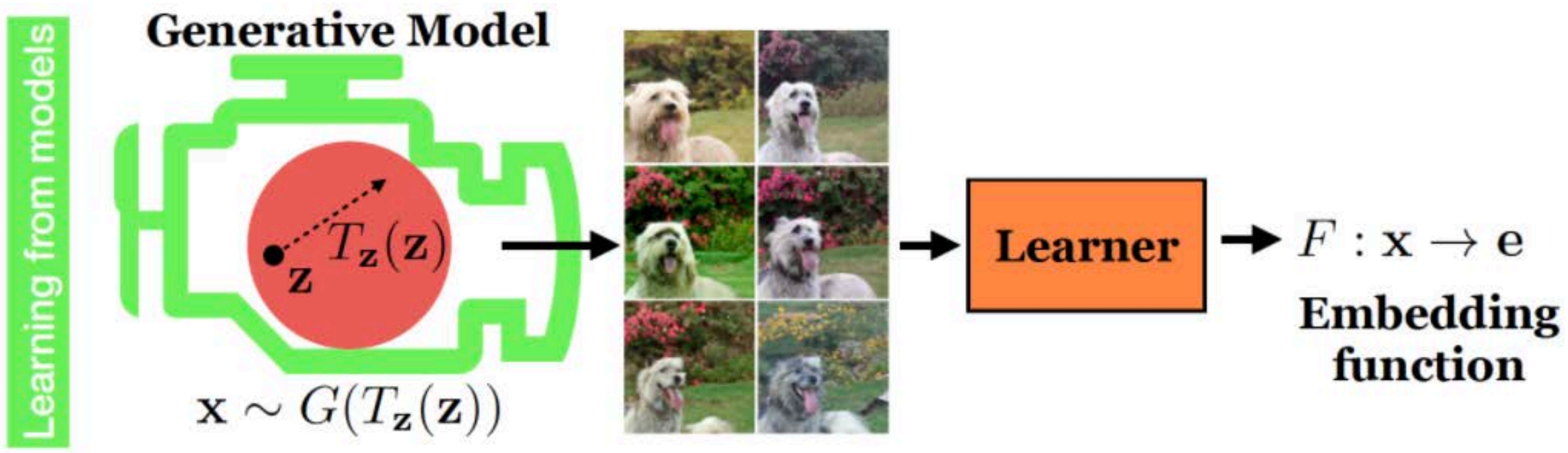
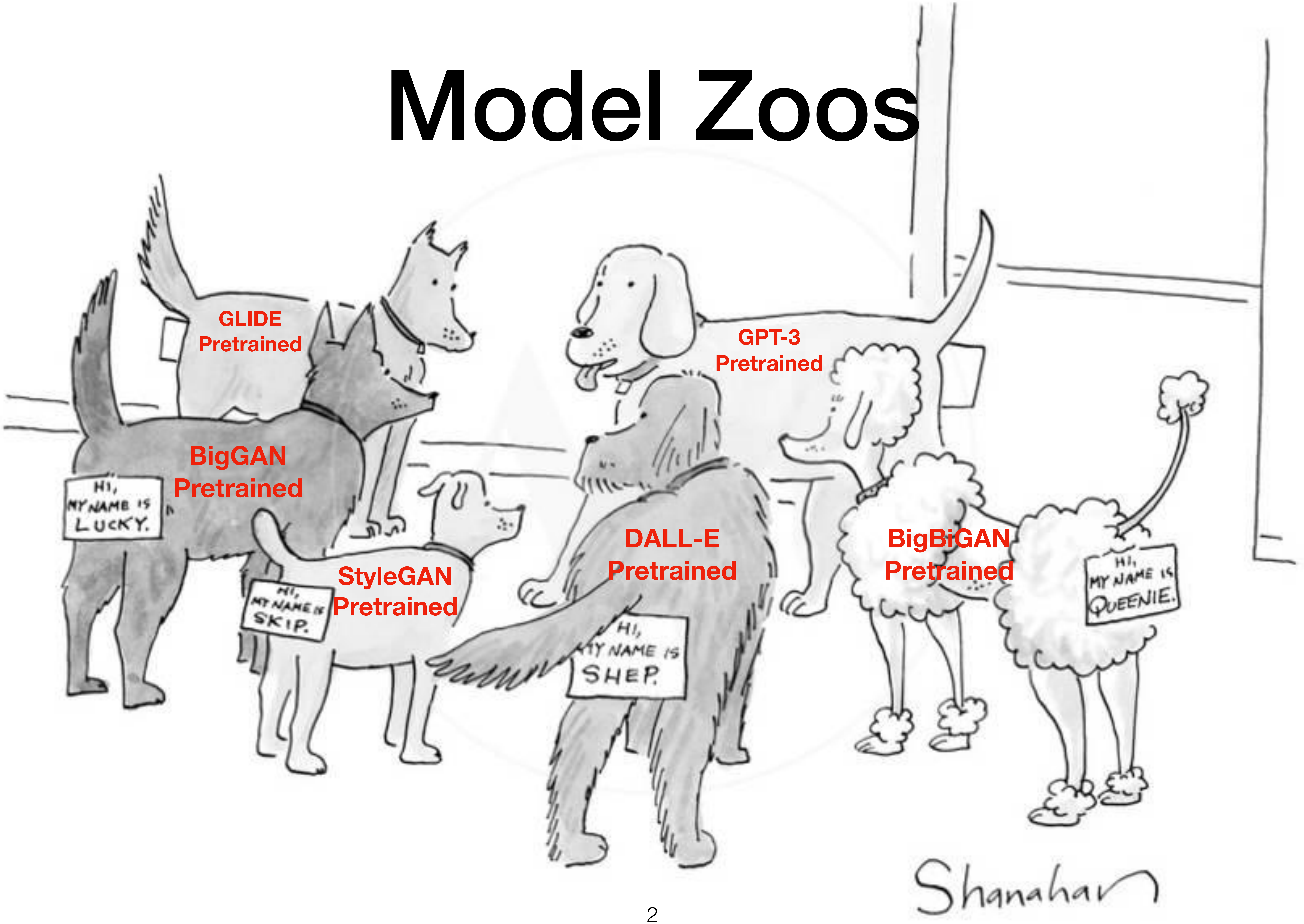
GENERATIVE MODELS AS A DATA SOURCE FOR MULTIVIEW REPRESENTATION LEARNING

