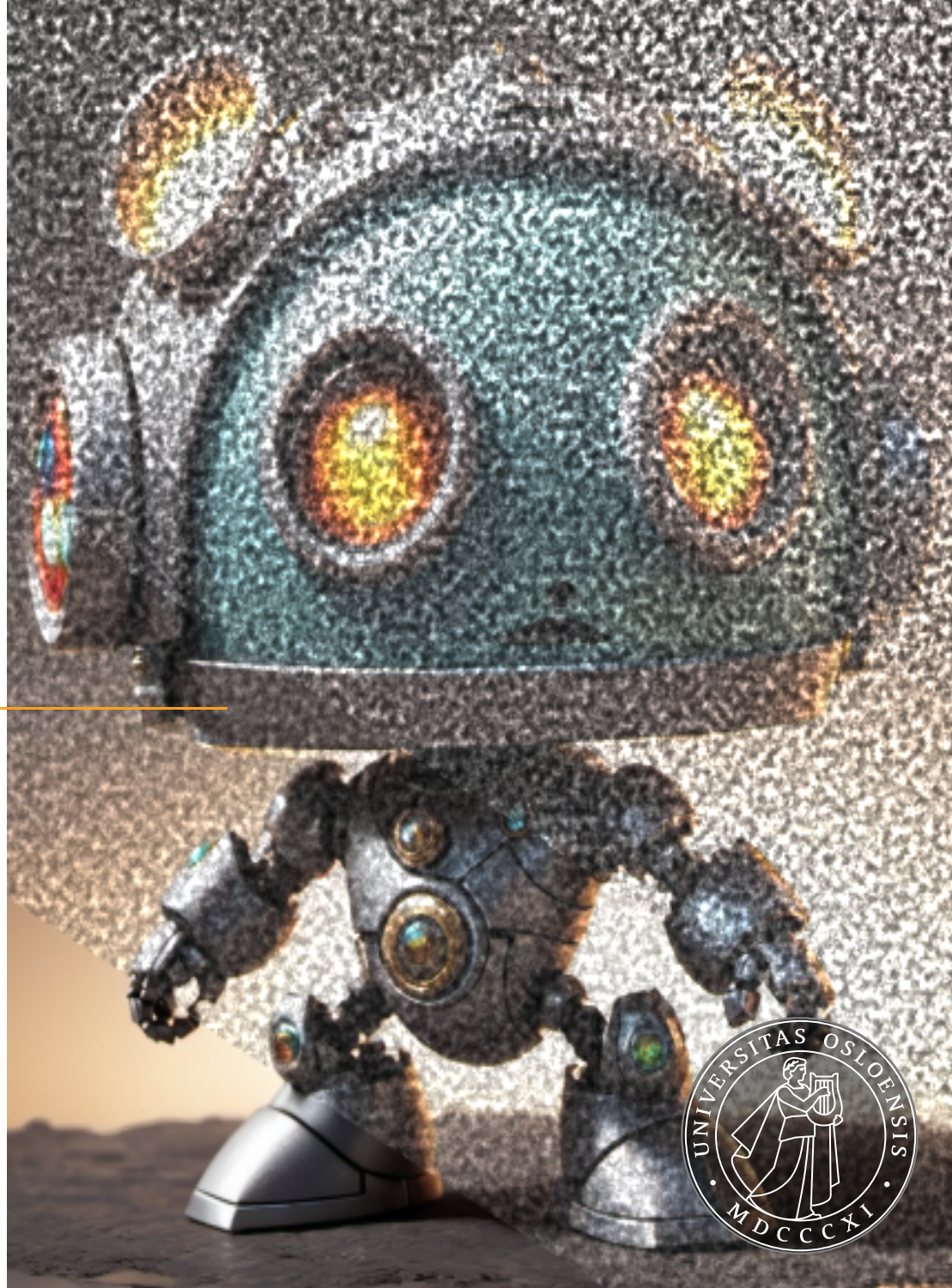
UNIVERSITY
OF OSLO

Generative Models

A birds-eye view

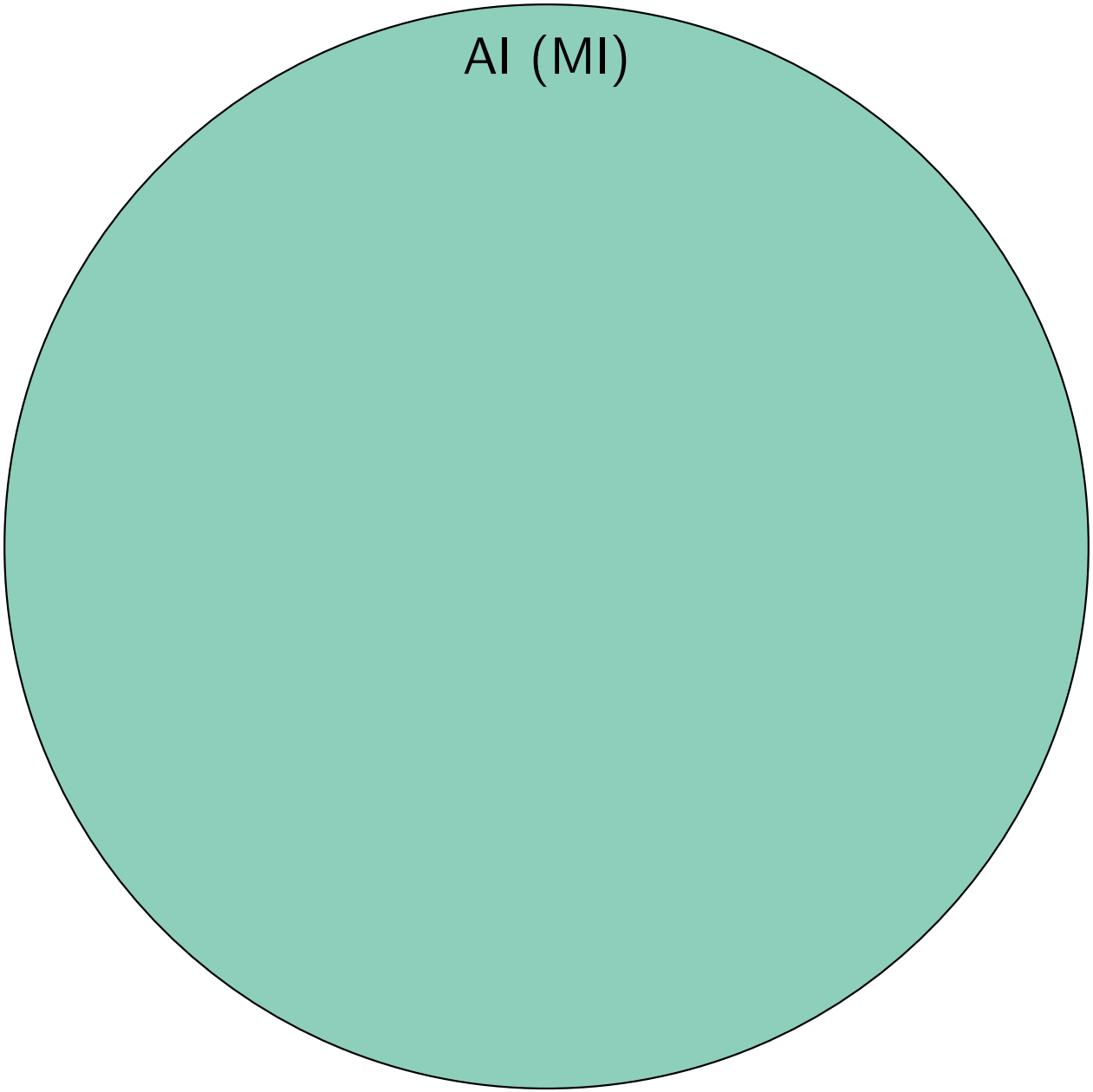
Adín Ramírez Rivera

adinr@uio.no
adin@gitlab.io
@adn_twitts
@adnrv.bsky.social

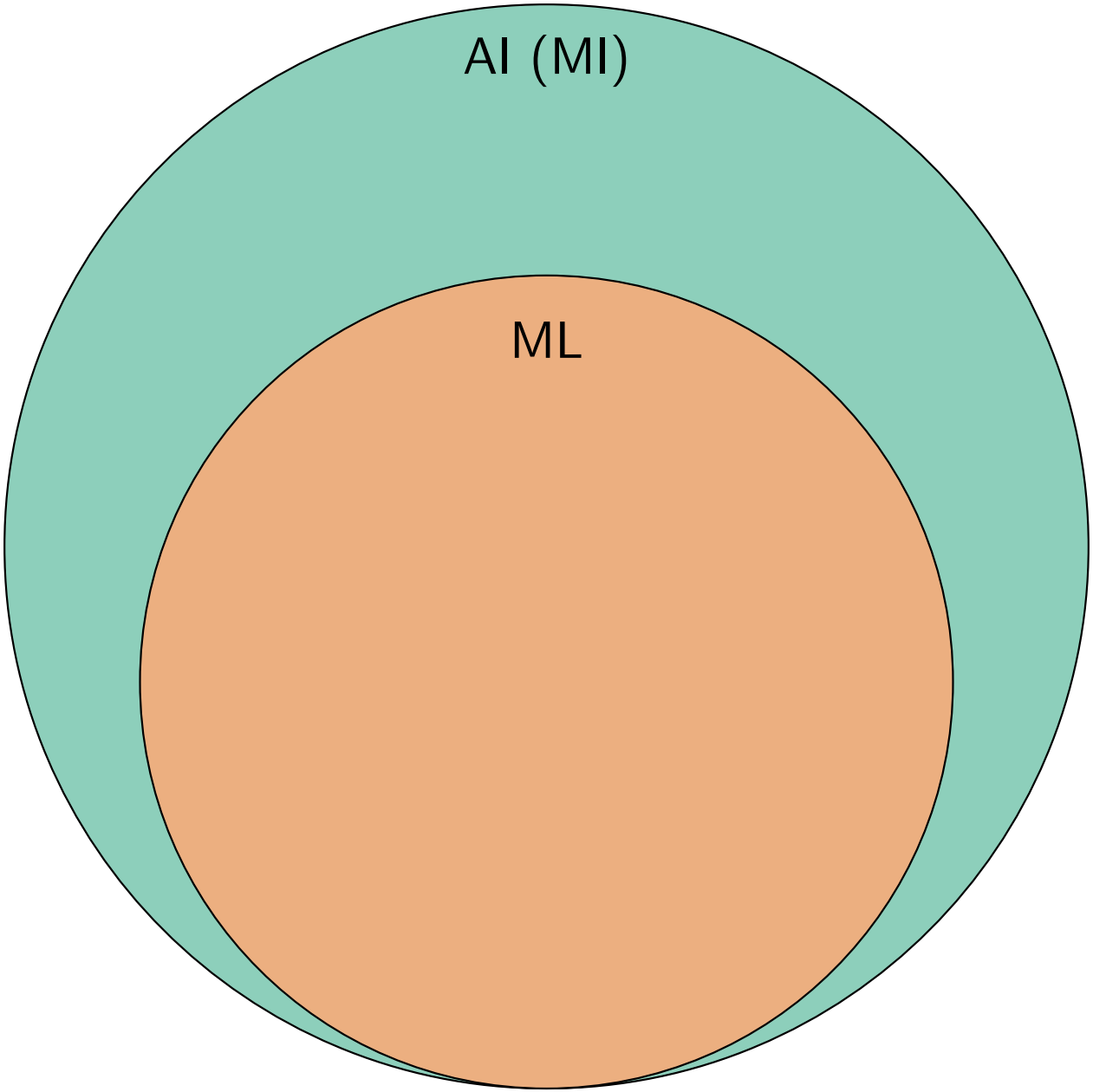


Gen-AI

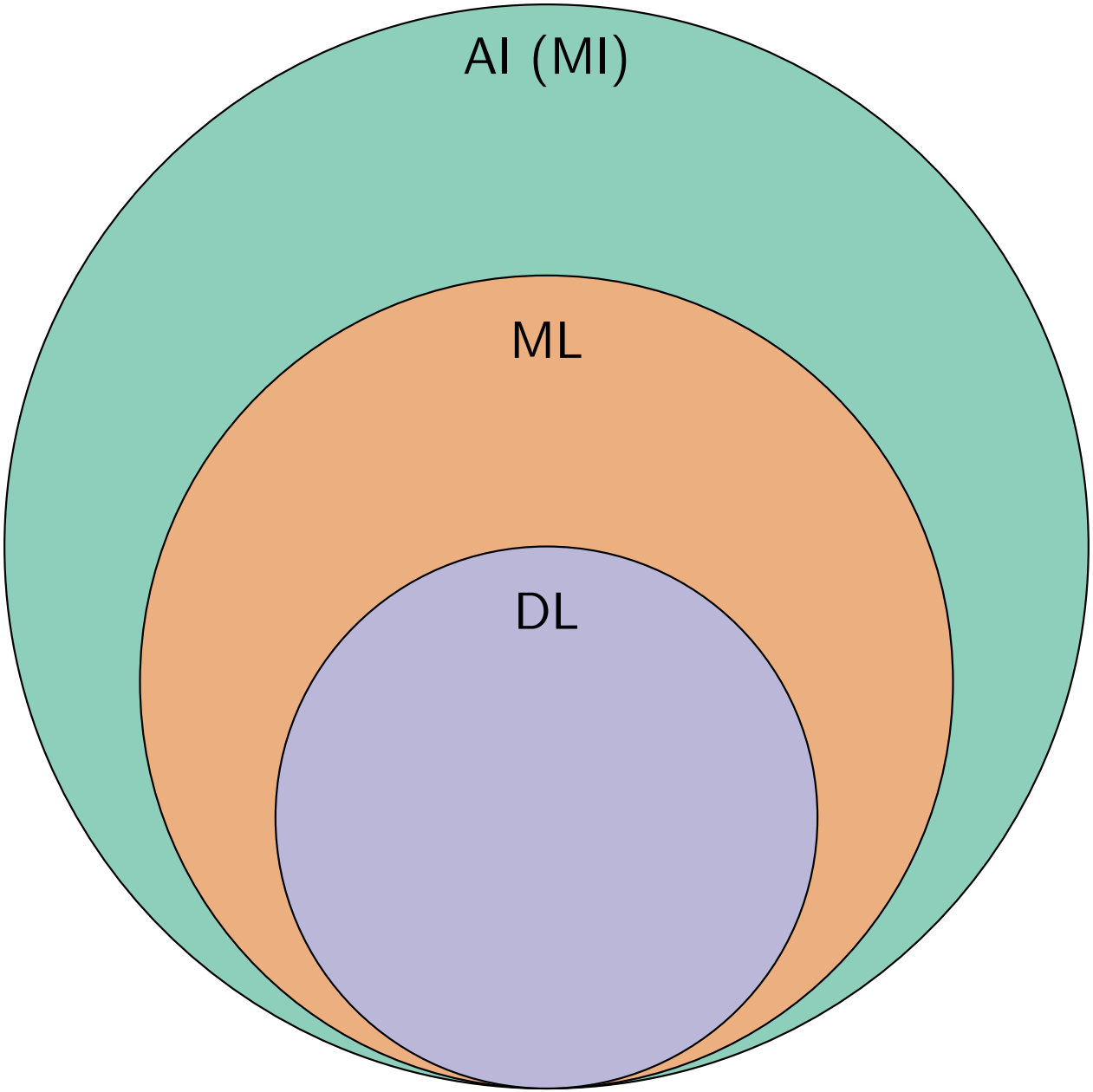
Gen-AI?



Gen-AI?

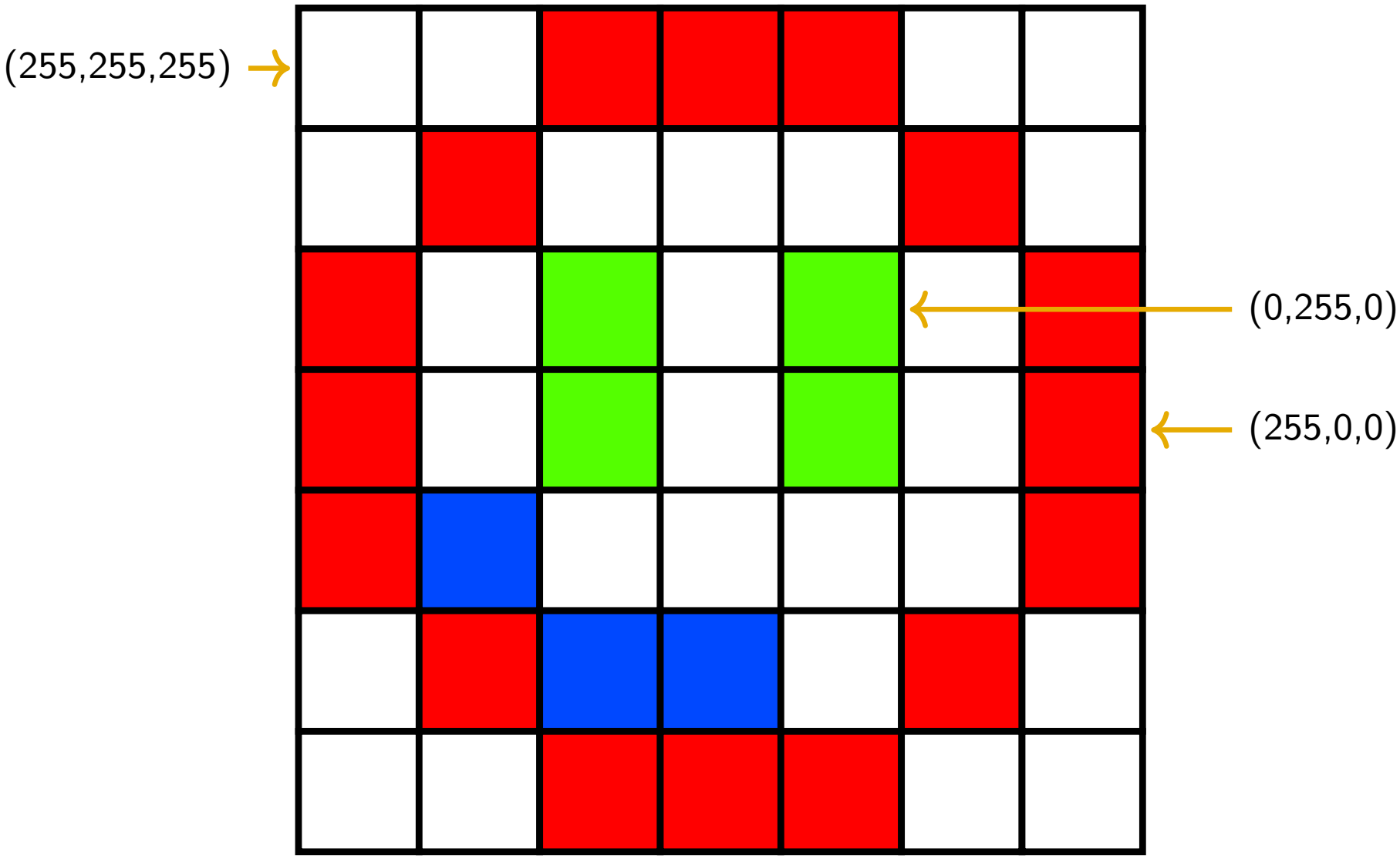


Gen-AI?

