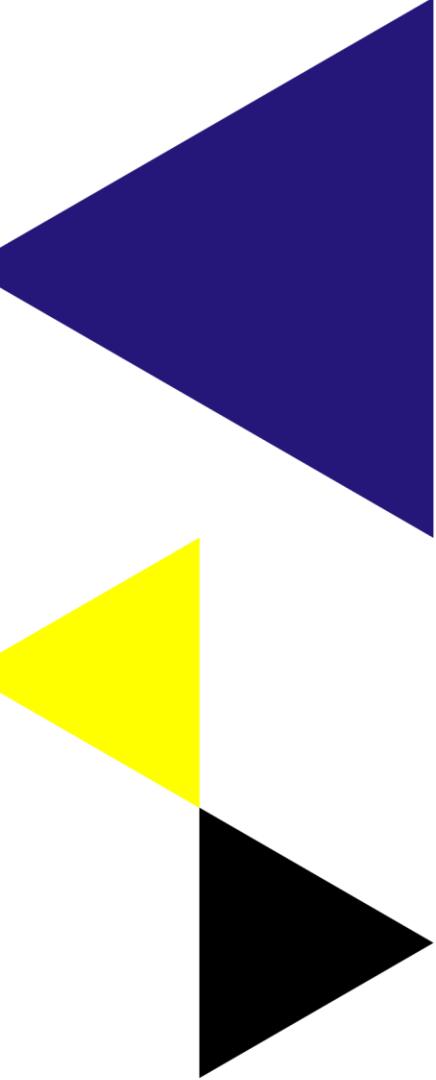
Aan de slag

- Maak het notebook op DLO:

Huggingface_diffusers.ipynb

- Extra stof:

- https://huggingface.co/docs/diffusers/using-diffusers/write_own_pipeline

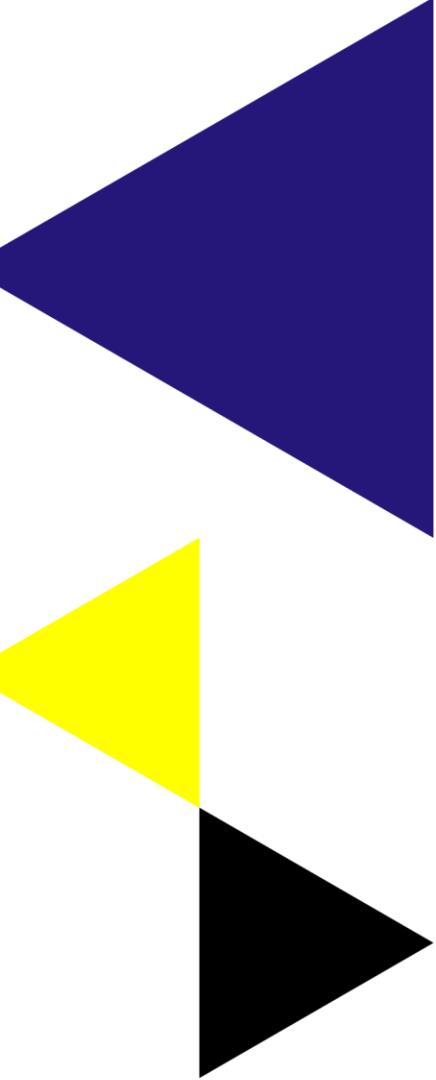


Meer lezen

- <https://www.assemblyai.com/blog/how-dall-e-2-actually-works/>
- <https://jalammar.github.io/illustrated-stable-diffusion/>
- <https://huggingface.co/docs/diffusers> Diffusers course van Huggingface 😊

Paper:

- Denoising Diffusion Probabilistic Models, <https://arxiv.org/pdf/2006.11239.pdf>, juni 2020



Einde

Creating Tomorrow