Ali Jahanian, Xavier Puig, Yonglong Tian, Phillip Isola

ICLR 2022

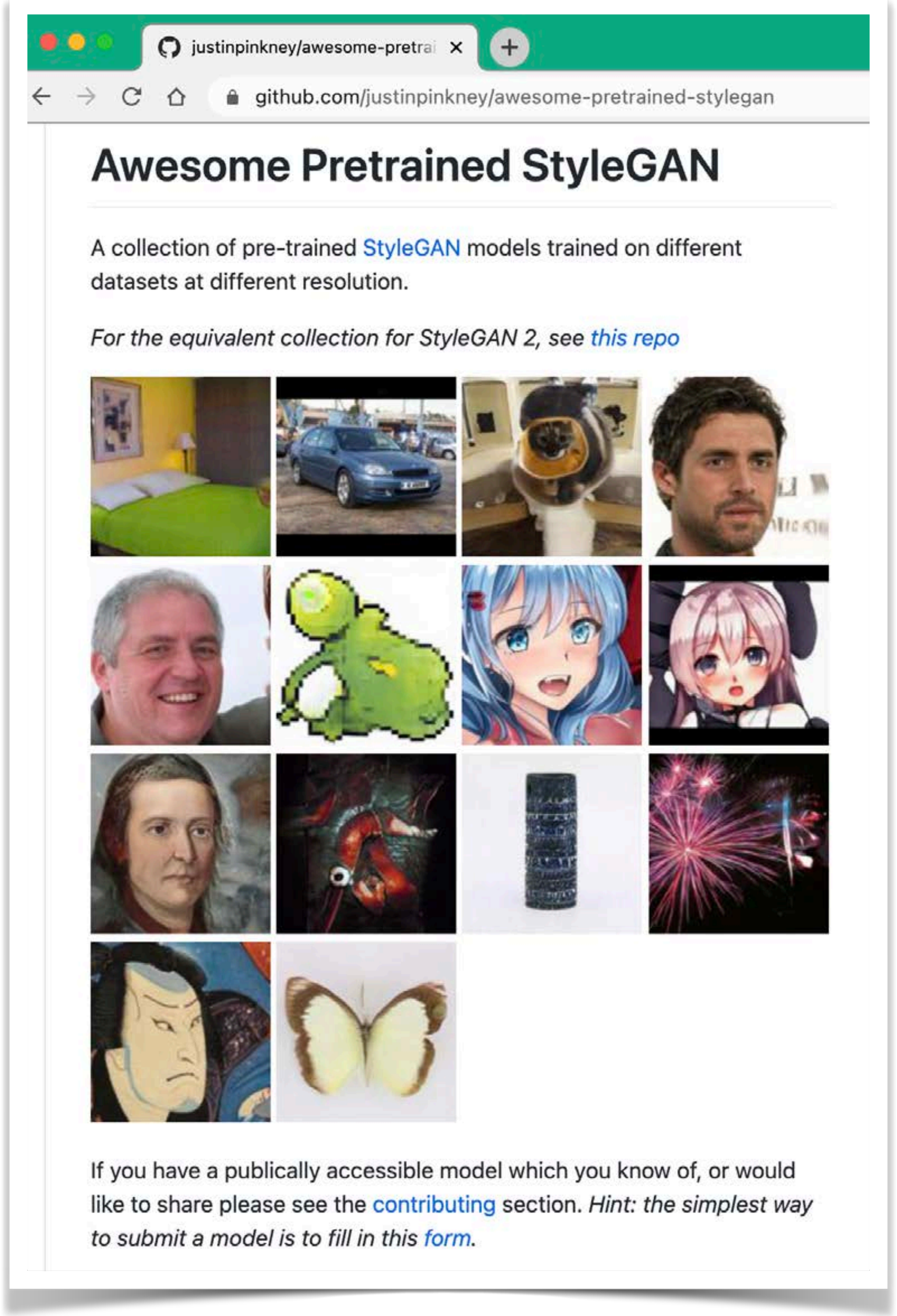


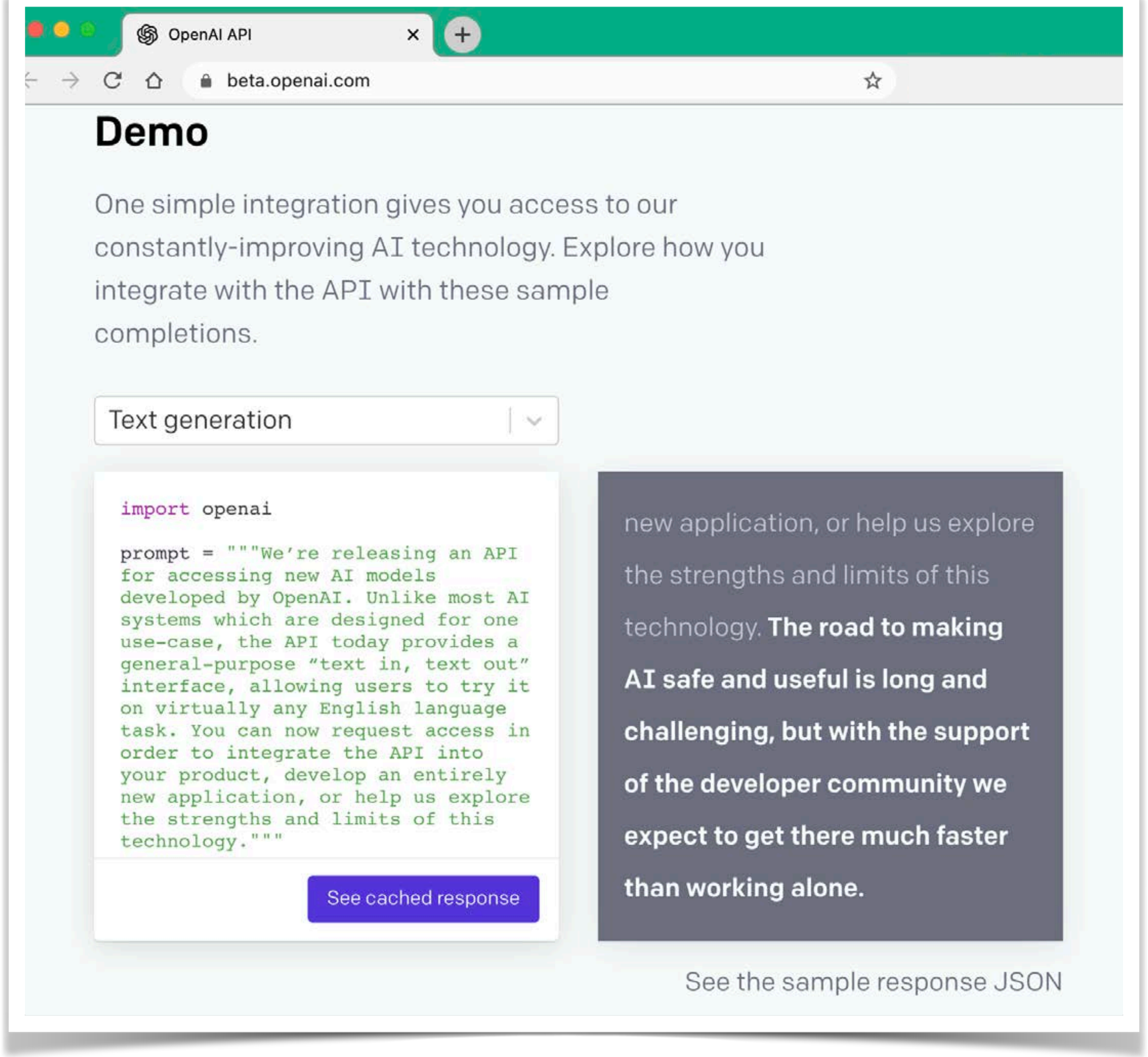
Model Zoos

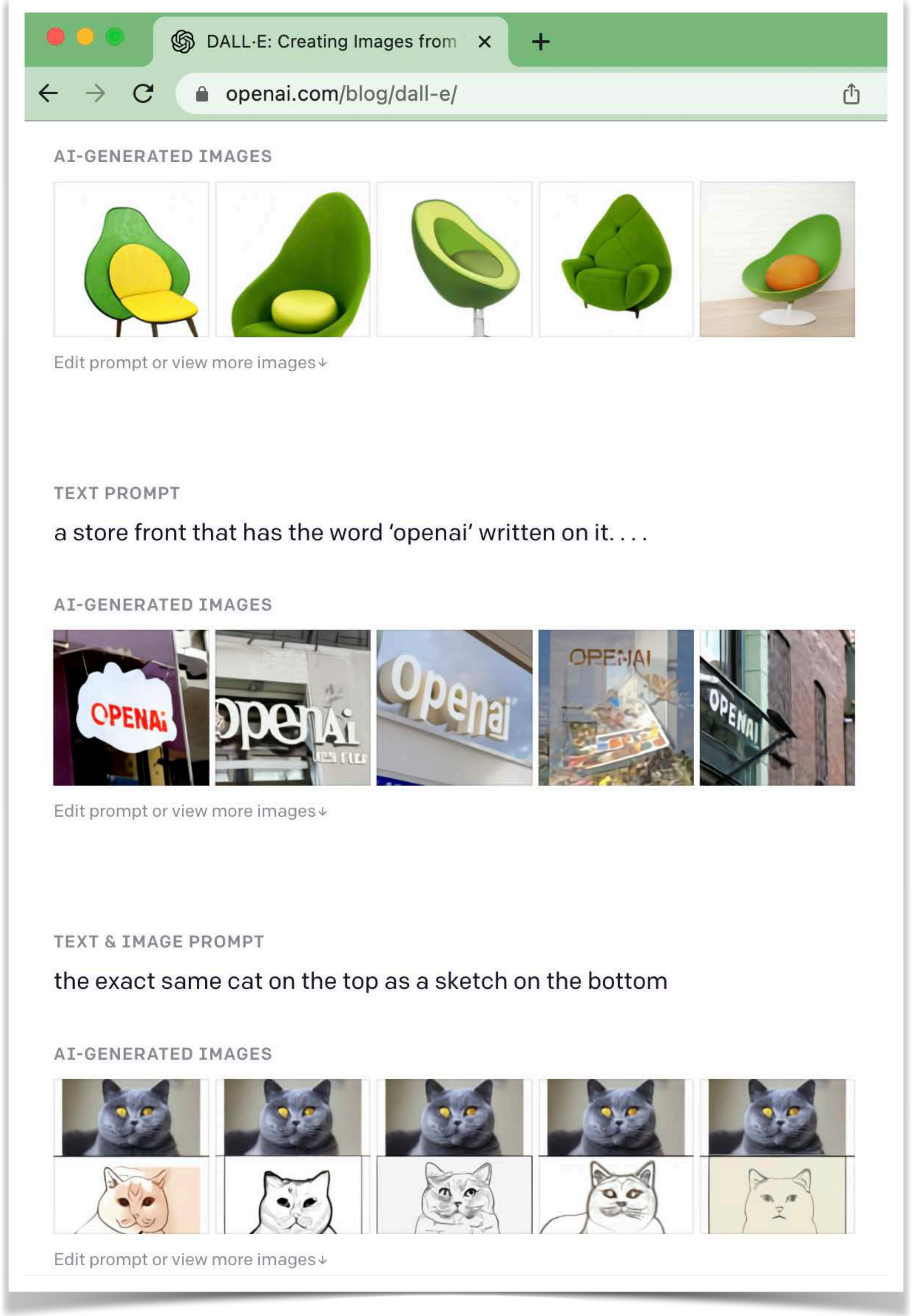

GLIDE
Pretrained
BigGAN
Pretrained
HI,
MY NAME IS
LUCKY.
StyleGAN
Pretrained
HI,
MY NAME IS
SKIP.
GPT-3
Pretrained
DALL-E
Pretrained
HI,
MY NAME IS
SHEP.
BigBiGAN
Pretrained
HI,
MY NAME IS
QUEENIE.
Shanahan