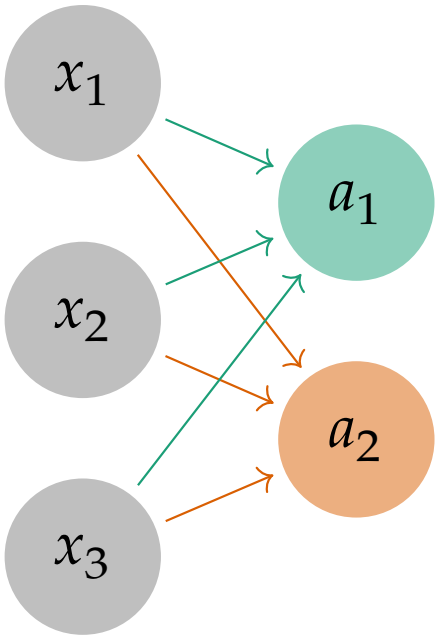


Some concepts before we start

Images are numbers



Neural Networks are numbers

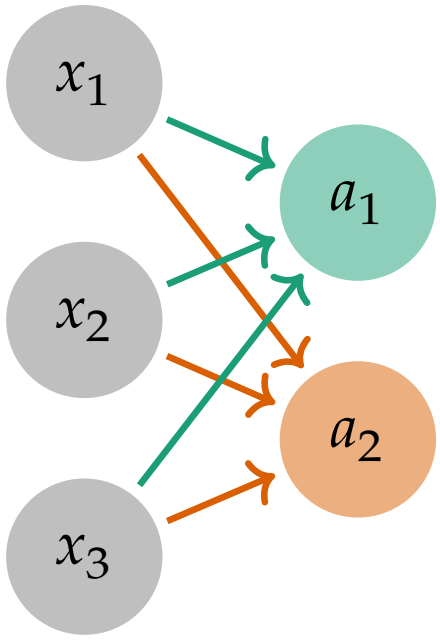


$$\begin{bmatrix} w_1 & w_2 & w_3 \\ w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b \\ b \end{bmatrix} =$$

$$\begin{bmatrix} w_1 x_1 + w_2 x_2 + w_3 x_3 + b \\ w_1 x_1 + w_2 x_2 + w_3 x_3 + b \end{bmatrix} =$$

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Neural Networks are numbers



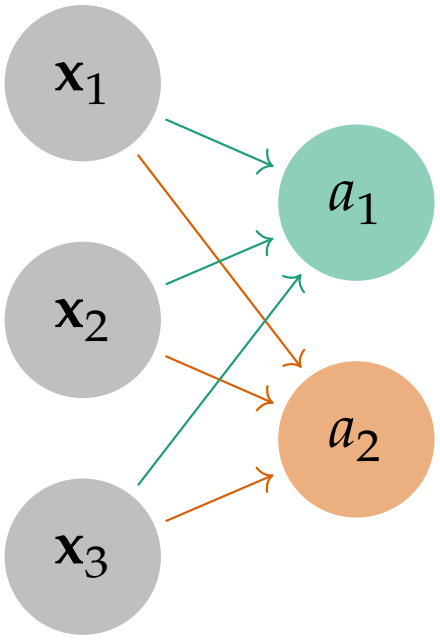
Model

$$\begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \\ \mathbf{w}_1 & \mathbf{w}_2 & \mathbf{w}_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \mathbf{b} \\ \mathbf{b} \end{bmatrix} =$$

$$\begin{bmatrix} \mathbf{w}_1 x_1 + \mathbf{w}_2 x_2 + \mathbf{w}_3 x_3 + \mathbf{b} \\ \mathbf{w}_1 x_1 + \mathbf{w}_2 x_2 + \mathbf{w}_3 x_3 + \mathbf{b} \end{bmatrix} =$$

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Neural Networks are numbers



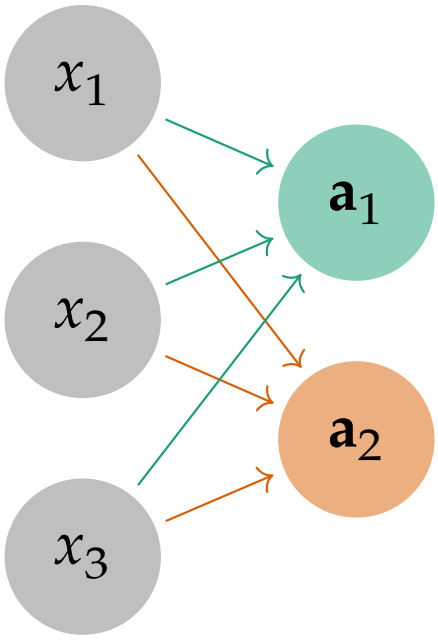
Inputs

$$\begin{bmatrix} w_1 & w_2 & w_3 \\ w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} + \begin{bmatrix} b \\ b \end{bmatrix} =$$

$$\begin{bmatrix} w_1\mathbf{x}_1 + w_2\mathbf{x}_2 + w_3\mathbf{x}_3 + b \\ w_1\mathbf{x}_1 + w_2\mathbf{x}_2 + w_3\mathbf{x}_3 + b \end{bmatrix} =$$

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Neural Networks are numbers



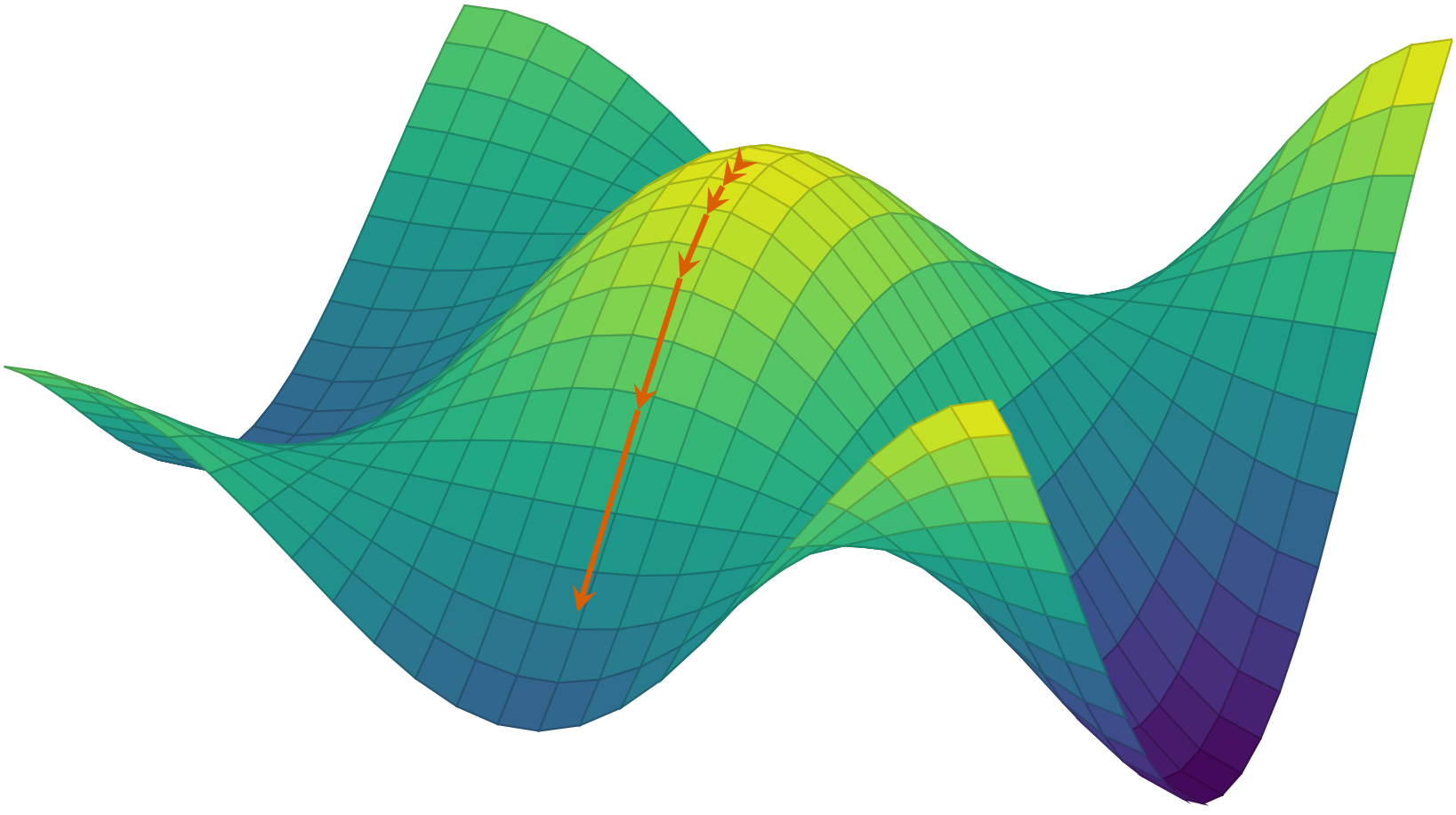
Outputs

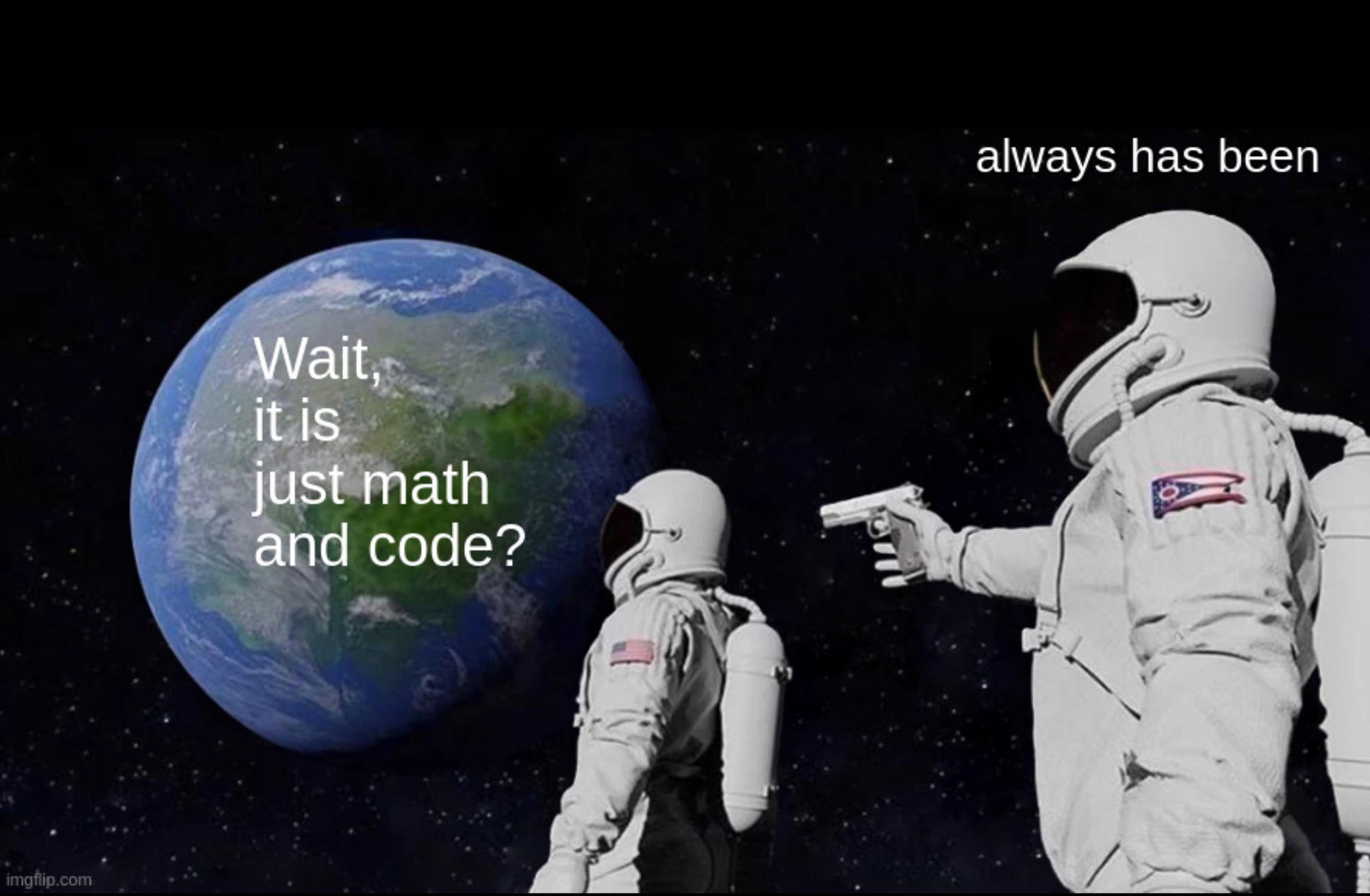
$$\begin{bmatrix} w_1 & w_2 & w_3 \\ w_1 & w_2 & w_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} b \\ b \end{bmatrix} =$$

$$\begin{bmatrix} w_1 x_1 + w_2 x_2 + w_3 x_3 + b \\ w_1 x_1 + w_2 x_2 + w_3 x_3 + b \end{bmatrix} =$$

$$\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix}$$

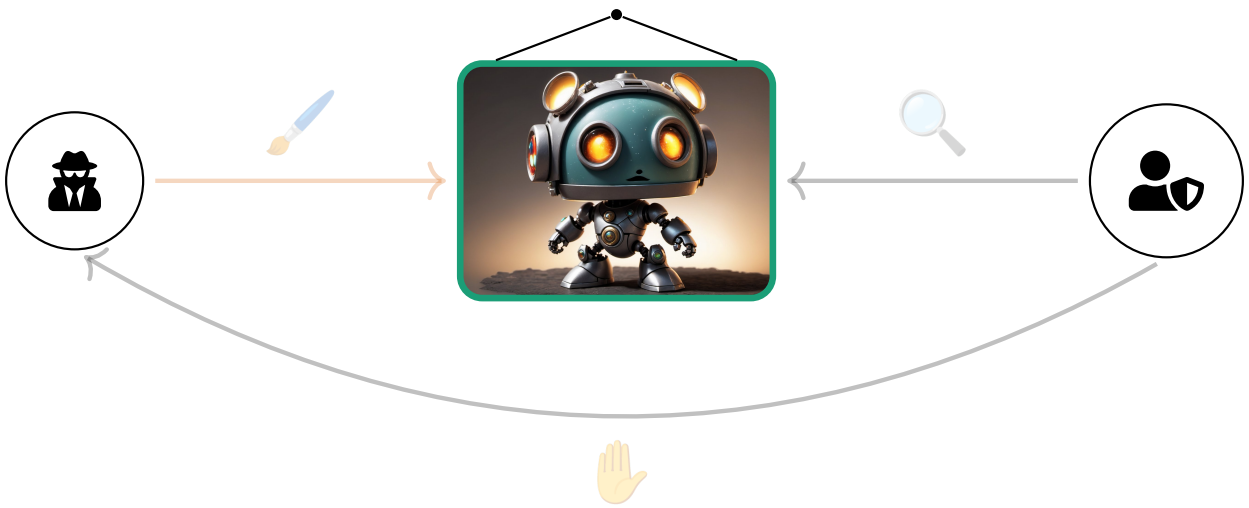
Optimization: finding our way through the landscape




always has been
Wait,
it is
just math
and code?
imgflip.com

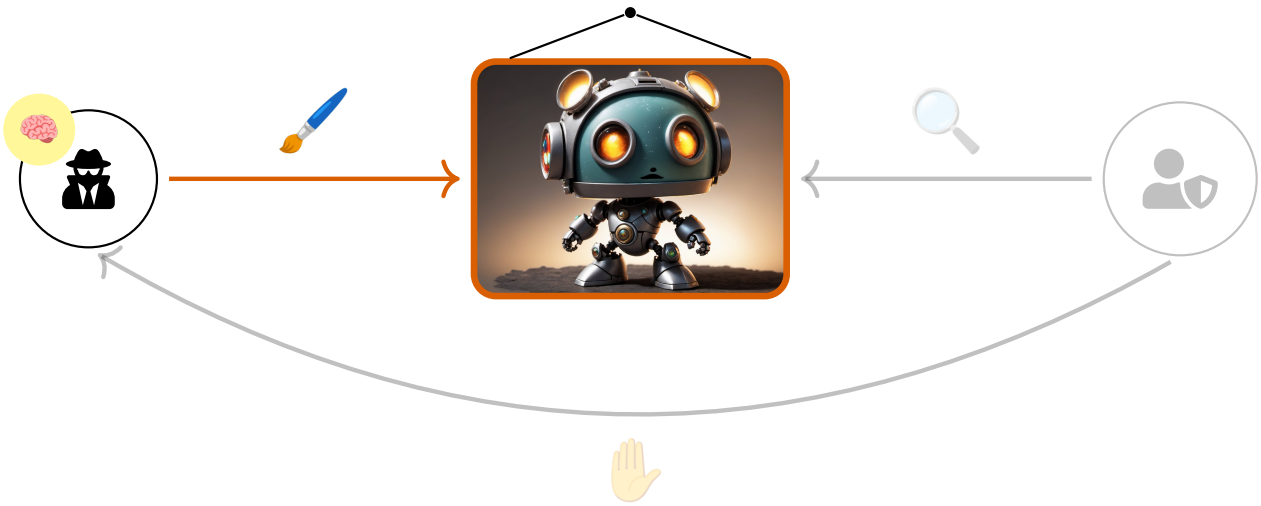
Generative Adversarial Networks (GANs)

A forger and a police: play a game



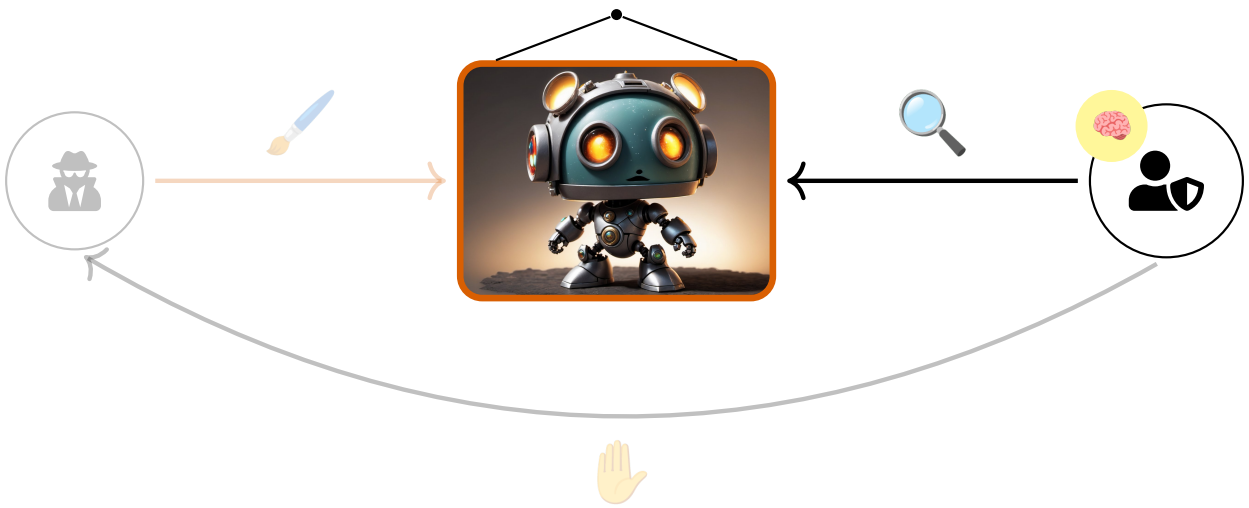
Goodfellow et al. “Generative Adversarial Networks.” NeurIPS, 2014.

A forger and a police: play a game



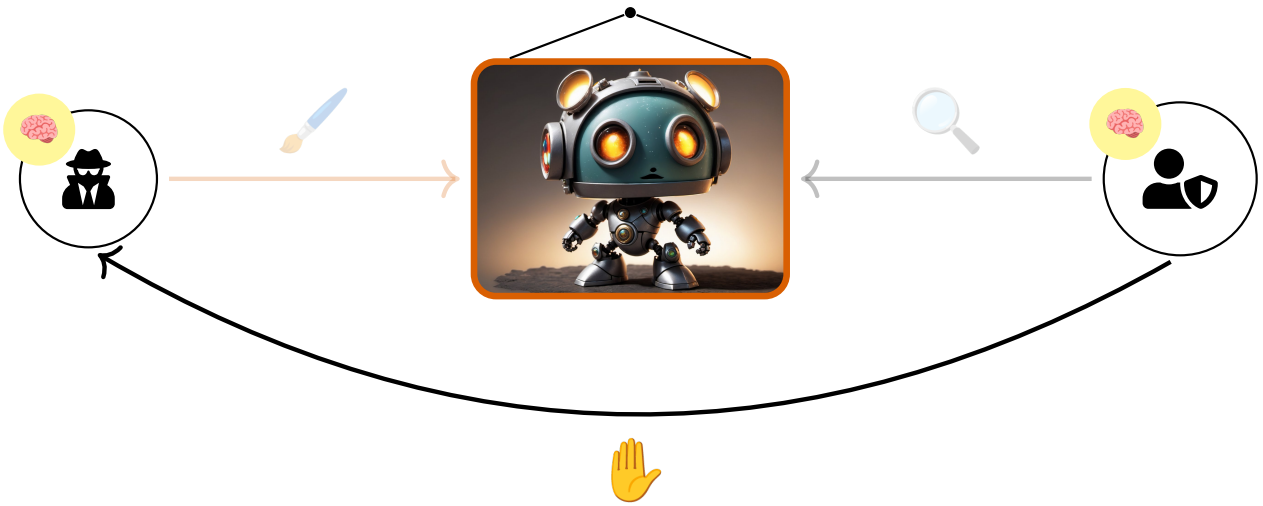
Goodfellow et al. “Generative Adversarial Networks.” NeurIPS, 2014.

A forger and a police: play a game



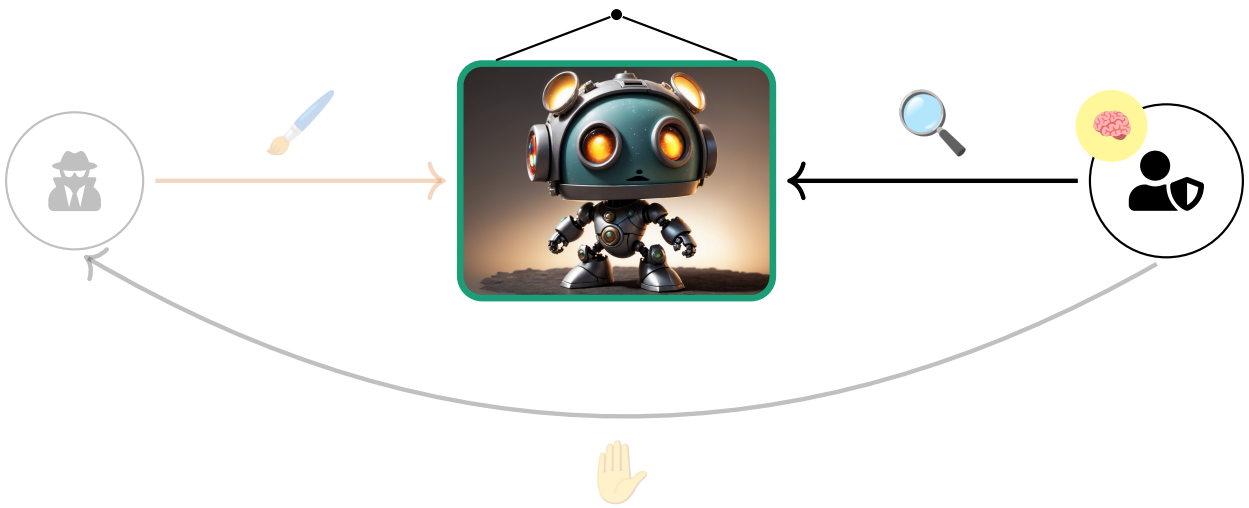
Goodfellow et al. “Generative Adversarial Networks.” NeurIPS, 2014.

A forger and a police: play a game



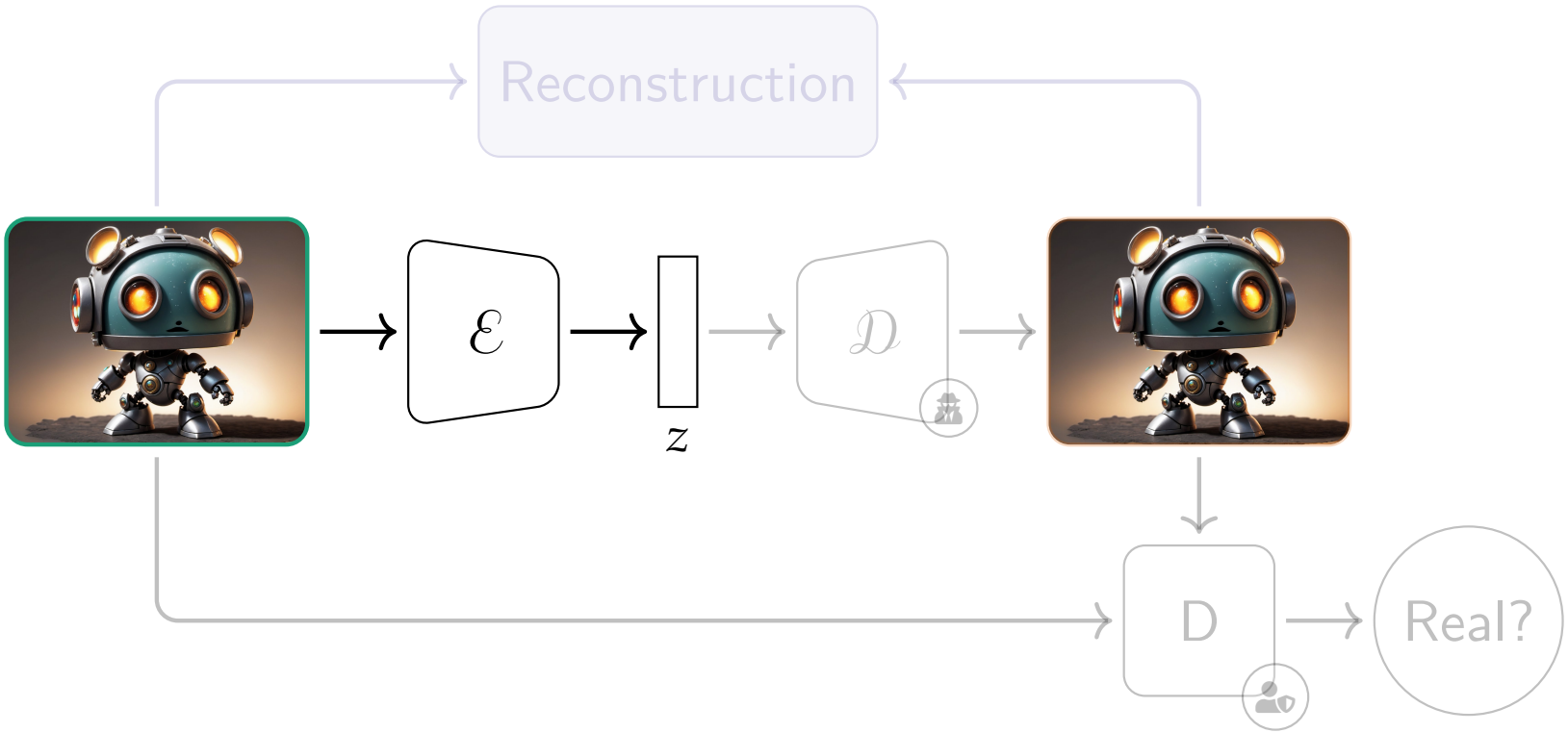
Goodfellow et al. “Generative Adversarial Networks.” NeurIPS, 2014.

A forger and a police: play a game

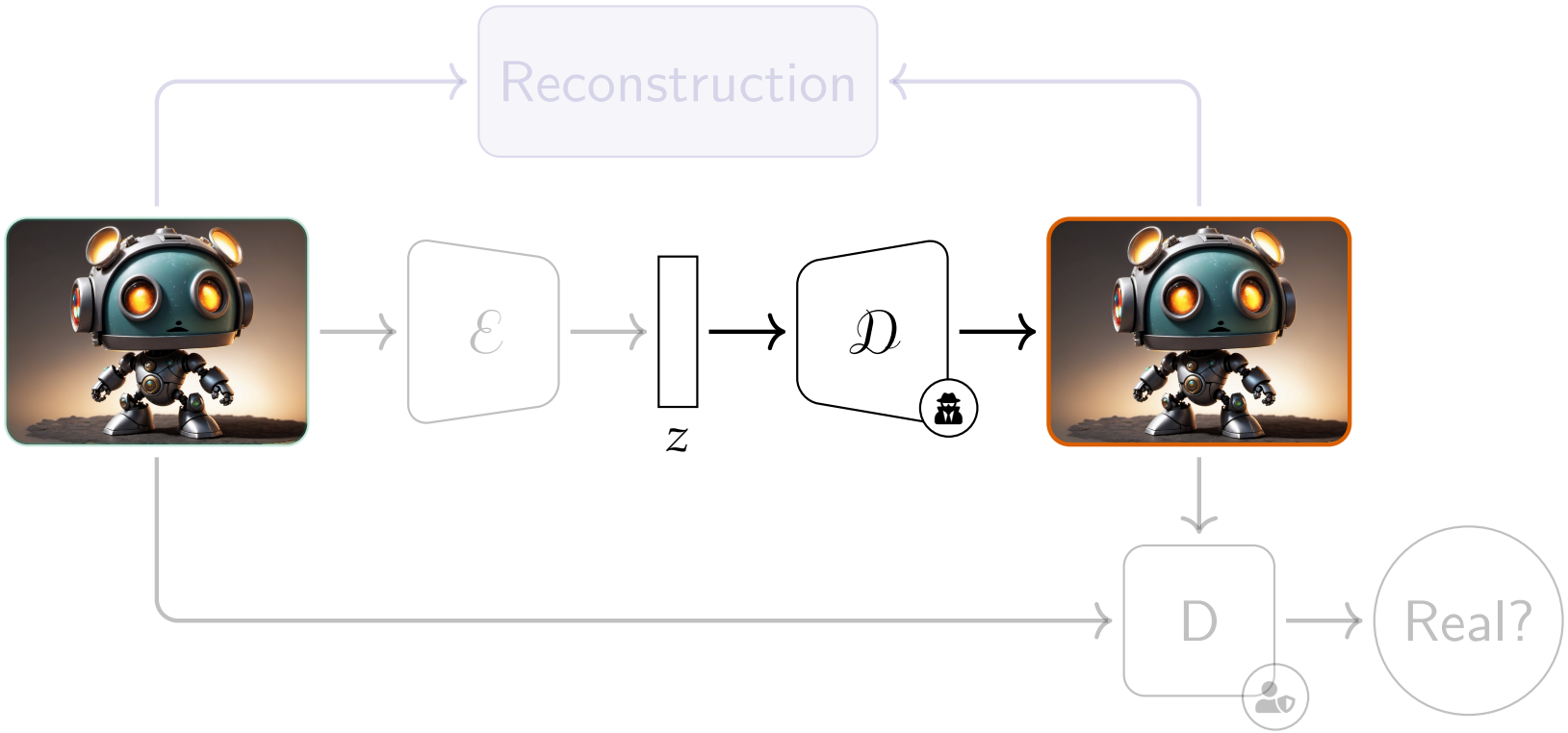


Goodfellow et al. “Generative Adversarial Networks.” NeurIPS, 2014.

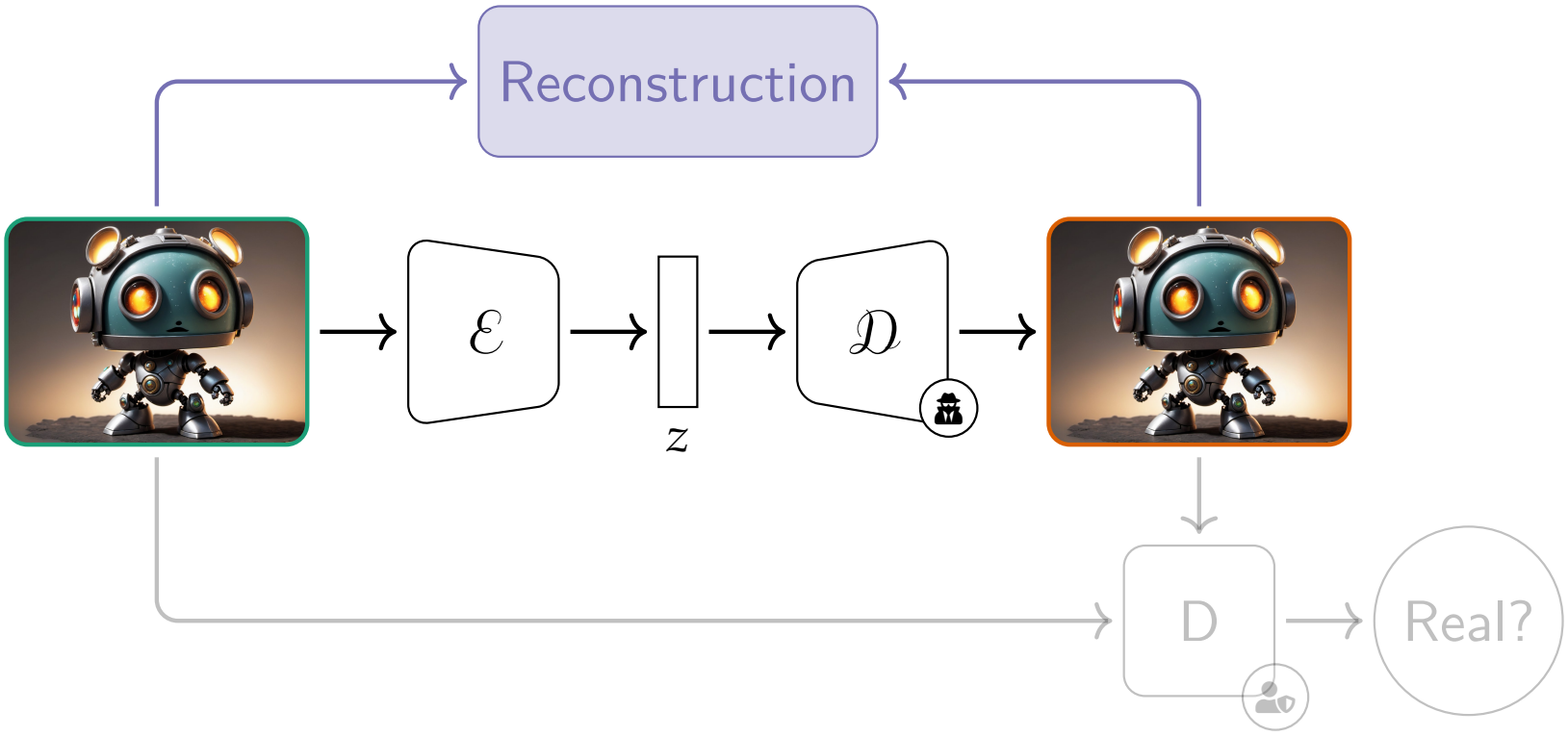
GAN



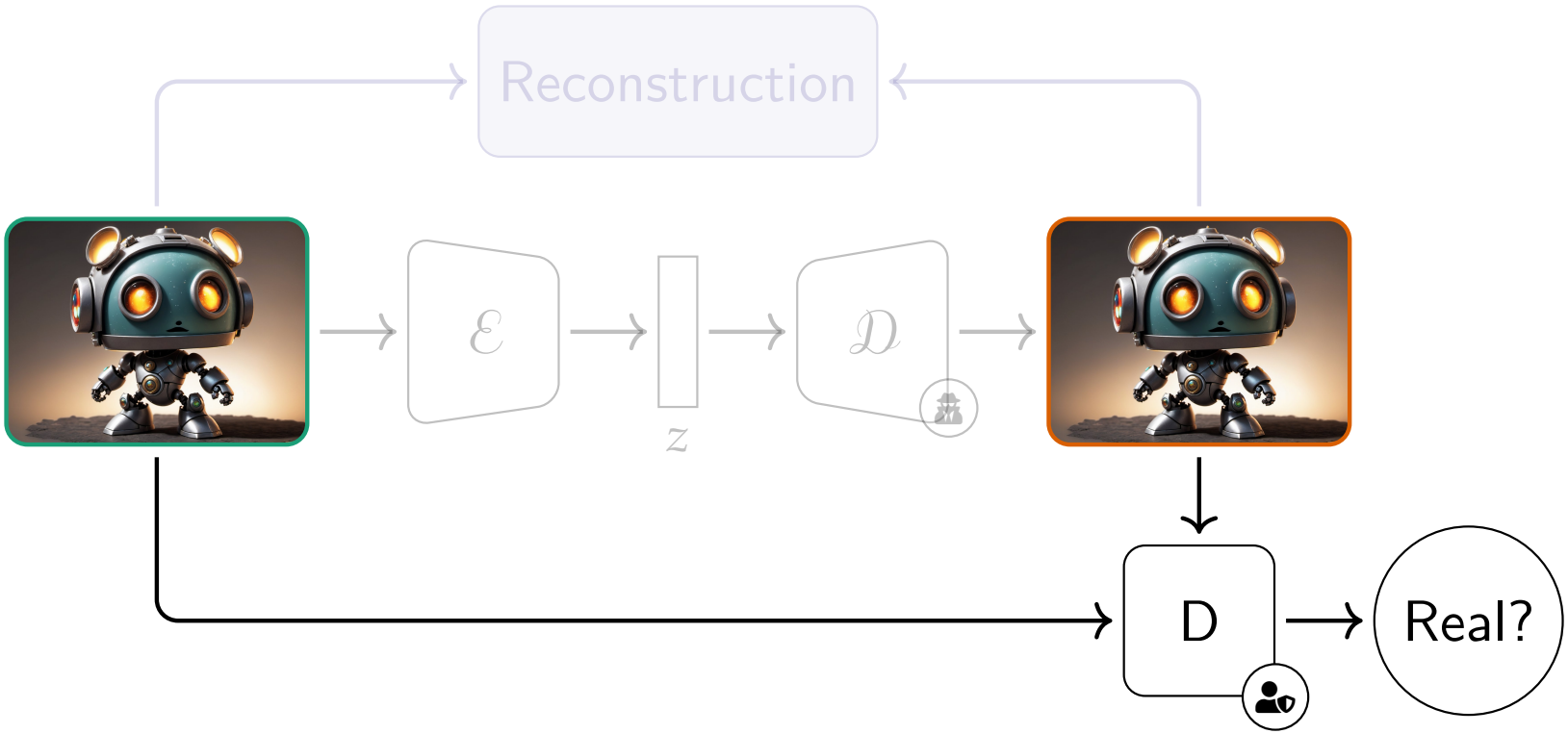
GAN



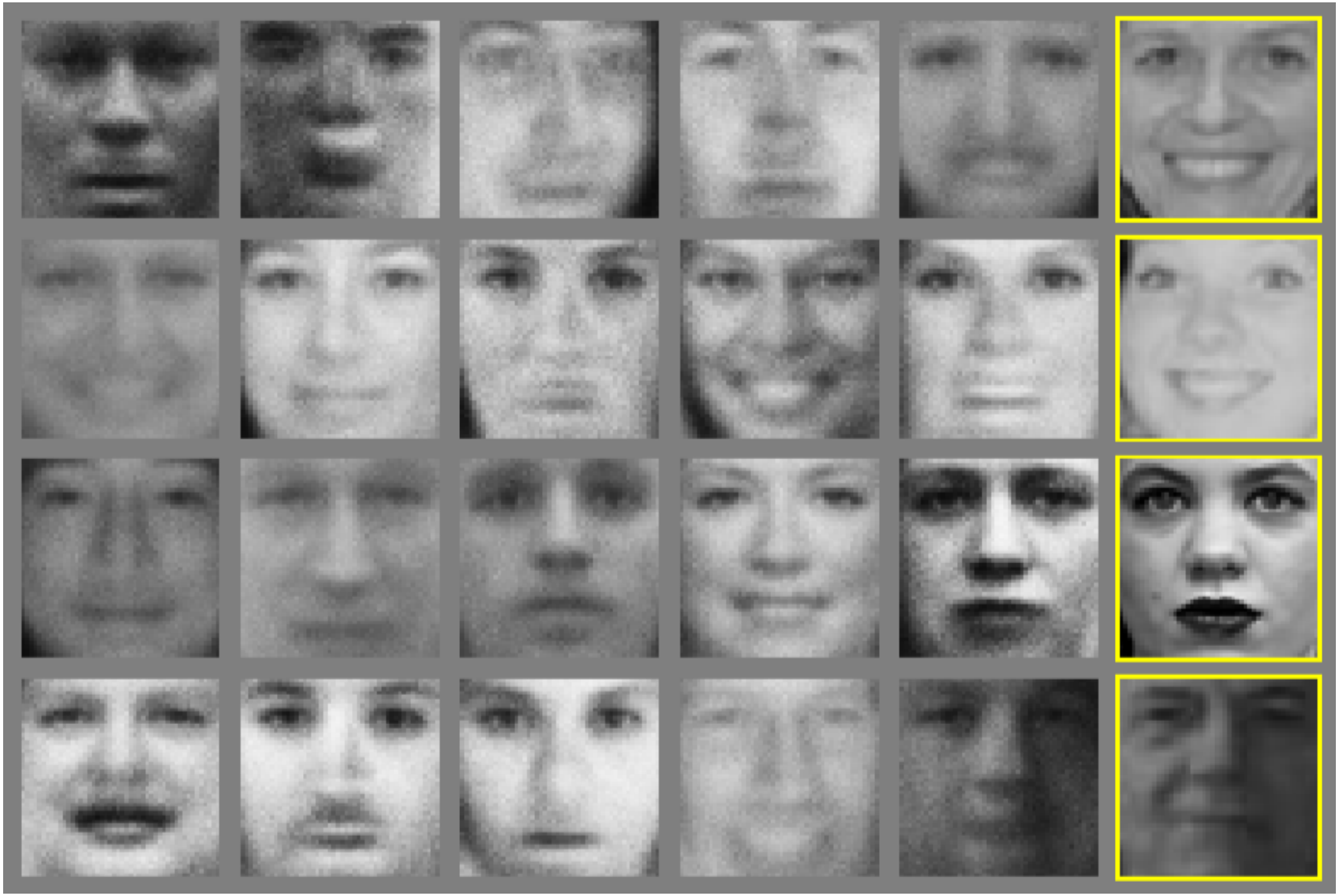
GAN



GAN



First results



Goodfellow et al. "Generative Adversarial Networks." NeurIPS, 2014.