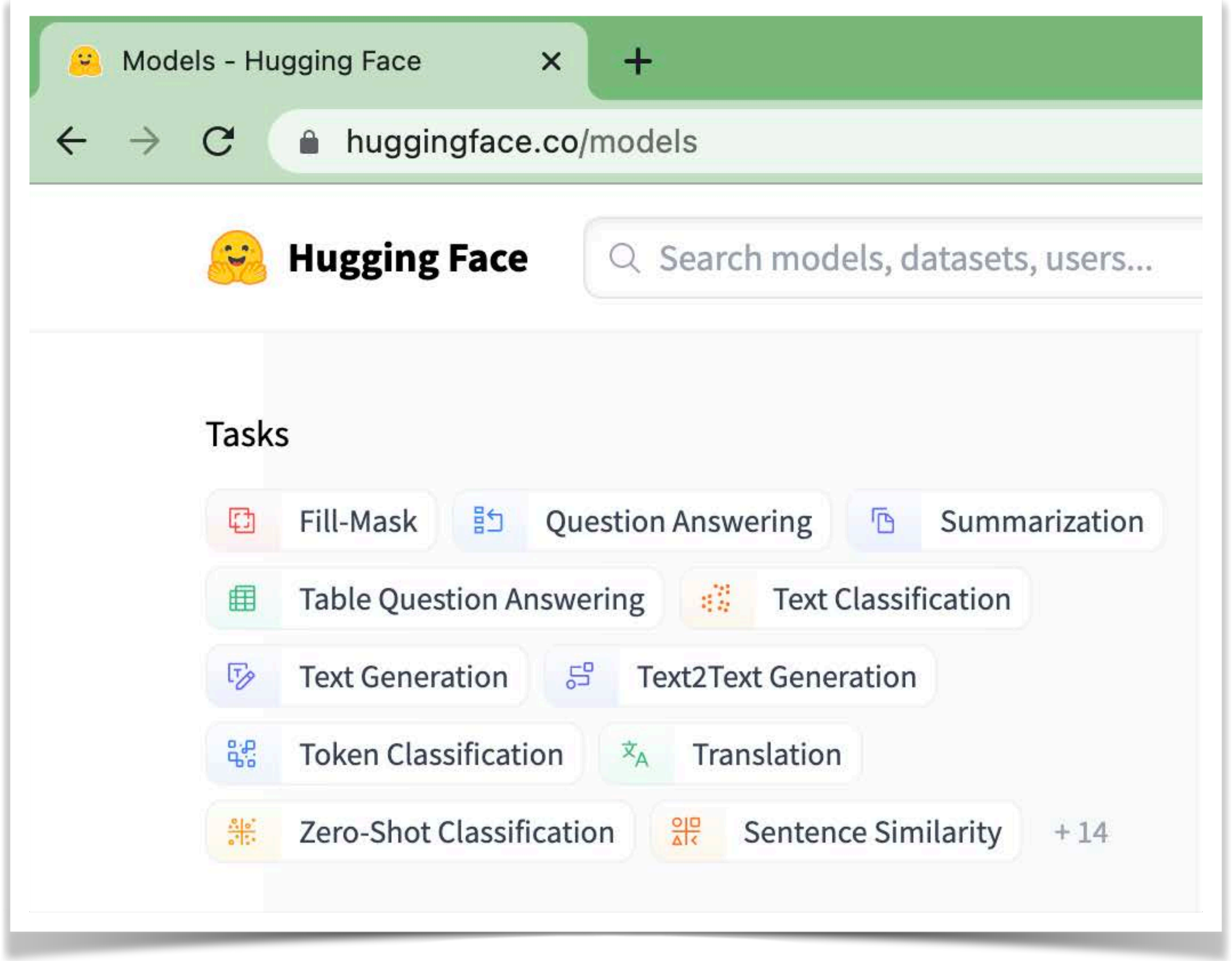
Model Zoos

The Era of Big Data —> The Era of Big Models?



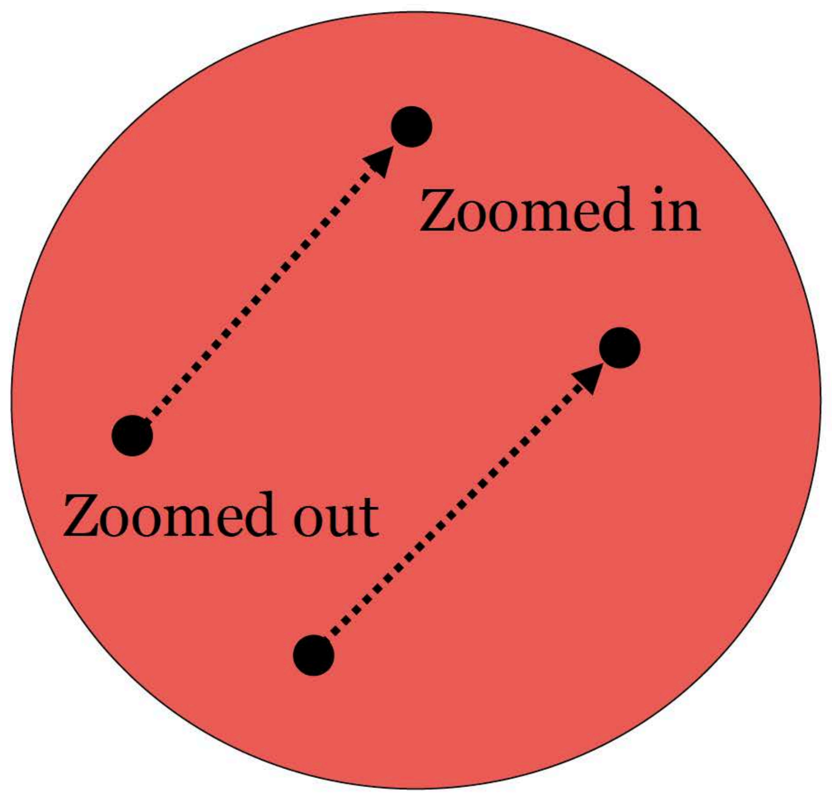






These models are steerable

By steering in the latent space of IGMs (Implicit Generative Models) the data all of the sudden become alive!