“Older” results



Karras et al. “Progressive Growing of GANs for Improved Quality, Stability, and Variation.” ICLR, 2018.

"Old" results



Vahdat and Kautz. "NVAE: A Deep Hierarchical Variational Autoencoder." NeurIPS, 2020.

Current results



Huang et al. "The GAN is dead; long live the GAN! A Modern Baseline GAN." NeurIPS, 2024.

Diffusion Models



dif · fu · sion

[*Latin*] diffundere
“to pour out over a wide surface, spread out, extend”






Diffusion processes














Reverse diffusion processes

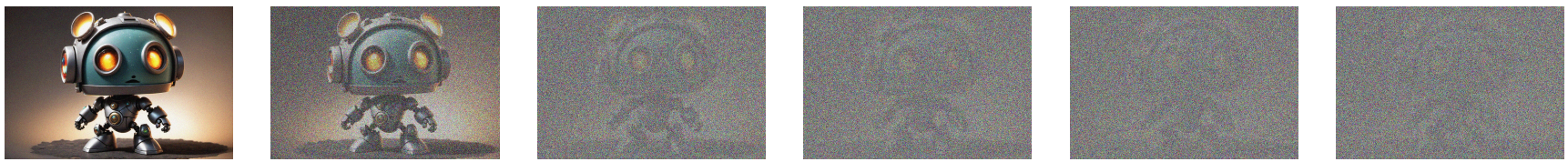



Easy




Hard

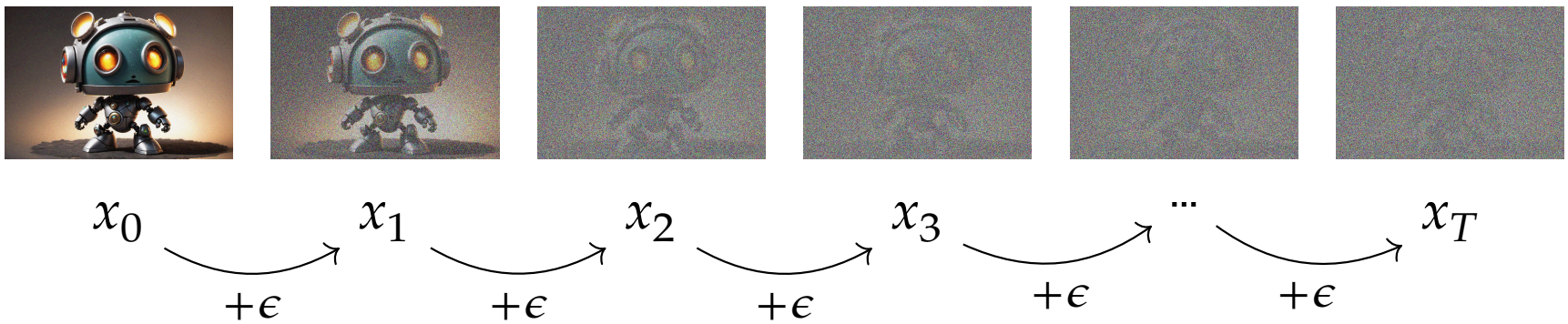
Diffusion process



x_0 x_1 x_2 x_3 ... x_T

Sohl-Dickstein et al., "Deep Unsupervised Learning using Nonequilibrium Thermodynamics." ICML, 2015.
Ho et al., "Denoising Diffusion Probabilistic Models." NeurIPS, 2020.

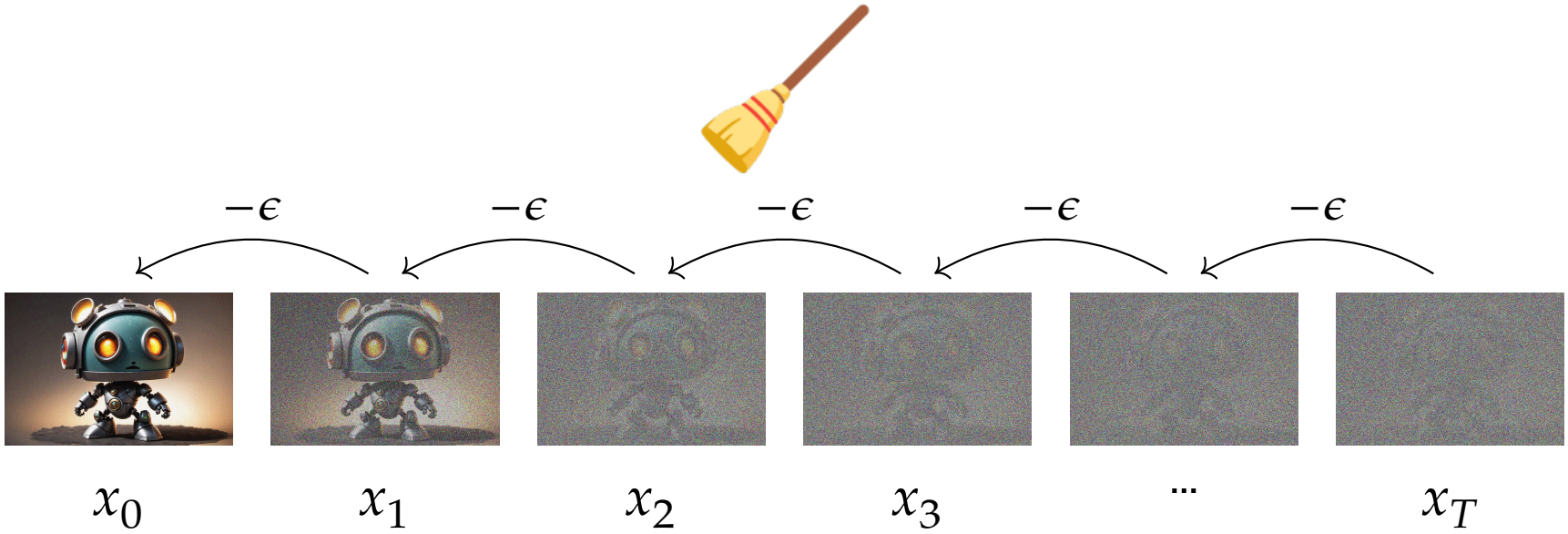
Diffusion process
⏩ Forward





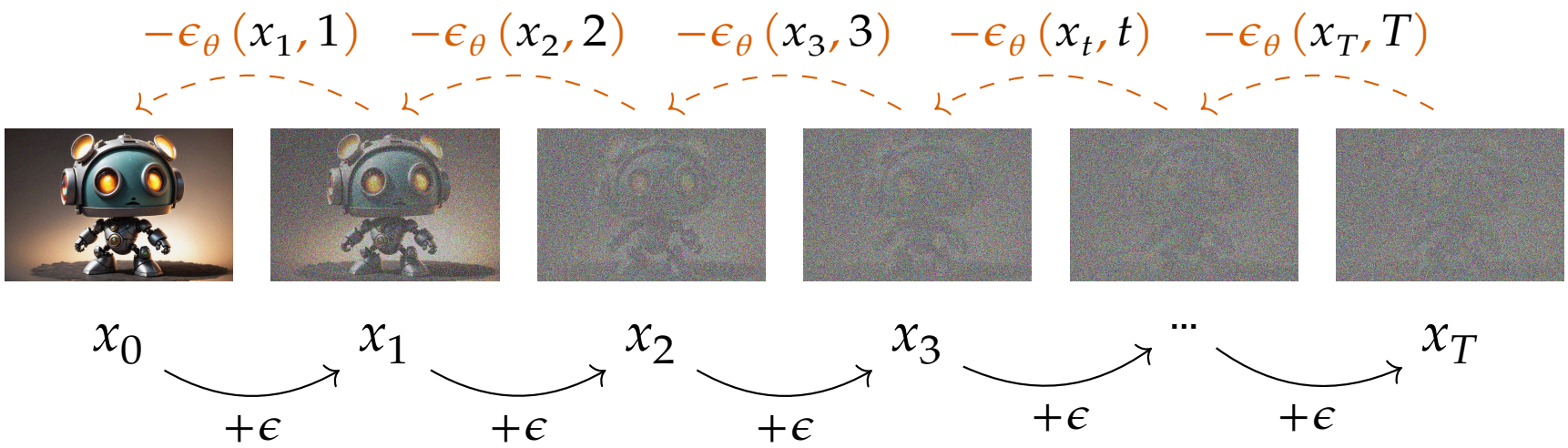
Sohl-Dickstein et al., "Deep Unsupervised Learning using Nonequilibrium Thermodynamics." ICML, 2015.
Ho et al., "Denoising Diffusion Probabilistic Models." NeurIPS, 2020.

Diffusion process
⏪ Reverse



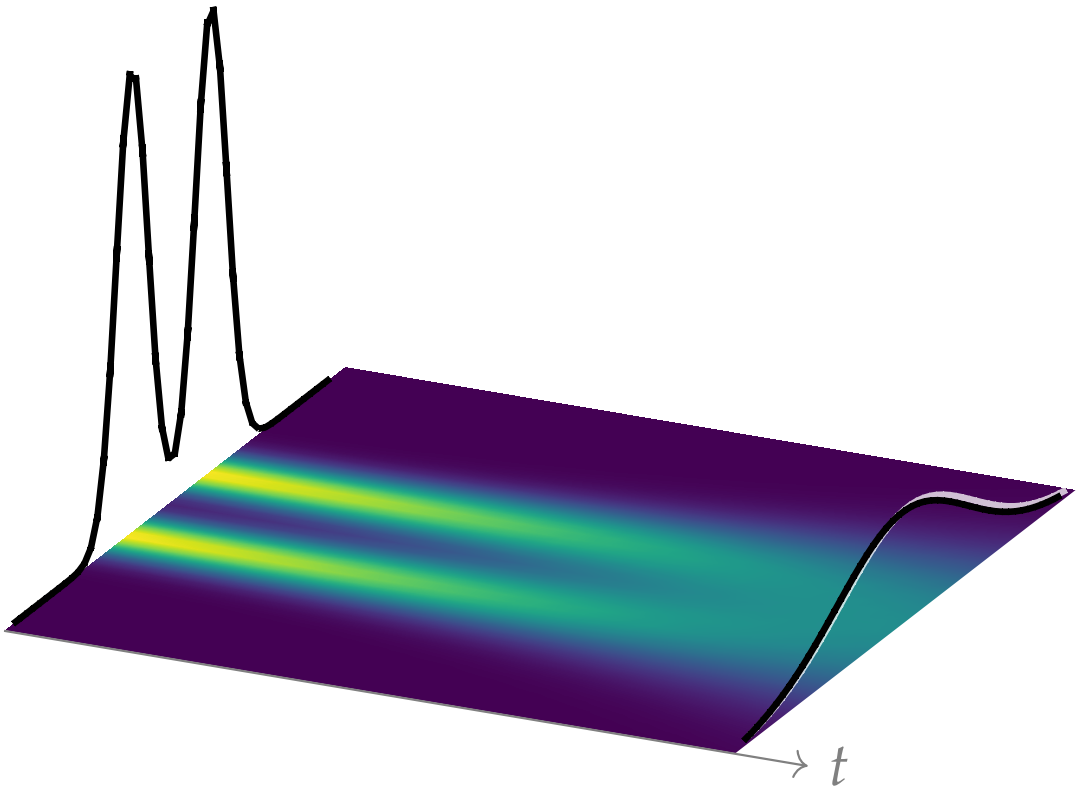
Sohl-Dickstein et al., "Deep Unsupervised Learning using Nonequilibrium Thermodynamics." ICML, 2015.
Ho et al., "Denoising Diffusion Probabilistic Models." NeurIPS, 2020.

Diffusion process



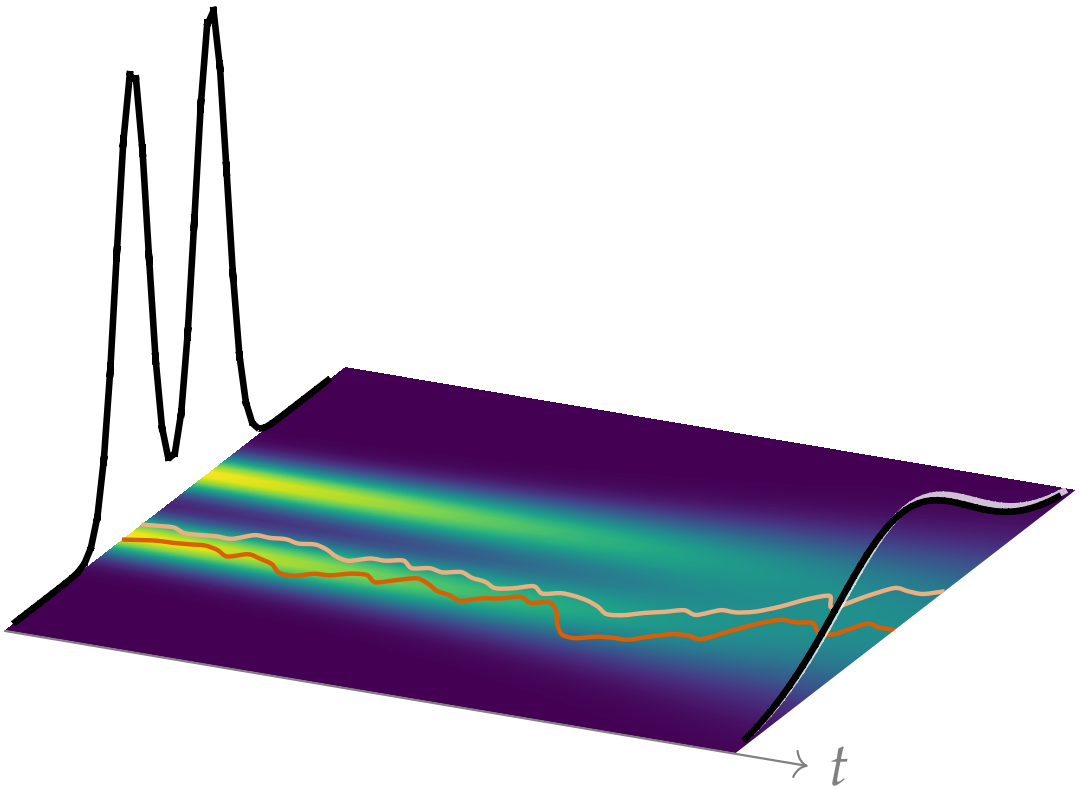
Sohl-Dickstein et al., "Deep Unsupervised Learning using Nonequilibrium Thermodynamics." ICML, 2015.
Ho et al., "Denoising Diffusion Probabilistic Models." NeurIPS, 2020.

Data diffusion with SDEs



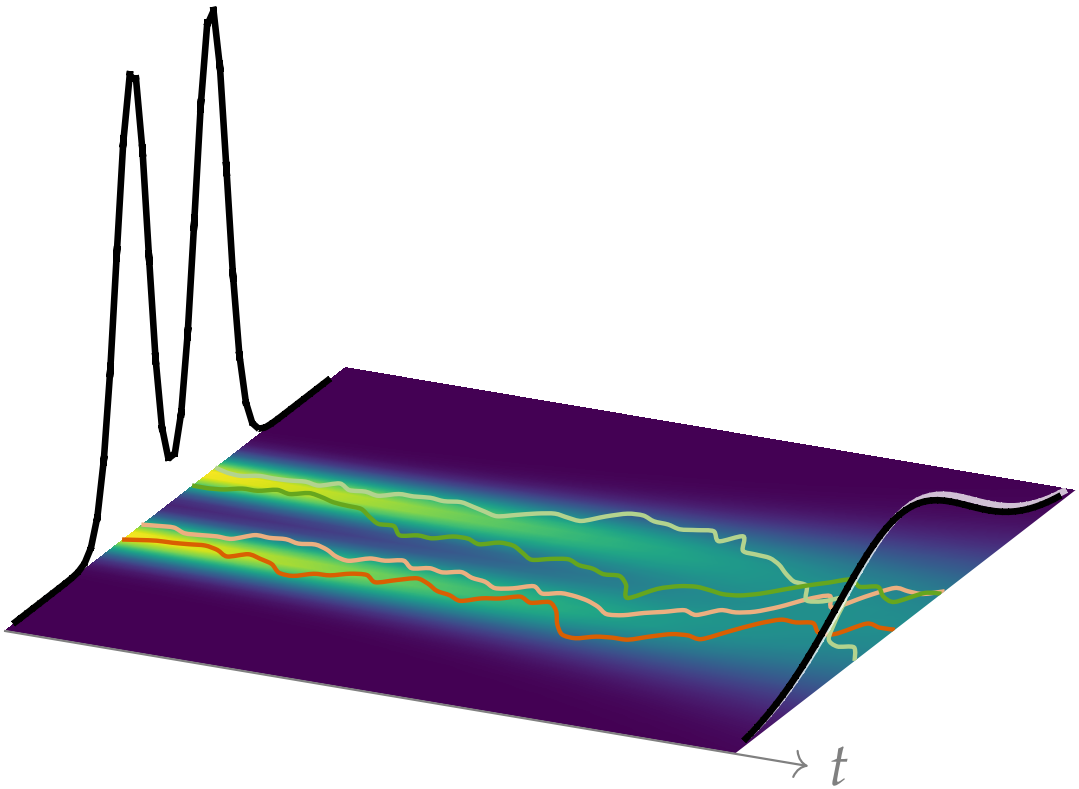
Distribution from $1\,000\,000$ diffused samples.

Data diffusion with SDEs



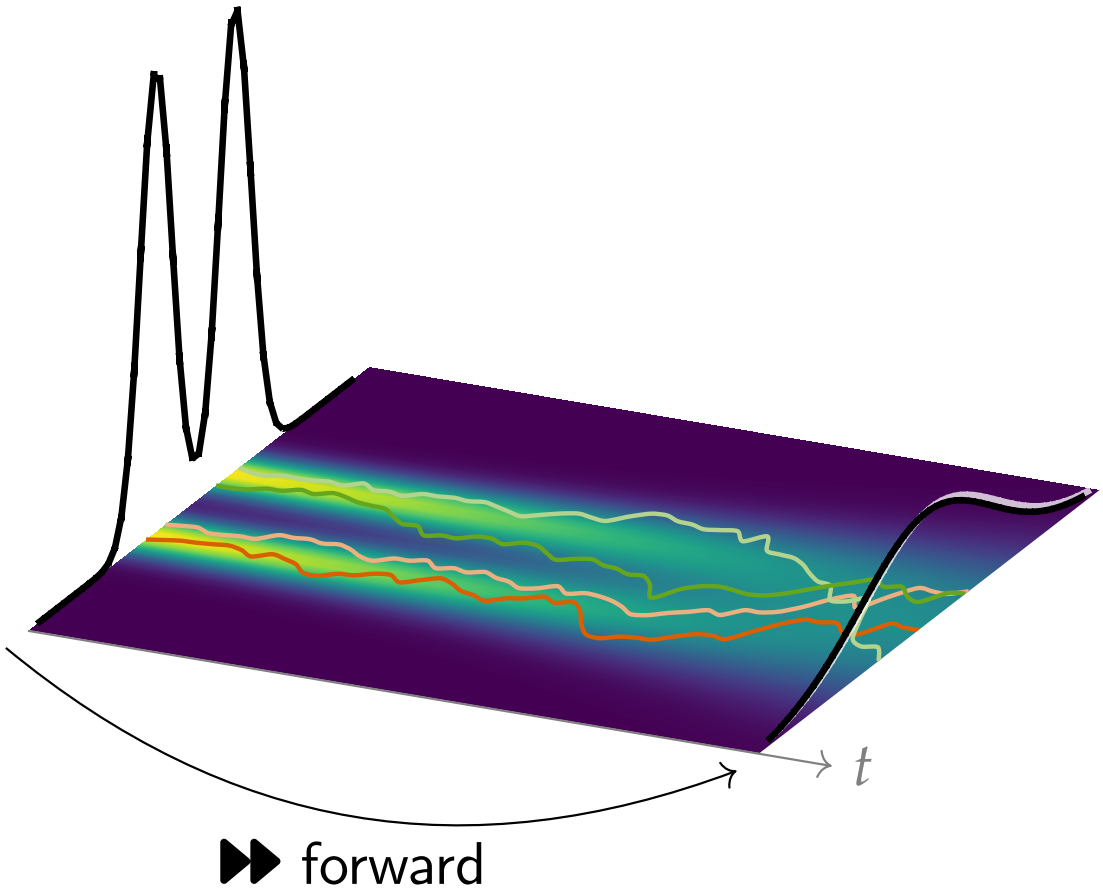
Distribution from $1\,000\,000$ diffused samples.

Data diffusion with SDEs



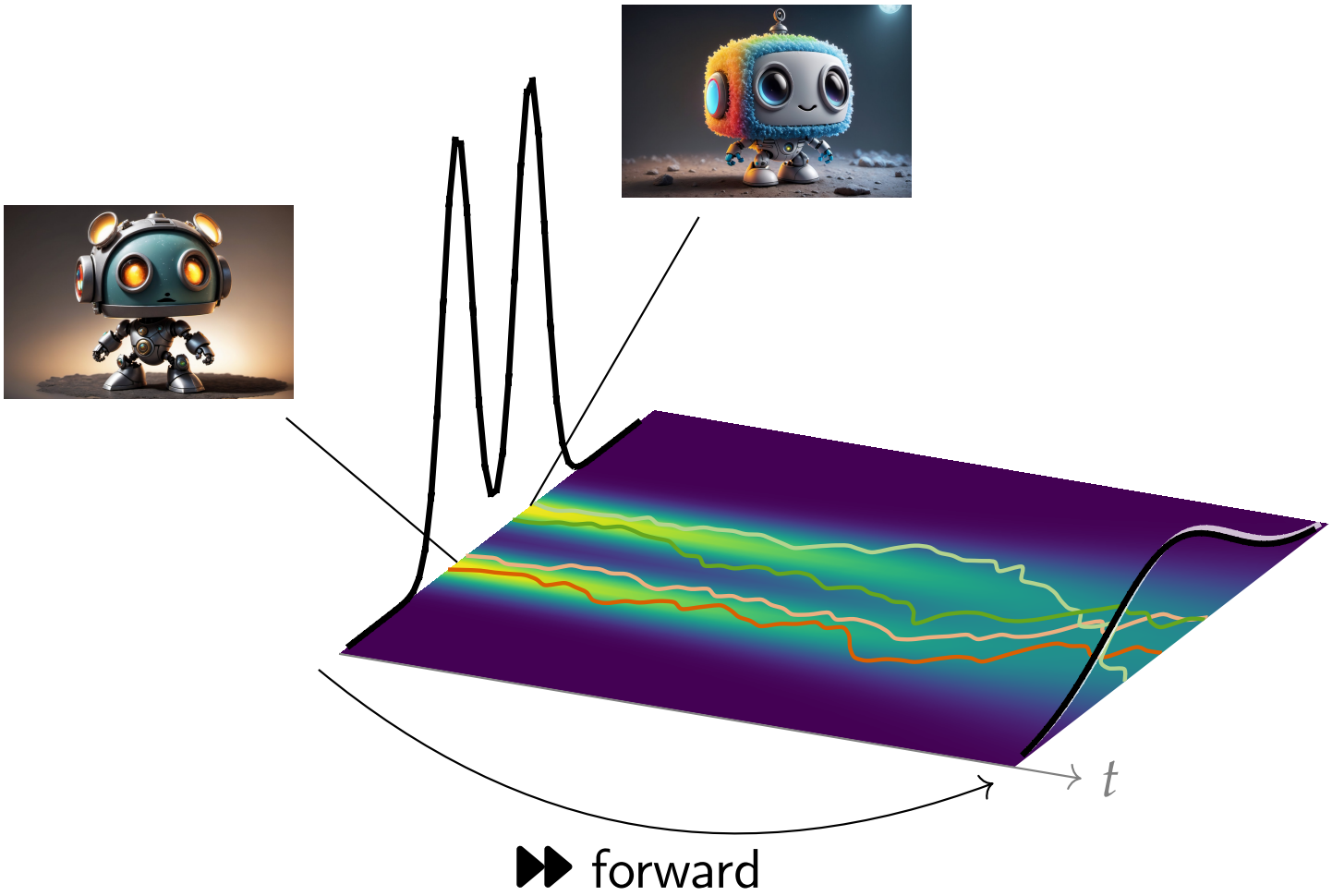
Distribution from $1\,000\,000$ diffused samples.

Data diffusion with SDEs



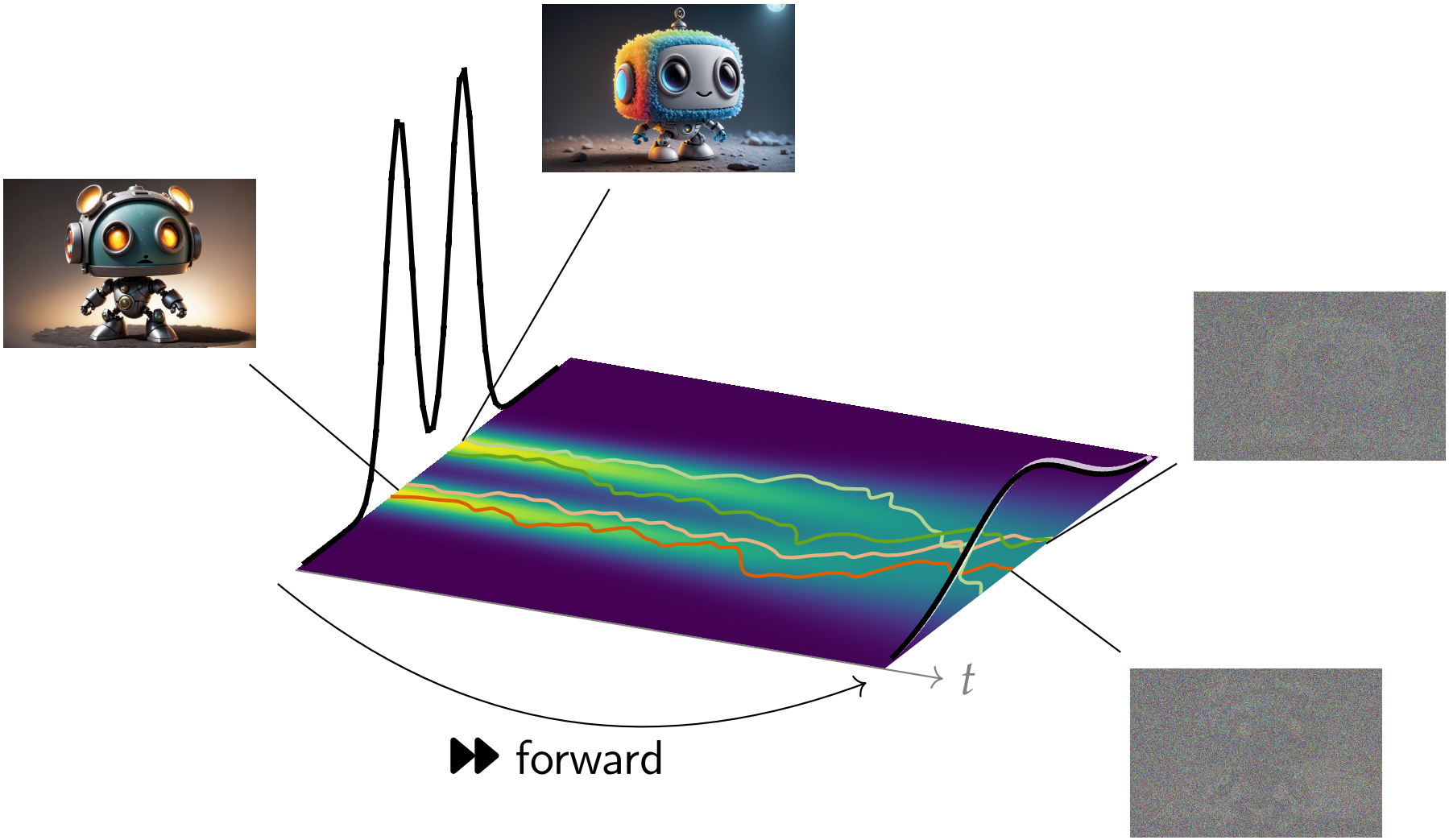
Distribution from $1\,000\,000$ diffused samples.

Data diffusion with SDEs



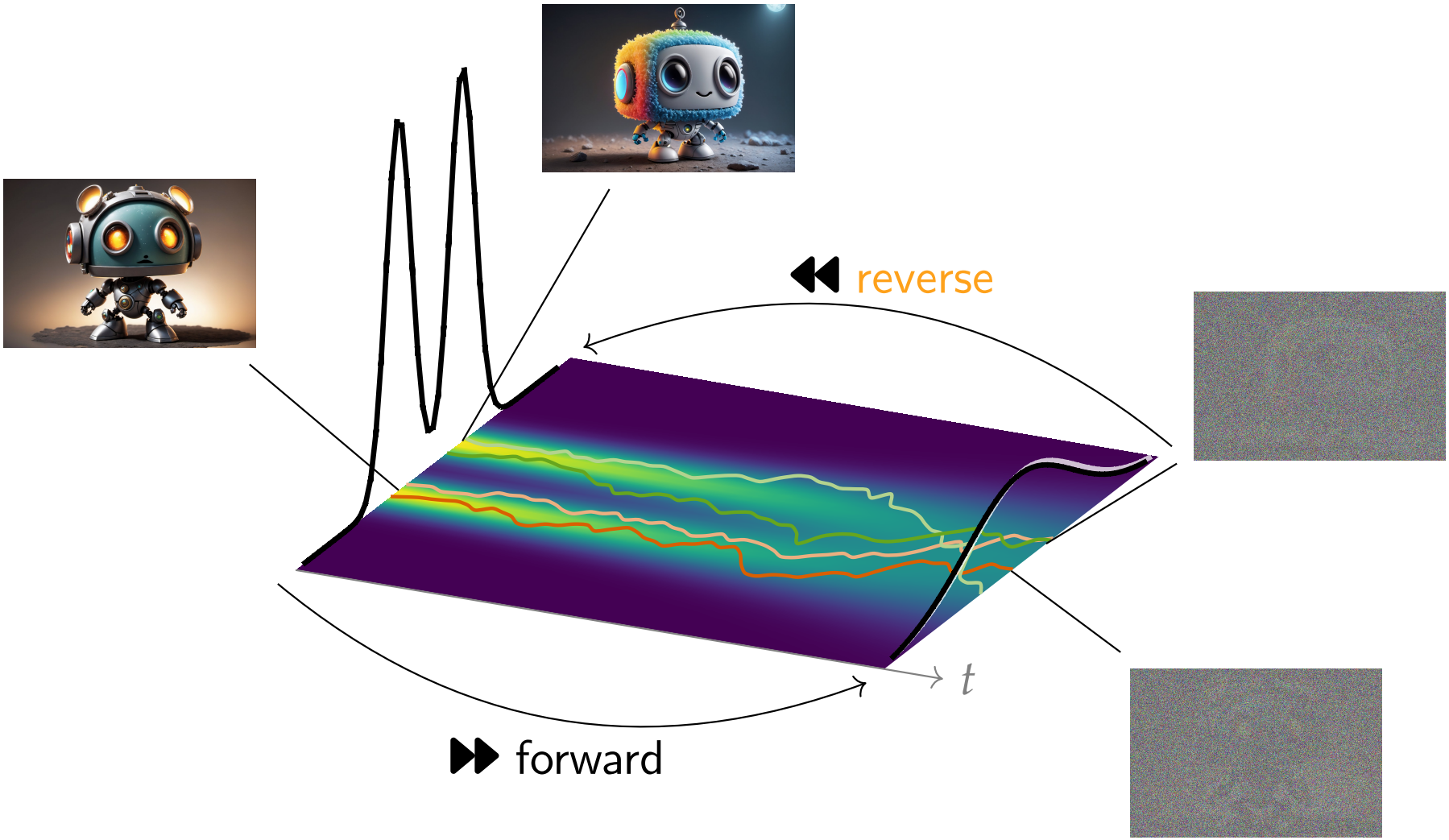
Distribution from $1\,000\,000$ diffused samples.

Data diffusion with SDEs



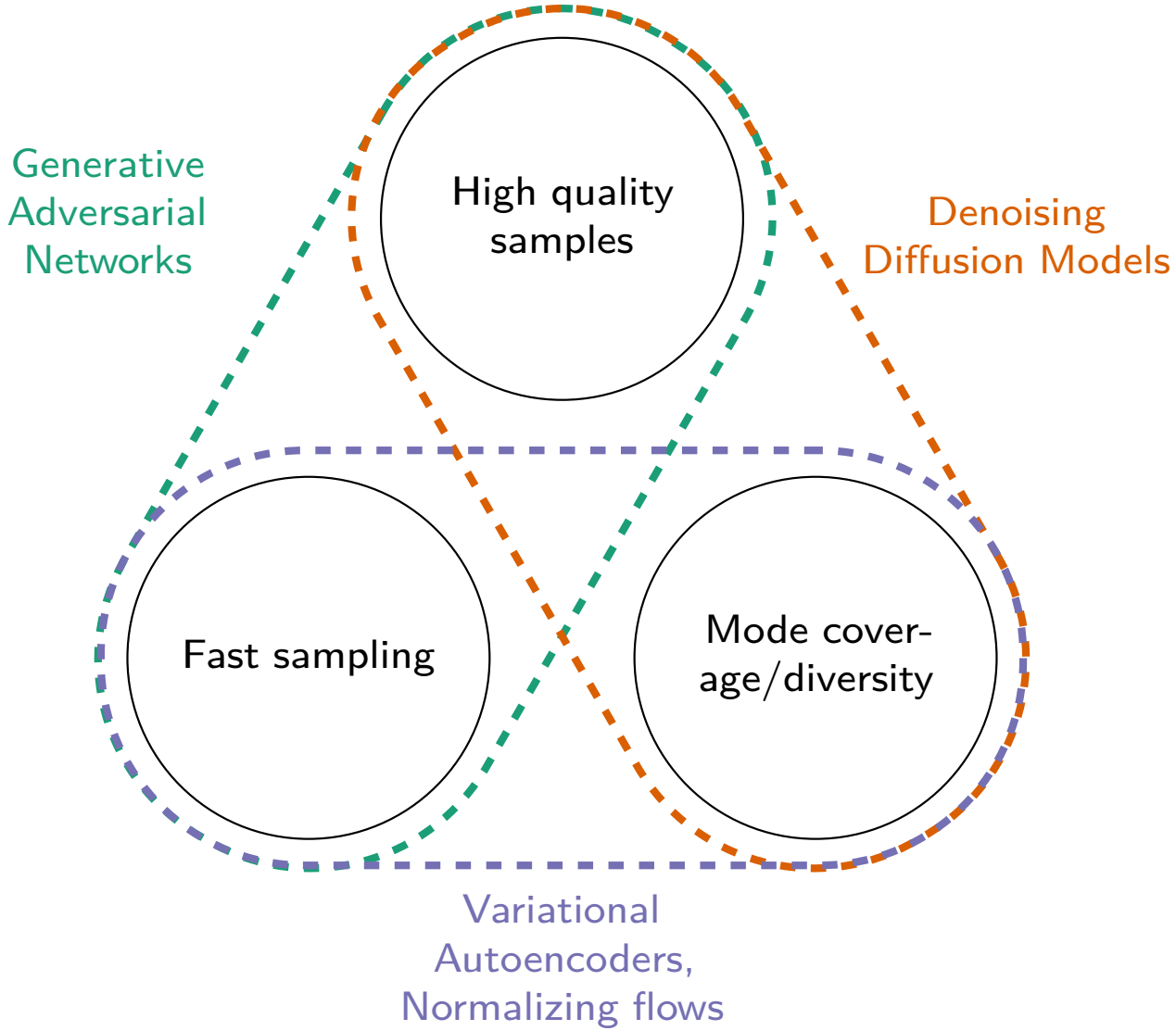
Distribution from $1\,000\,000$ diffused samples.

Data diffusion with SDEs



Distribution from $1\,000\,000$ diffused samples.

Generative learning trilemma



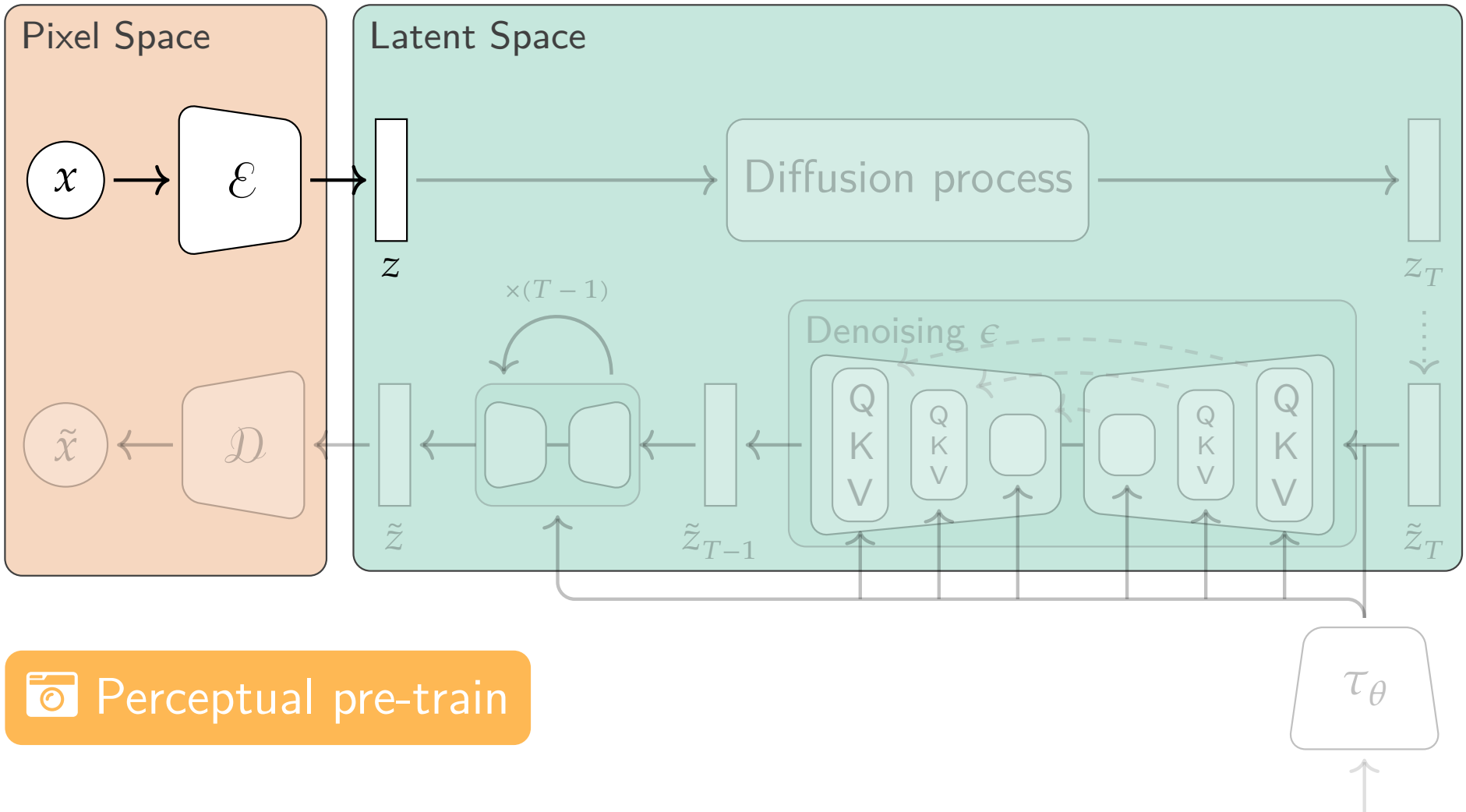
Xiao et al., "Tackling the Generative Learning Trilemma with Denoising Diffusion GANs." ICLR, 2022.