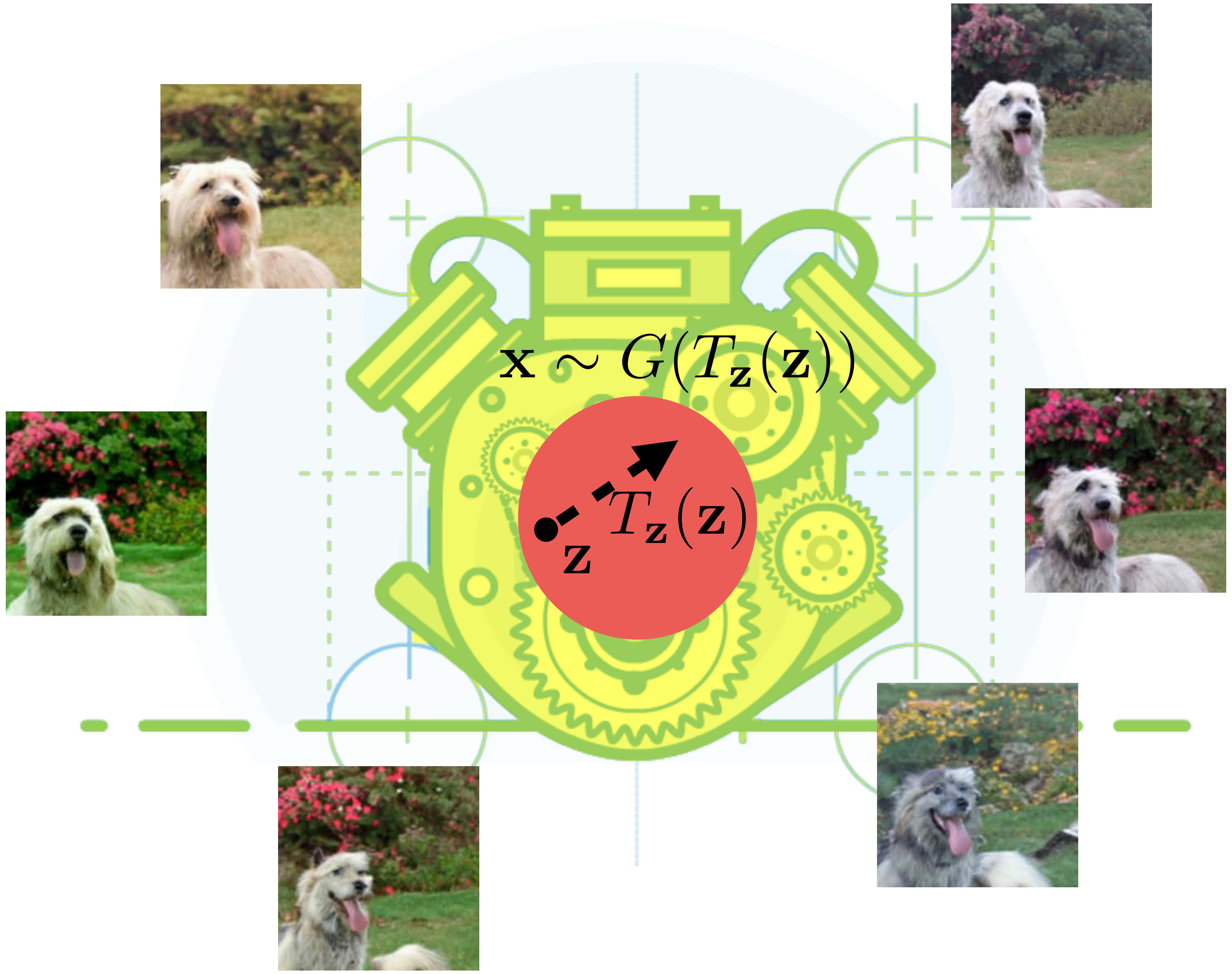
Latent Space

[On the "steerability" of GANs: Jahanian*, Chai*, Isola, ICLR 2020]

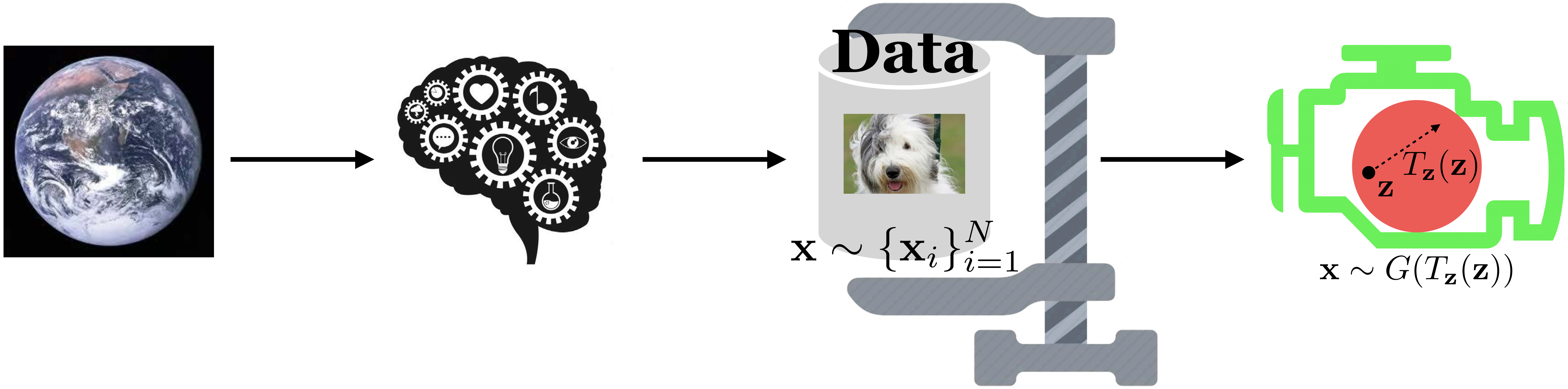
Data is static, an IGM is not!

$$\mathbf{x} \sim \{\mathbf{x}_i\}_{i=1}^{N}$$

Generative Model



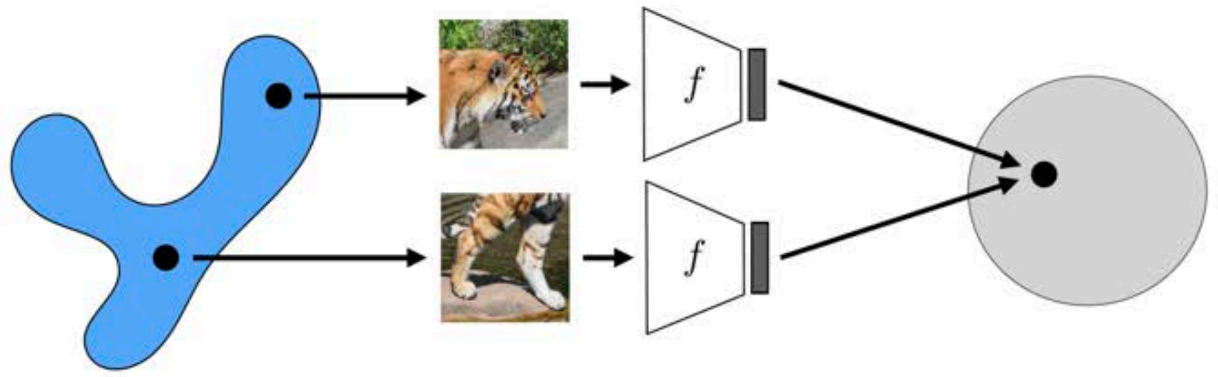
IGM as a compressed and organized copy of data



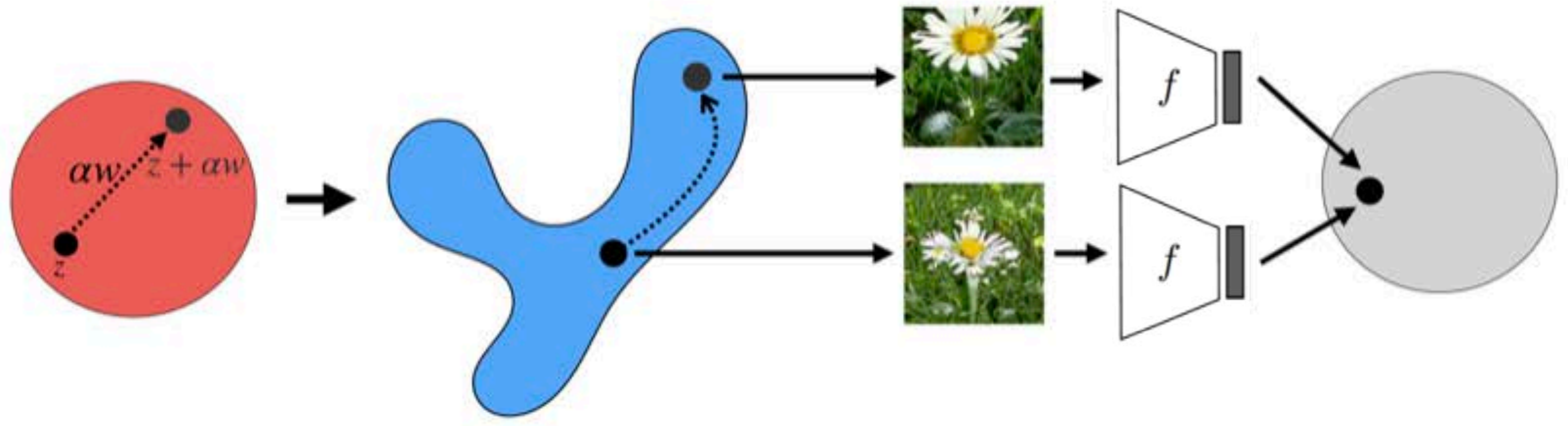
Datasets become increasingly unwieldy, missing, or private?