Acceleration techniques

- Advanced ODE/SDE solvers
- Distillation techniques
- Low-dimensional diffusion process
- Advanced diffusion process

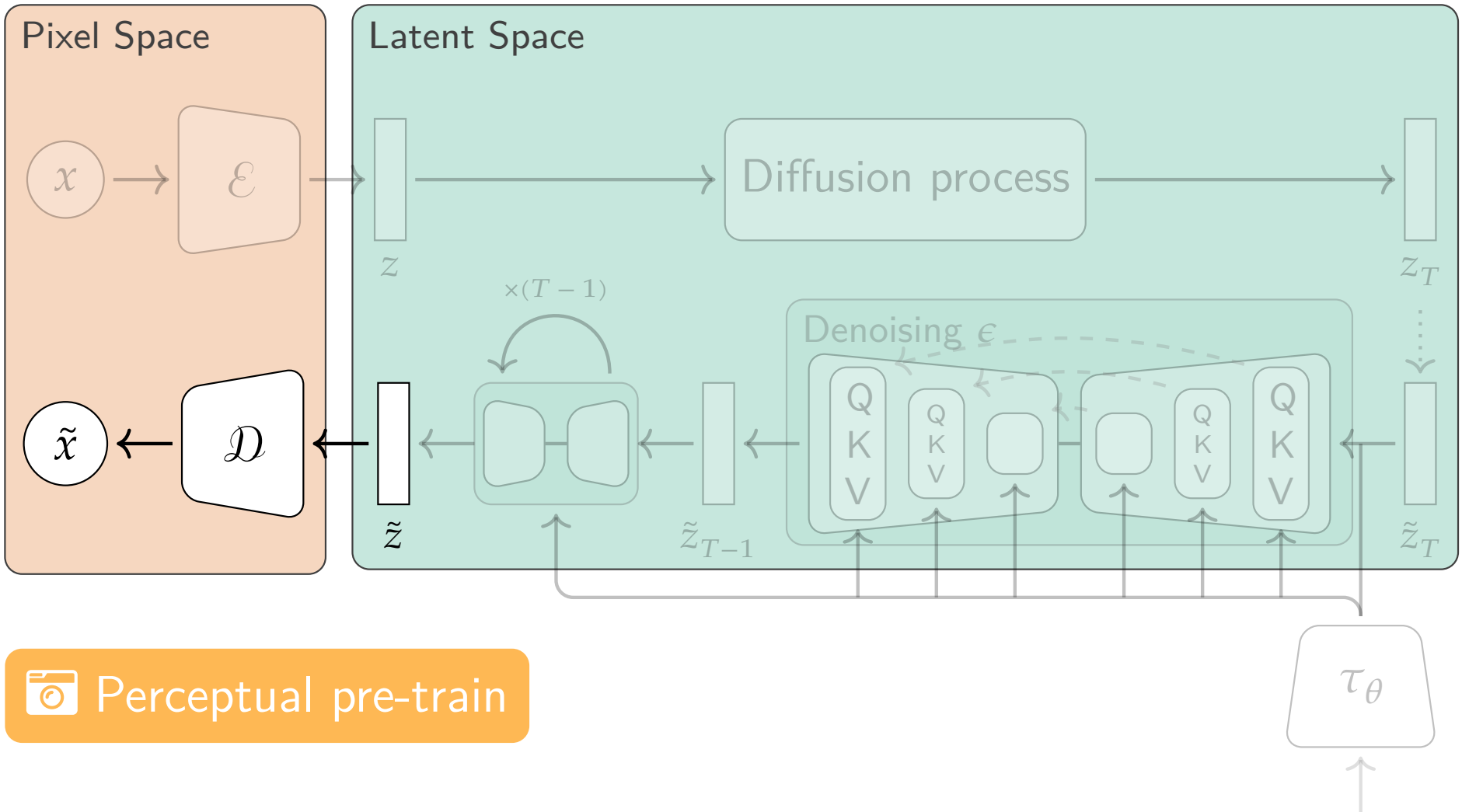
Latent Diffusion

Latent Diffusion



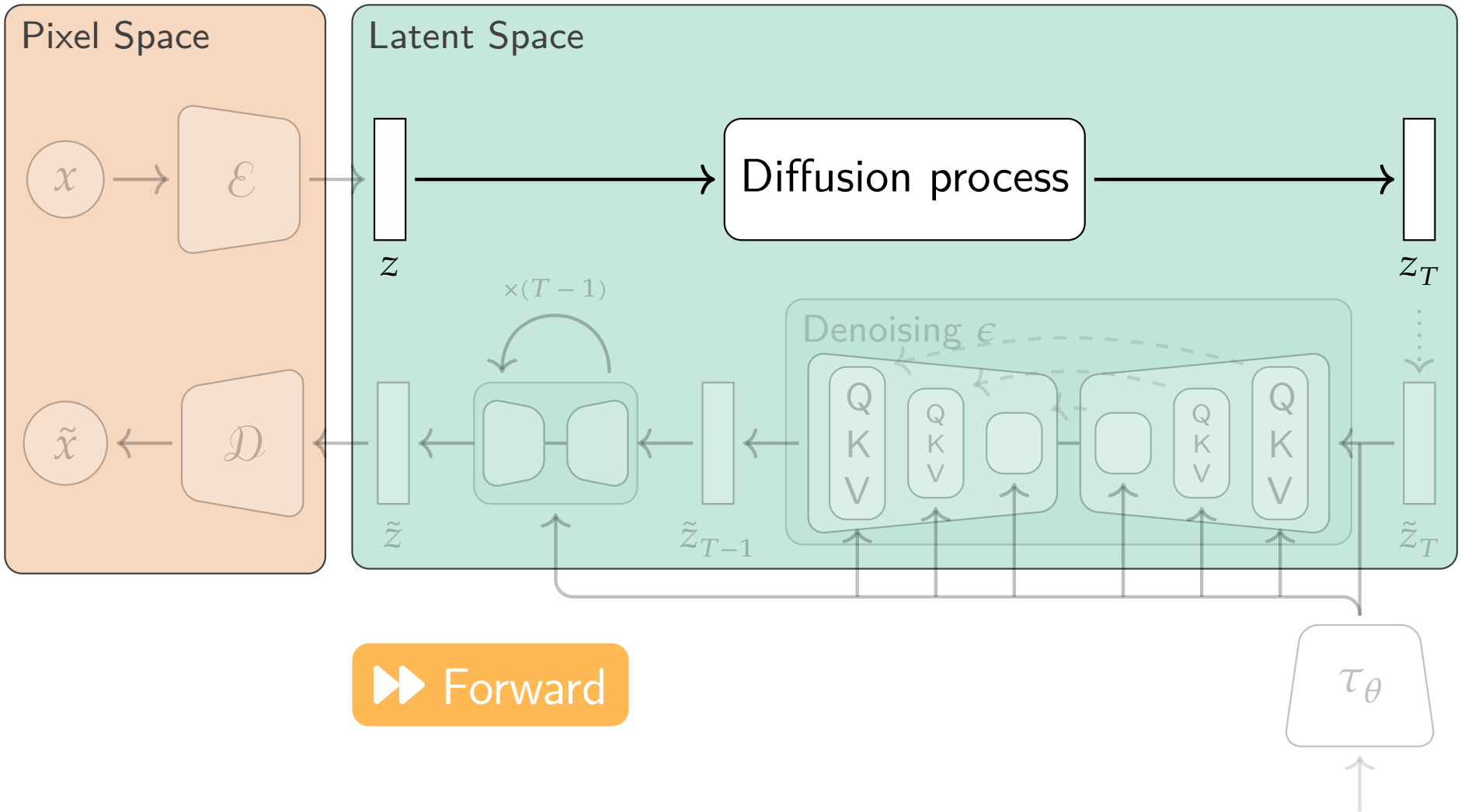
Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models." CVPR, 2022.

Latent Diffusion



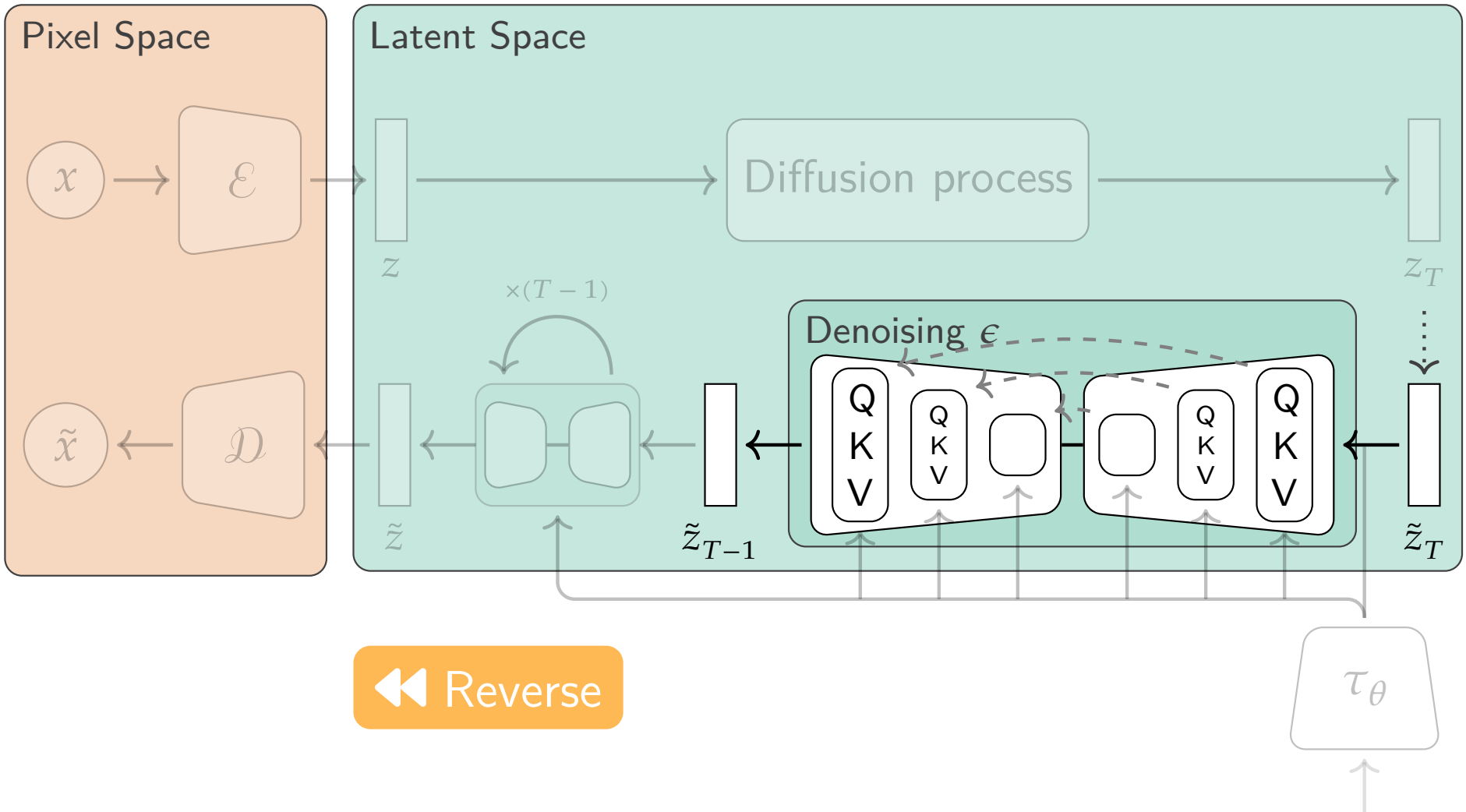
Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models." CVPR, 2022.

Latent Diffusion



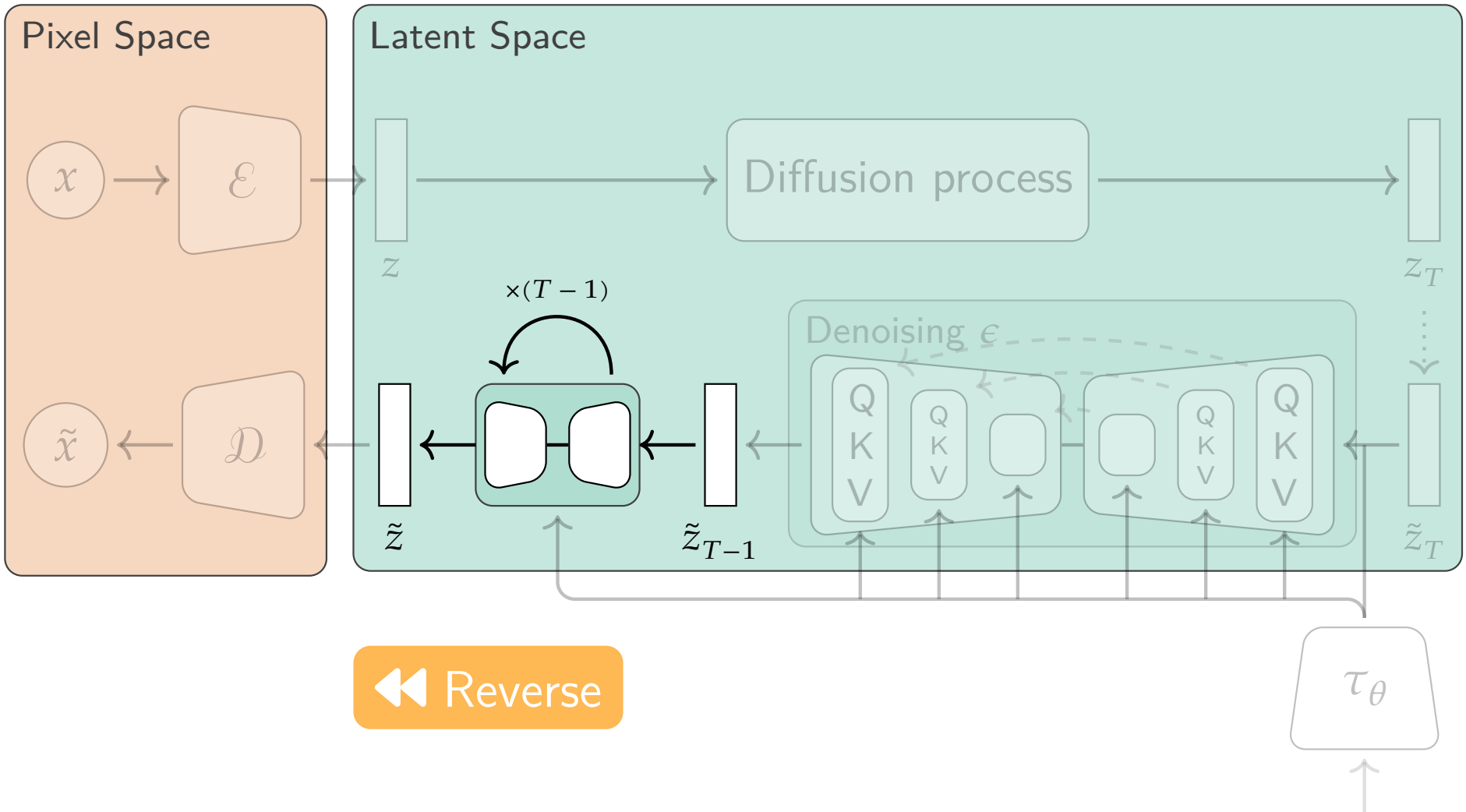
Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models." CVPR, 2022.

Latent Diffusion



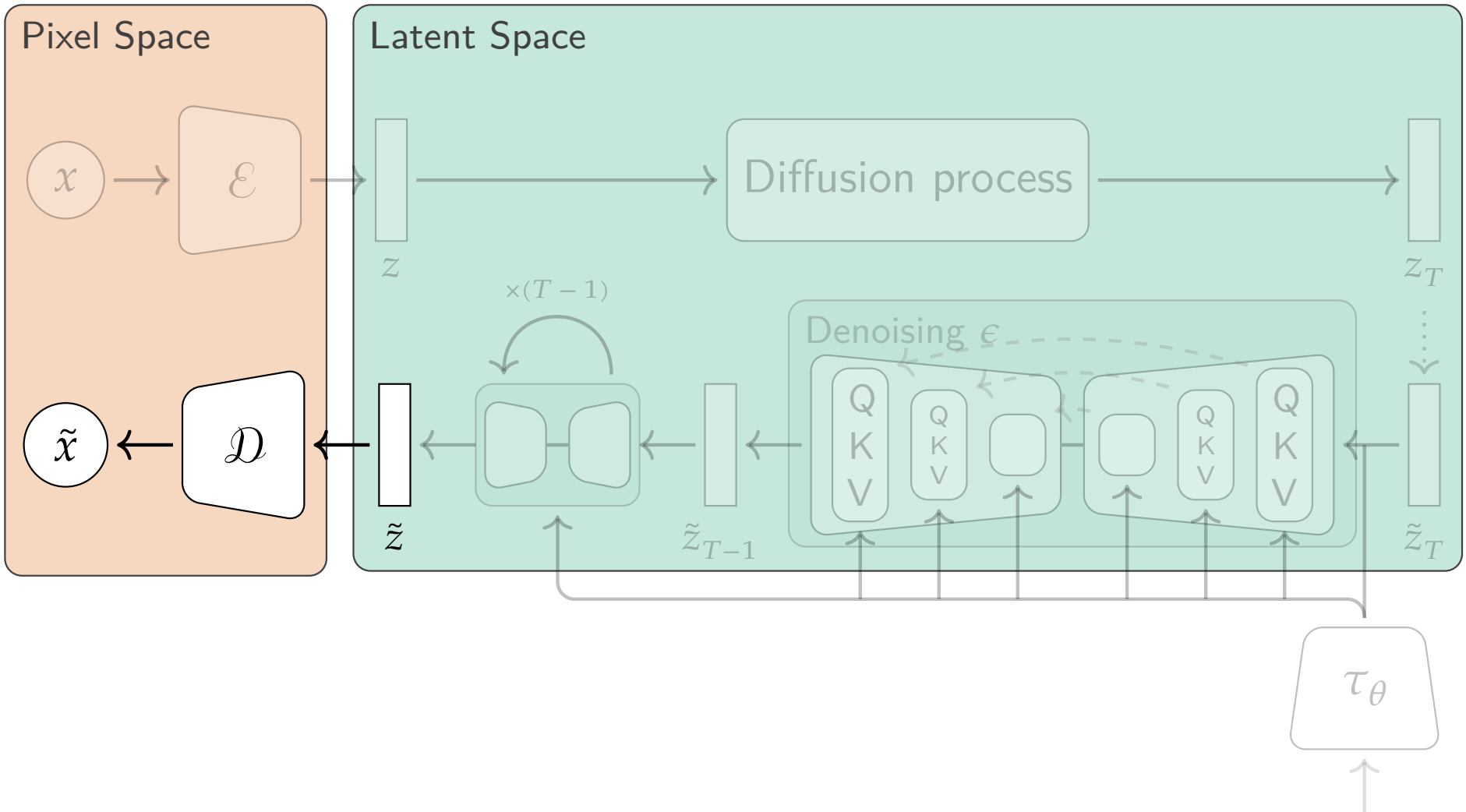
Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models." CVPR, 2022.

Latent Diffusion



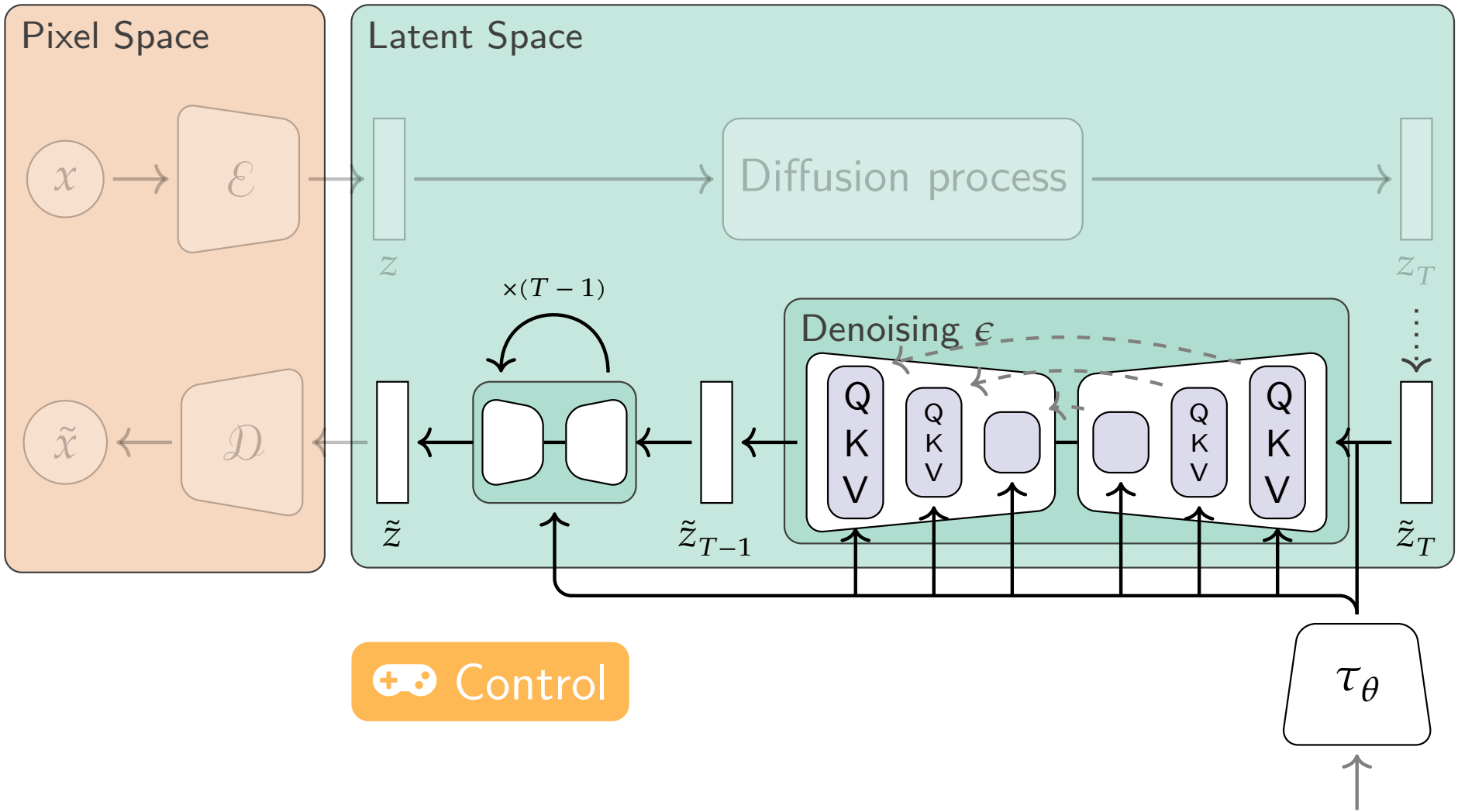
Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models." CVPR, 2022.

Latent Diffusion

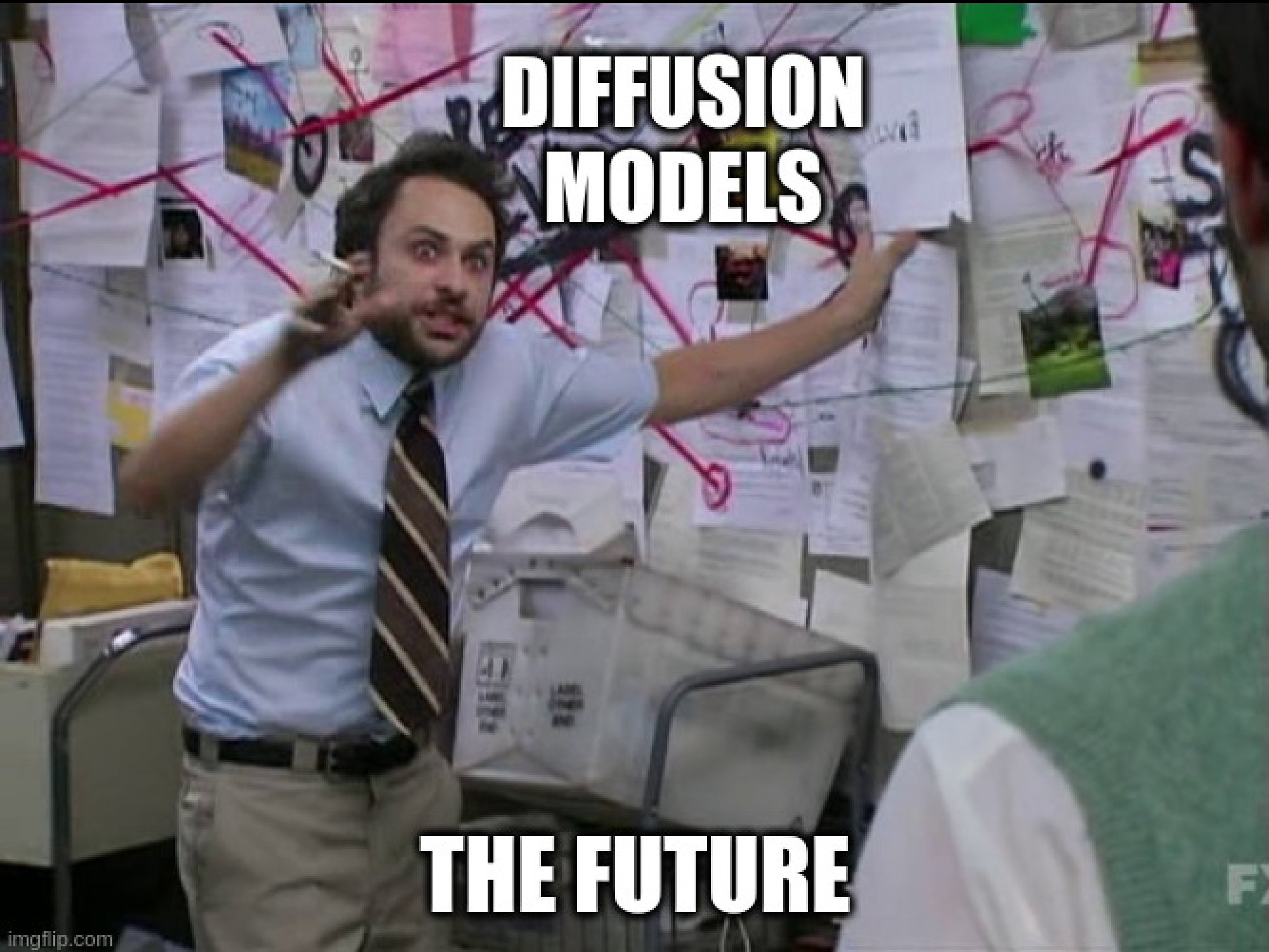


Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models." CVPR, 2022.

Latent Diffusion



Rombach et al., "High-Resolution Image Synthesis with Latent Diffusion Models." CVPR, 2022.


DIFFUSION MODELS
THE FUTURE
imgflip.com

Sora



Prompt: A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. Sh...



Anand Bhattad @anand_bhattad · Feb 17

Even Sora doesn't know projective geometry!🤔🤓

projective-geometry.github.io



Sarkar et al., "Shadows Don't Lie and Lines Can't Bend! Generative Models don't know Projective Geometry...for now." CVPR, 2024.

Sora as World Models



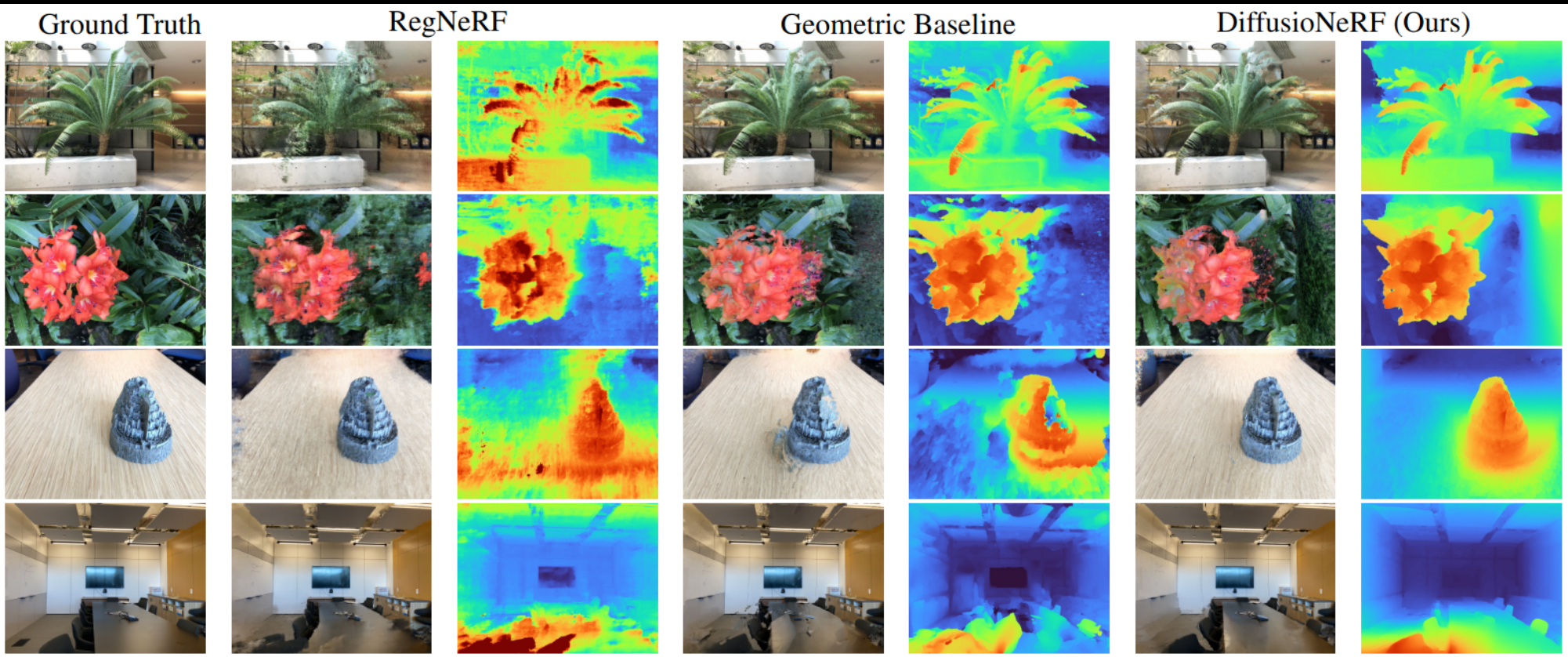


Real-time game engines



Alonso et al., "Diffusion for World Modeling: Visual Details Matter in Atari." arXiv, 2405.12399v1.
Valevski et al., "Diffusion Models Are Real-Time Game Engines." arXiv, 2408.14837v1.

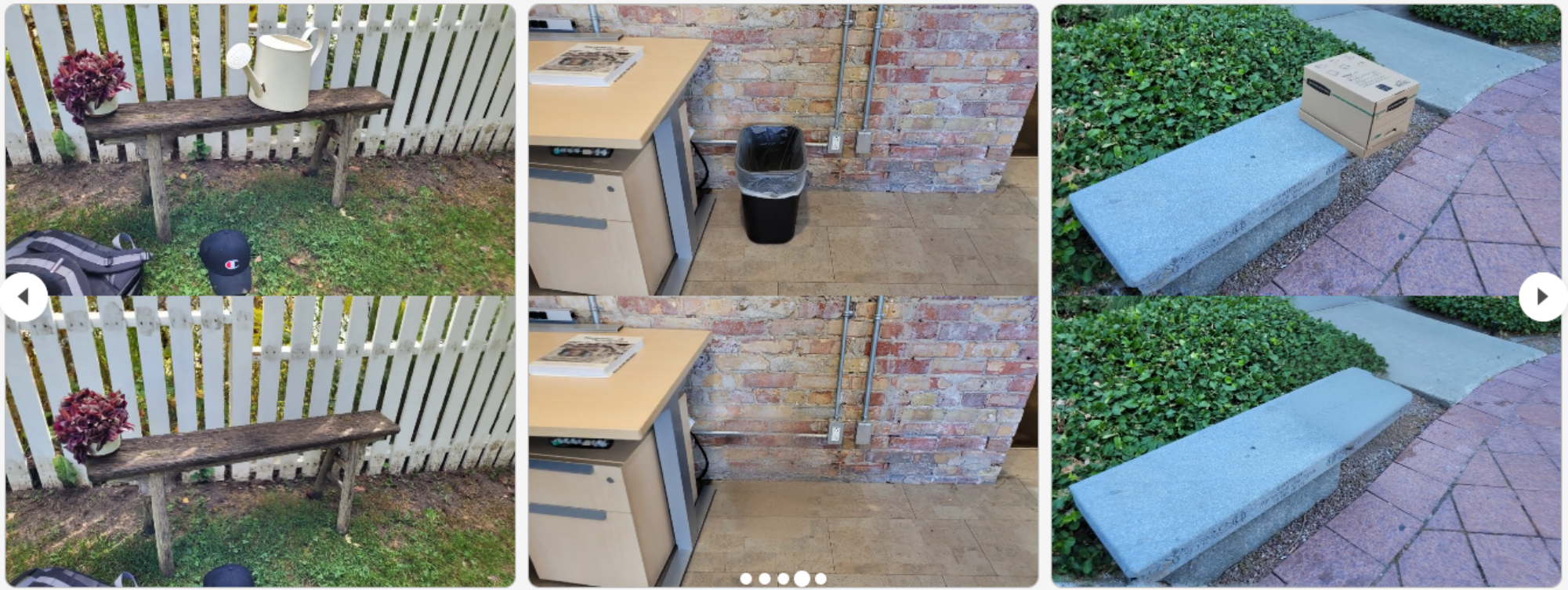
NeRF meets Diffusion



Wynn & Turmukhambetov, "DiffusioNeRF: Regularizing Neural Radiance Fields with Denoising Diffusion Models." CVPR, 2023.

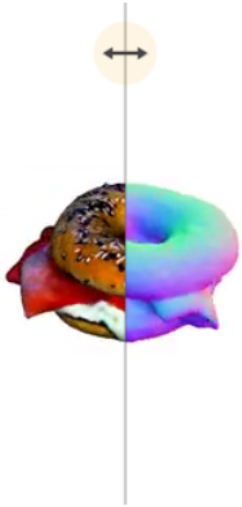


Shih et al., "ExtraNeRF: Visibility-Aware View Extrapolation of Neural Radiance Fields with Diffusion Models," CVPR, 2024.



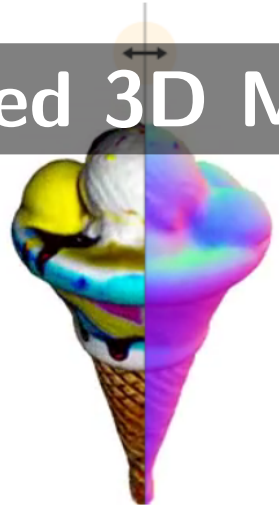
Lin et al., “Taming Latent Diffusion Model for Neural Radiance Field Inpainting.” ECCV, 2024.

Diffused 3D Models



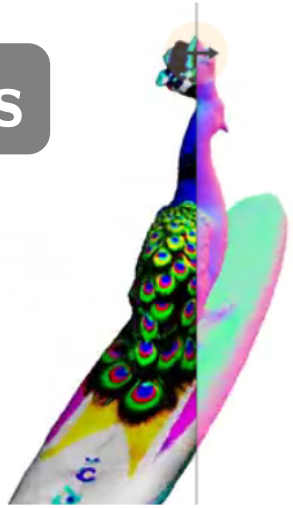
Reveal 3D mesh!

[...] a bagel filled with cream cheese and lox.



Reveal 3D mesh!

[...] an ice cream sundae.



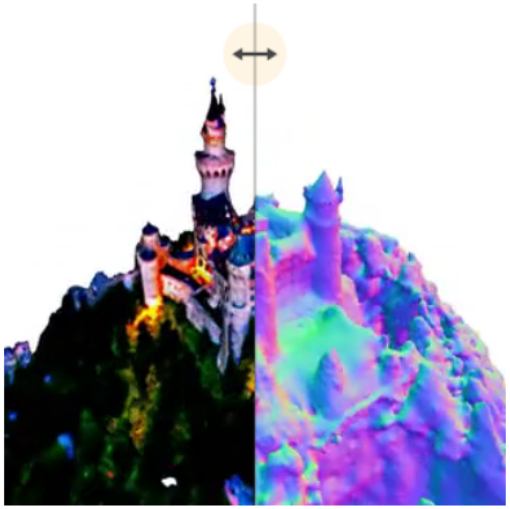
Reveal 3D mesh!

[...] a peacock on a surfboard.



Reveal 3D mesh!

[...] a plate piled high with chocolate chip cookies.



Reveal 3D mesh!

[...] Neuschwanstein Castle, aerial view.



Reveal 3D mesh!