Steerability meets contrastive learning

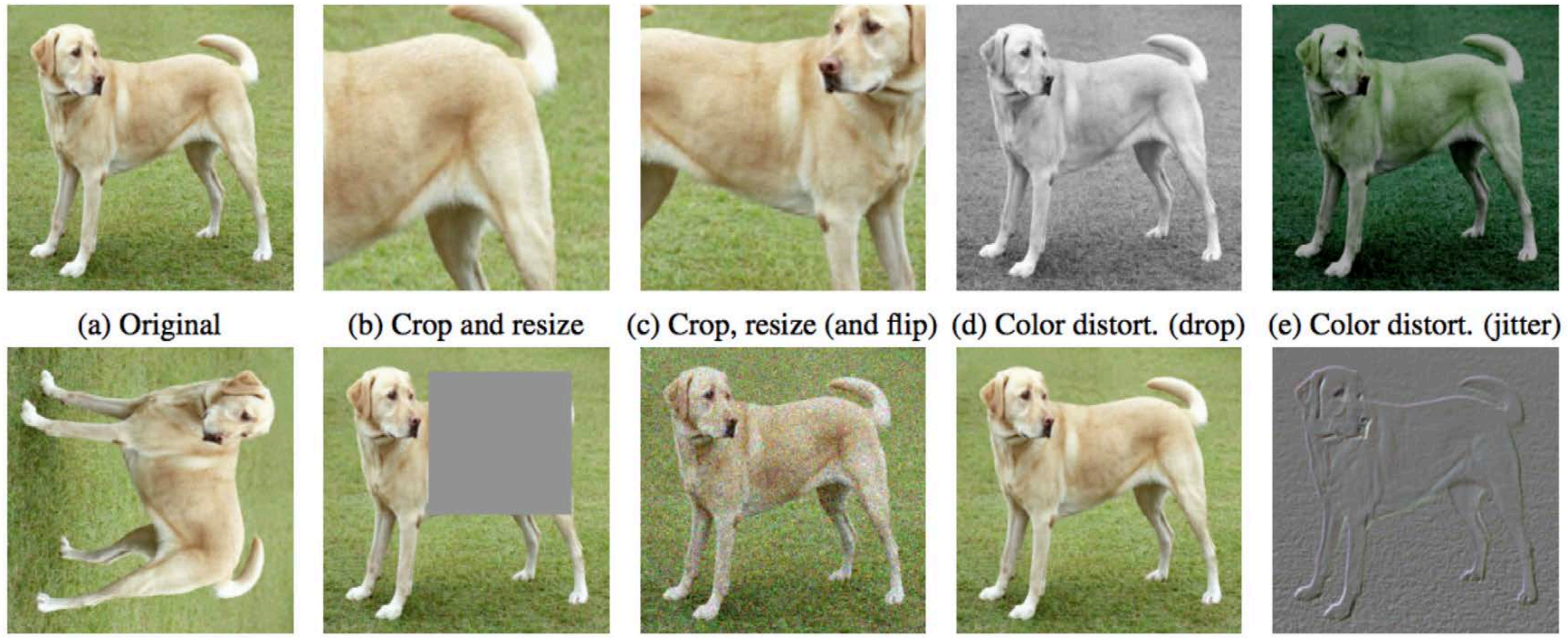
Real Dataset



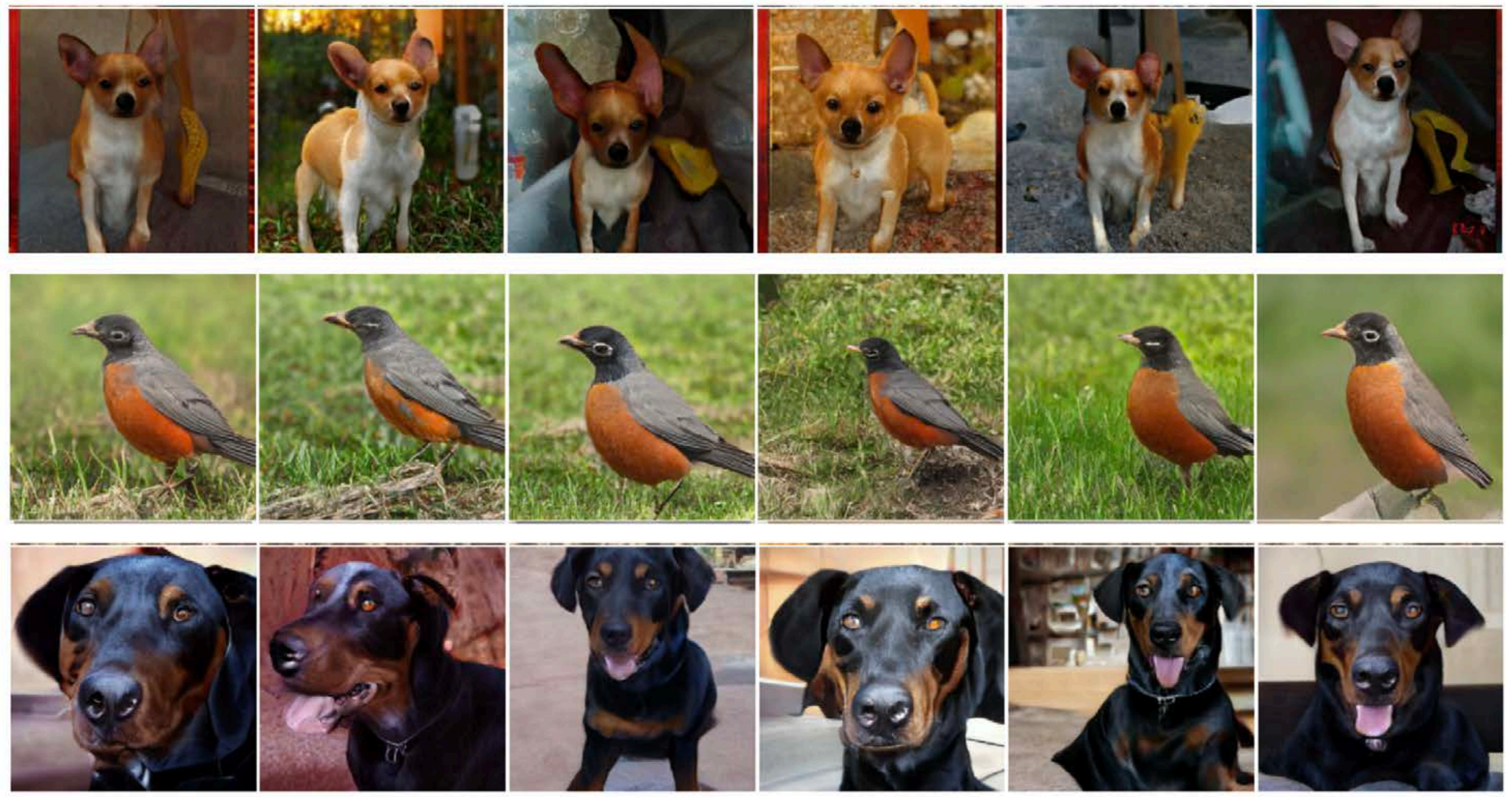
IGMs



These are all different views of the same thing

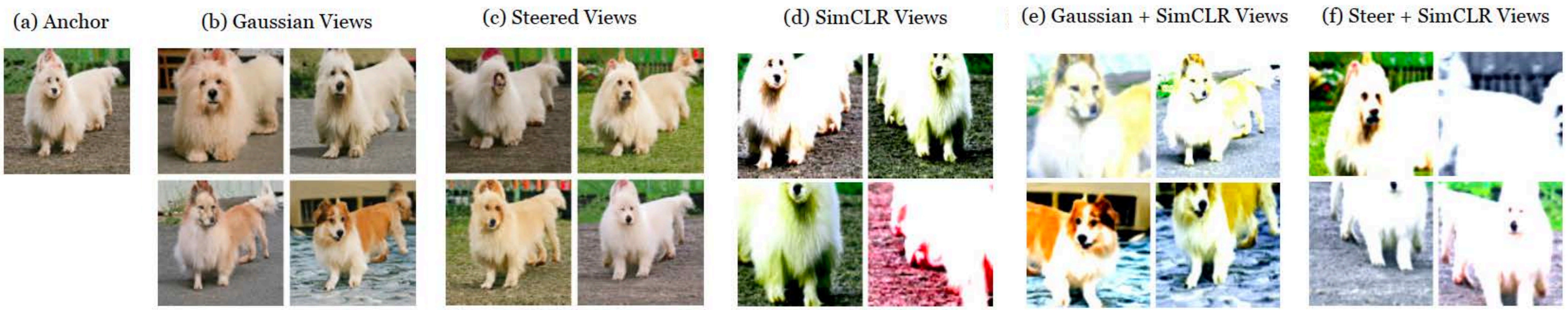


These are all different views of the same thing



[SimCLR, Chen et al. 2020]