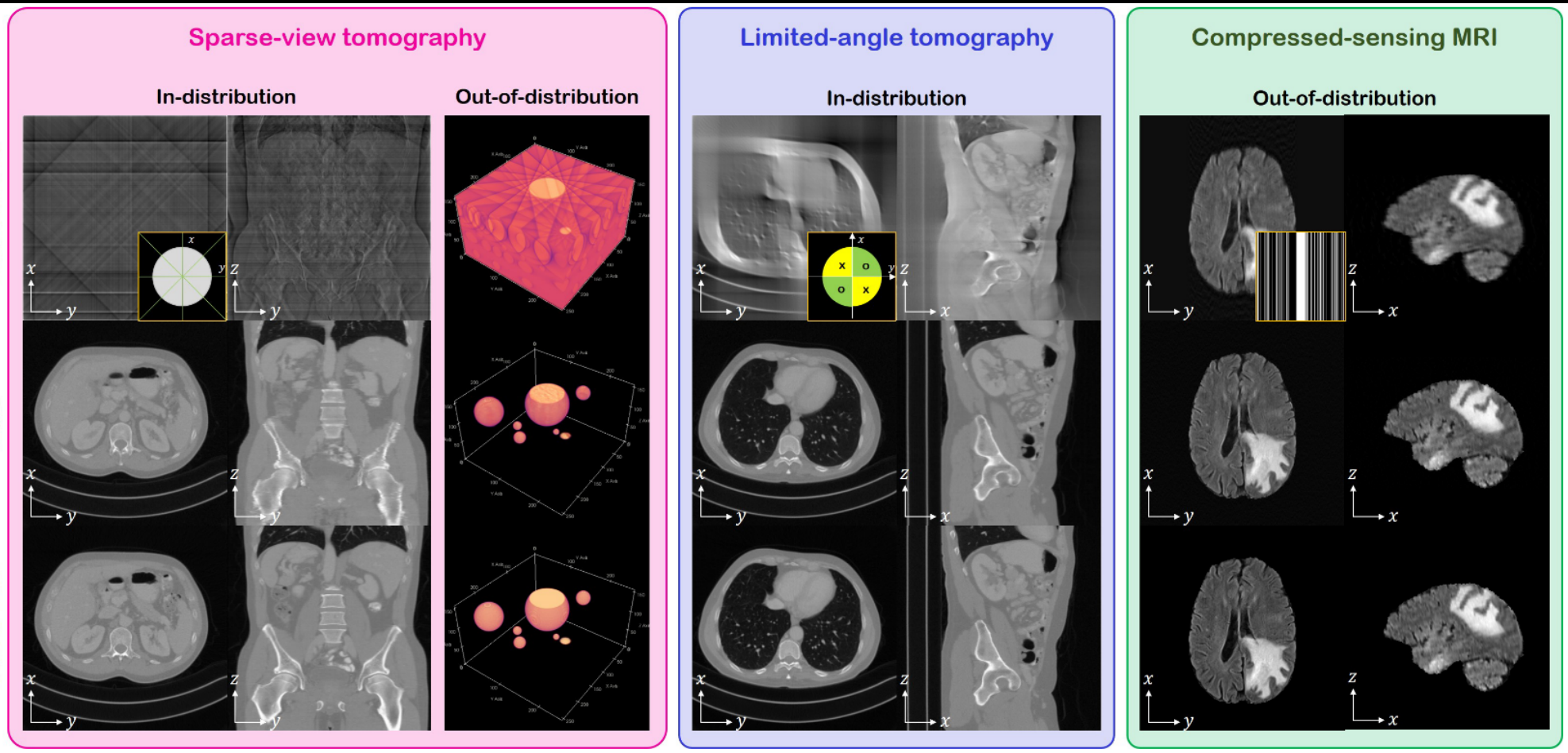
[...] the Imperial State Crown of England.

Lin et al., "Magic3D: High-Resolution Text-to-3D Content Creation." CVPR, 2023.

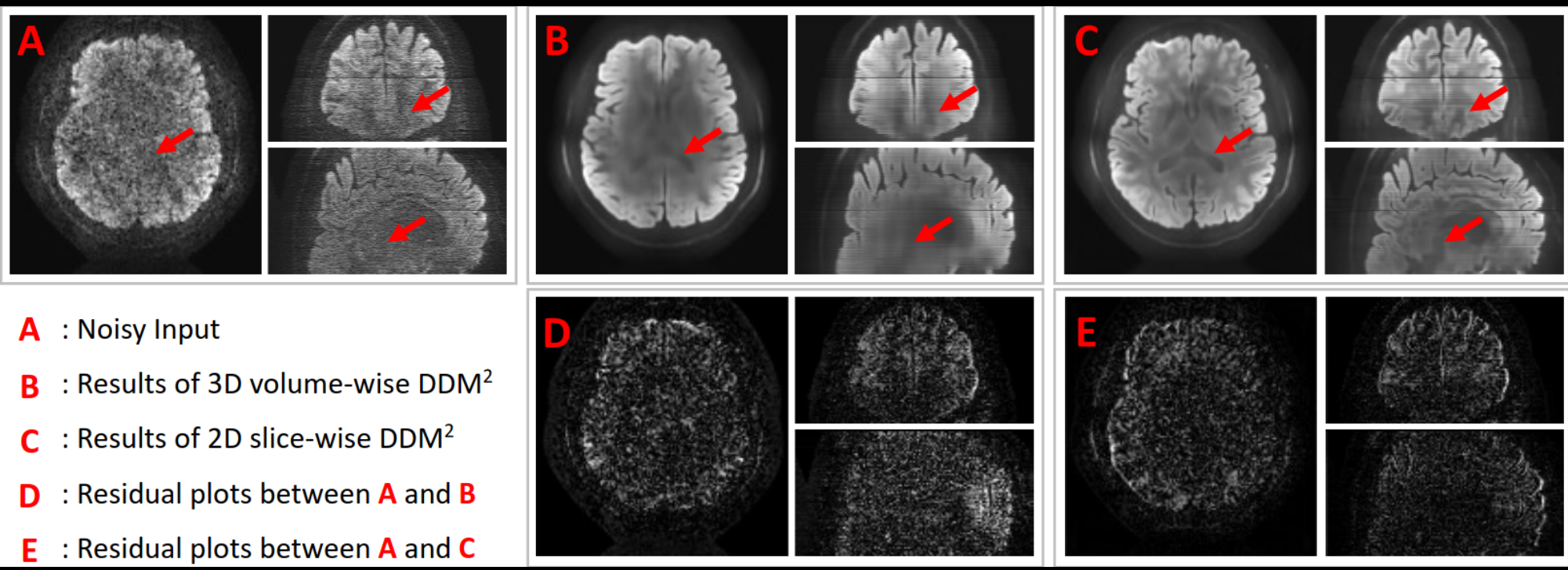


Liu et al., "Zero-1-to-3: Zero-shot One Image to 3D Object." ICCV, 2023.

Diffusion on Medical Imaging



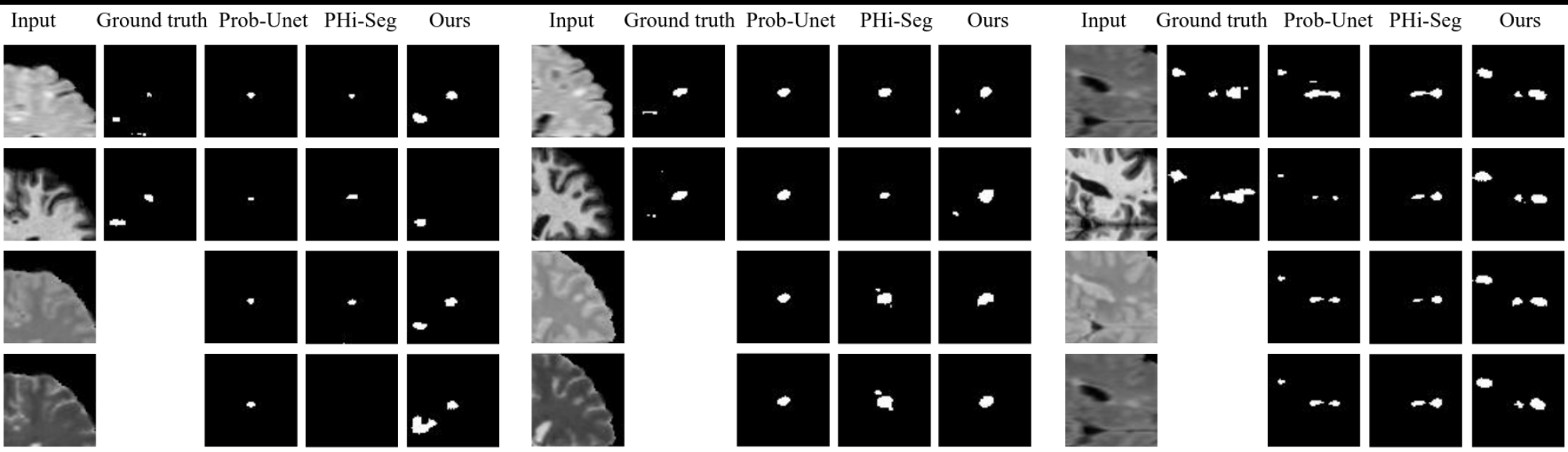
Chung et al., "Solving 3D Inverse Problems using Pre-trained 2D Diffusion Models." CVPR, 2023.



Xiang et al., "DDM2: Self-Supervised Diffusion MRI Denoising with Generative Diffusion Models." ICLR, 2023.

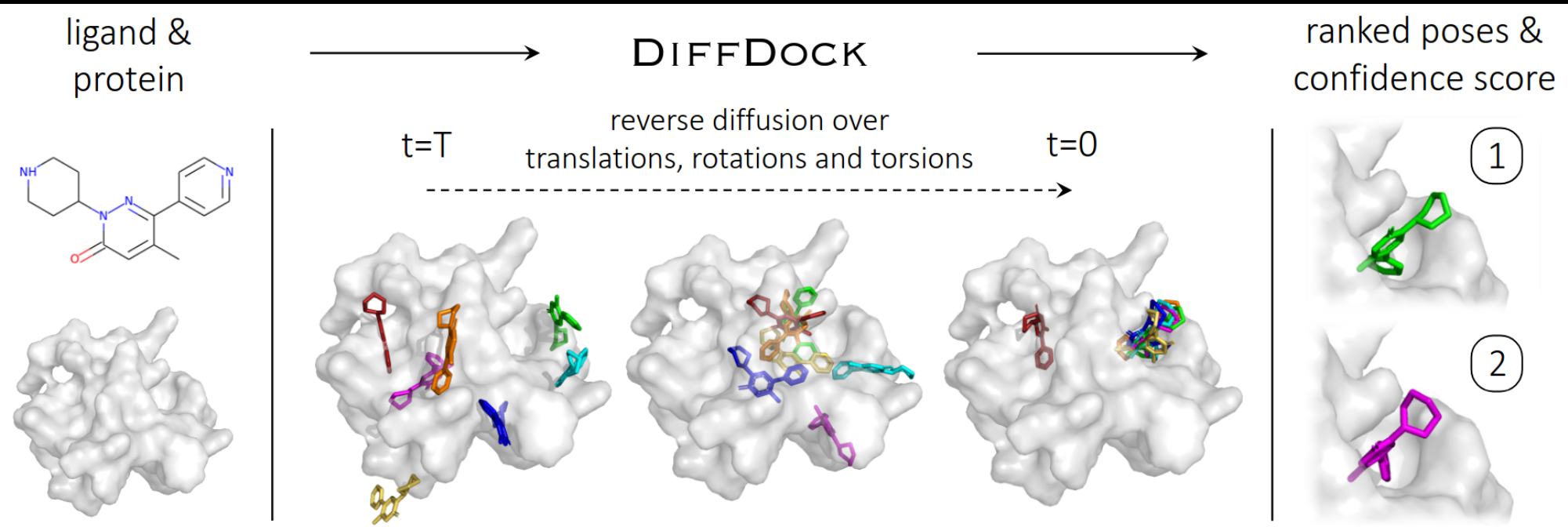


Chen et al., "HiDiff: Hybrid Diffusion Framework for Medical Image Segmentation." IEEE TMI, 2024.



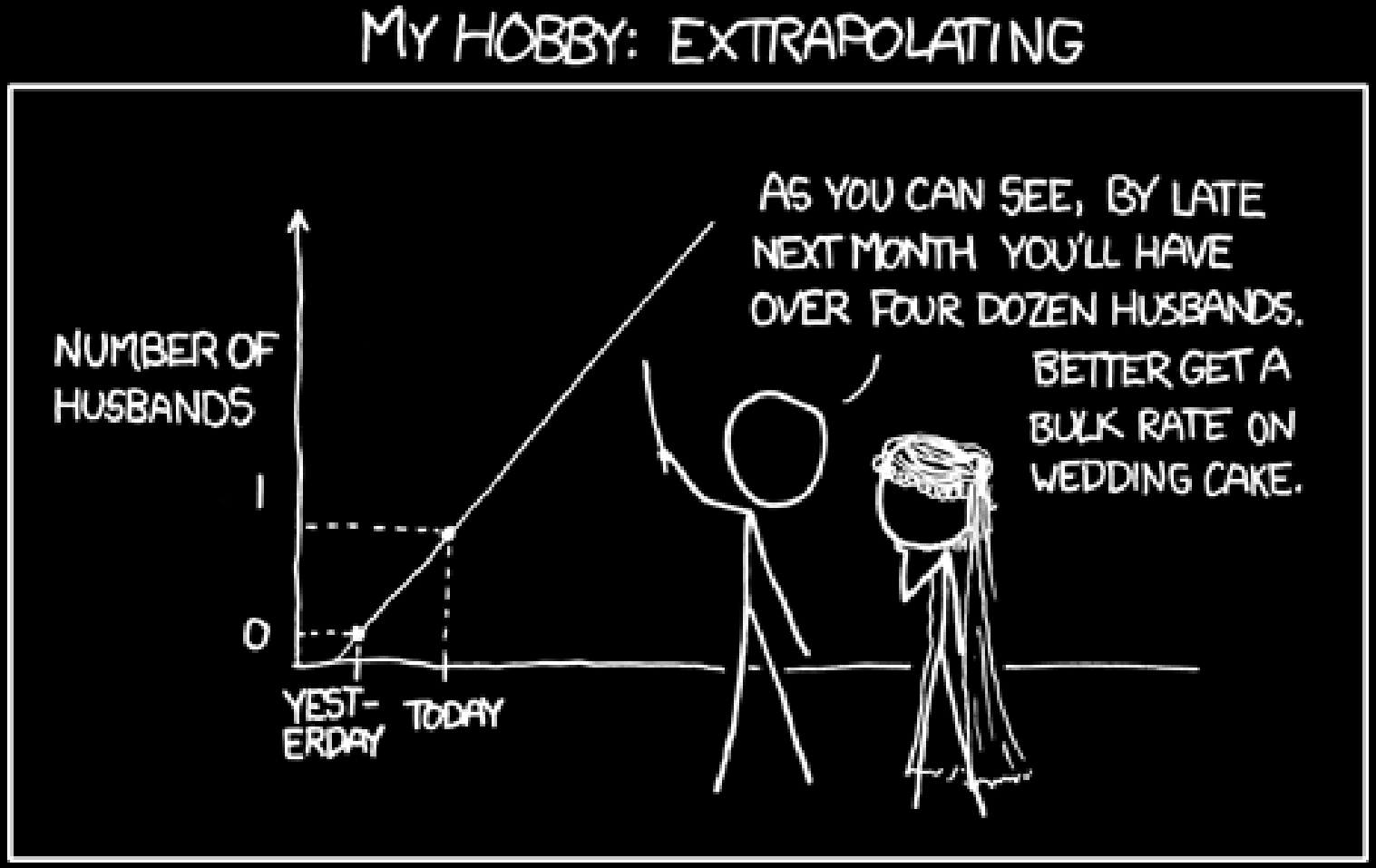
Rahman et al., "Ambiguous Medical Image Segmentation using Diffusion Models." CVPR, 2023.

And more



Corso et al., "DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking." ICLR, 2023.
Corso et al., "Deep Confident Steps to New Pockets: Strategies for Docking Generalization." ICLR, 2024.

And many more



https://xkcd.com/605/

The Nobel Prize in Chemistry 2024



Ill. Niklas Elmehed © Nobel Prize Outreach

David Baker

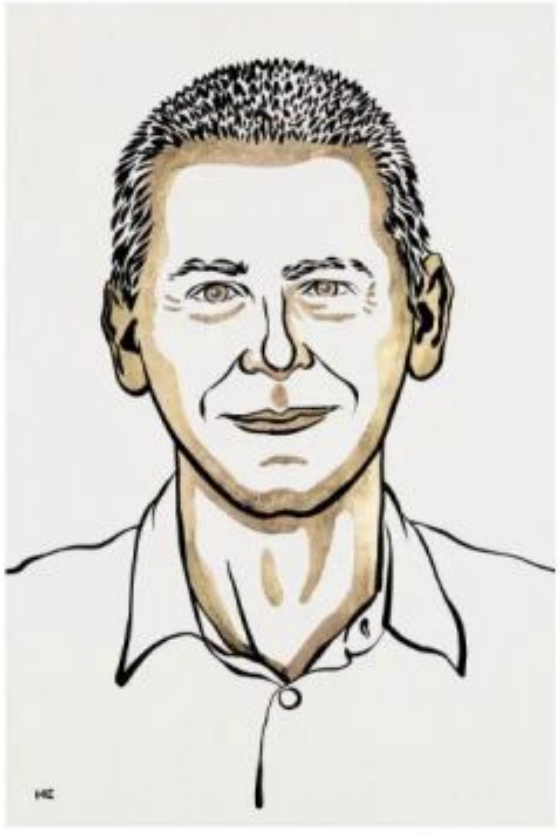
Prize share: 1/2



Ill. Niklas Elmehed © Nobel Prize Outreach

Demis Hassabis

Prize share: 1/4



Ill. Niklas Elmehed © Nobel Prize Outreach

John M. Jumper

Prize share: 1/4



Take away points

Main ideas

- Based on first principles: math strikes back
- Sped-up into success
- Data efficient: need less data
- Exploding number of applications, but not perfect

Main ideas

- Based on first principles: math strikes back
- Sped-up into success
- Data efficient: need less data
- Exploding number of applications, but not perfect

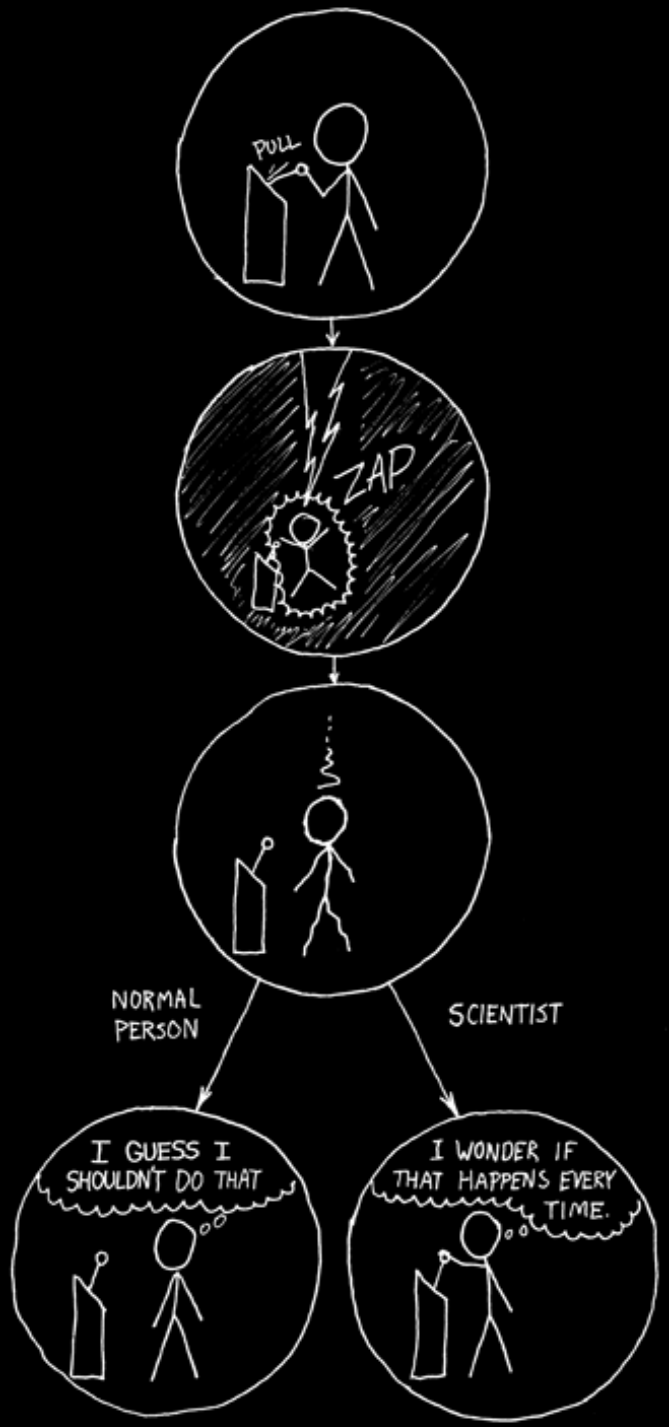
Main ideas

- Based on first principles: math strikes back
- Sped-up into success
- Data efficient: need less data
- Exploding number of applications, but not perfect

Main ideas

- Based on first principles: math strikes back
- Sped-up into success
- Data efficient: need less data
- Exploding number of applications, but not perfect

We must keep exploring



https://xkcd.com/242/

Main ideas

Thank you!

- Based on first principles: math strikes back
- Sped-up into success
- Data efficient: need less data
- Exploding number of applications, but not perfect

Let's chat:
adinr@uio.no adin@gitlab.io @adn_twitts @adnrv.bsky.social