Examples of different transformation methods



Unconditional IGM data

Conditional IGM data

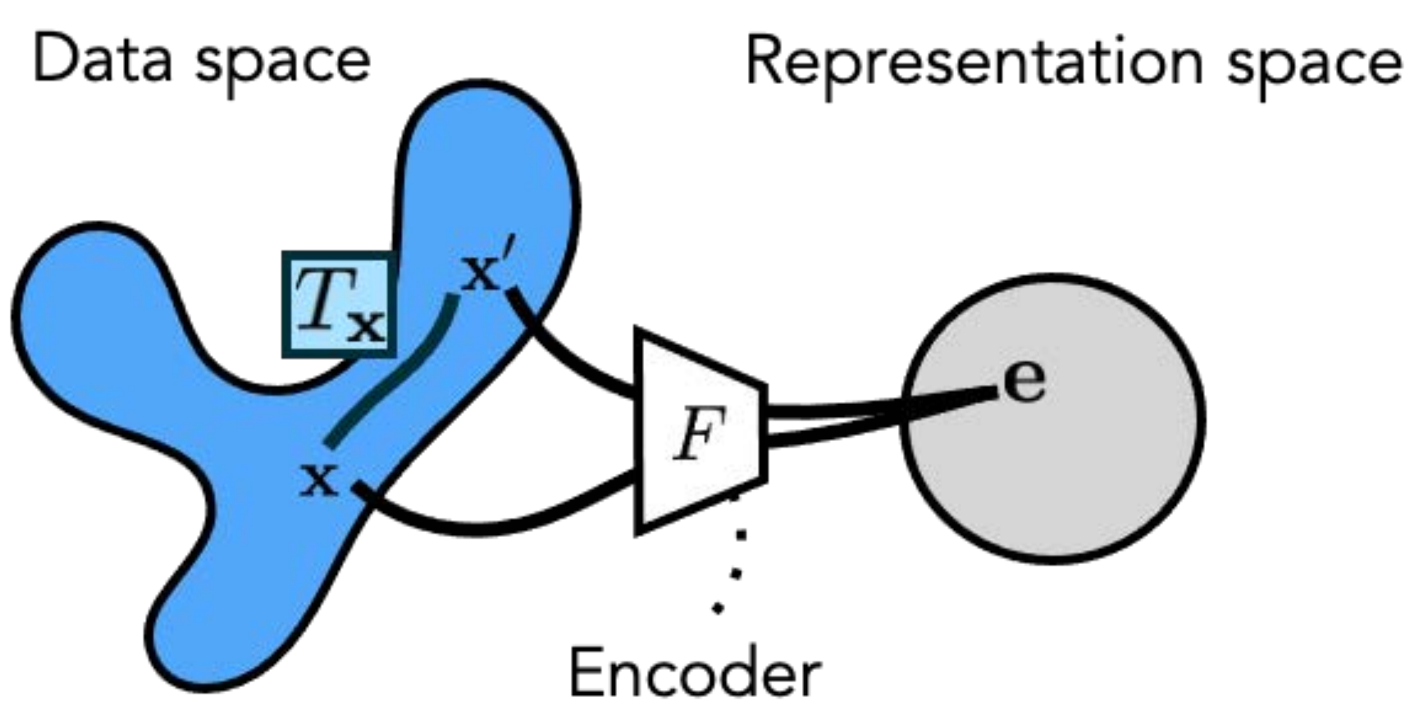
Contrastive learning + Generative modeling

Contrastive learning from real data (SimCLR, etc)

Given: $\{\mathbf{x}_i\}_{i=1}^{N}$, $T_{\mathbf{x}}$

Dataset

Learn: F



Top-1 Accuracy on ImageNet1000 linear transfer

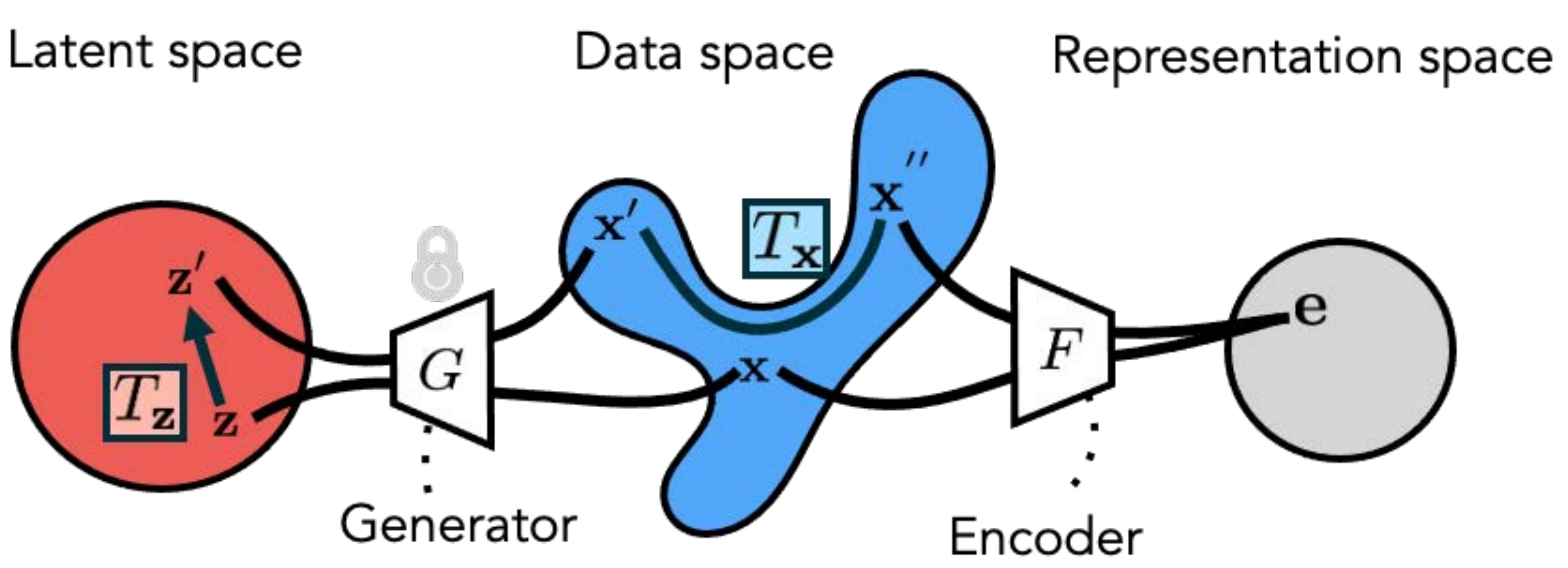
43.9%

Contrastive learning from generated data

Given: G , $T_{\mathbf{z}}$, $T_{\mathbf{x}}$

Generator

Learn: F



Only $T_{\mathbf{x}}$ $\longrightarrow$ 35.7%

$T_{\mathbf{z}}$ + $T_{\mathbf{x}}$ $\longrightarrow$ 42.6%

What if we keep generating data?

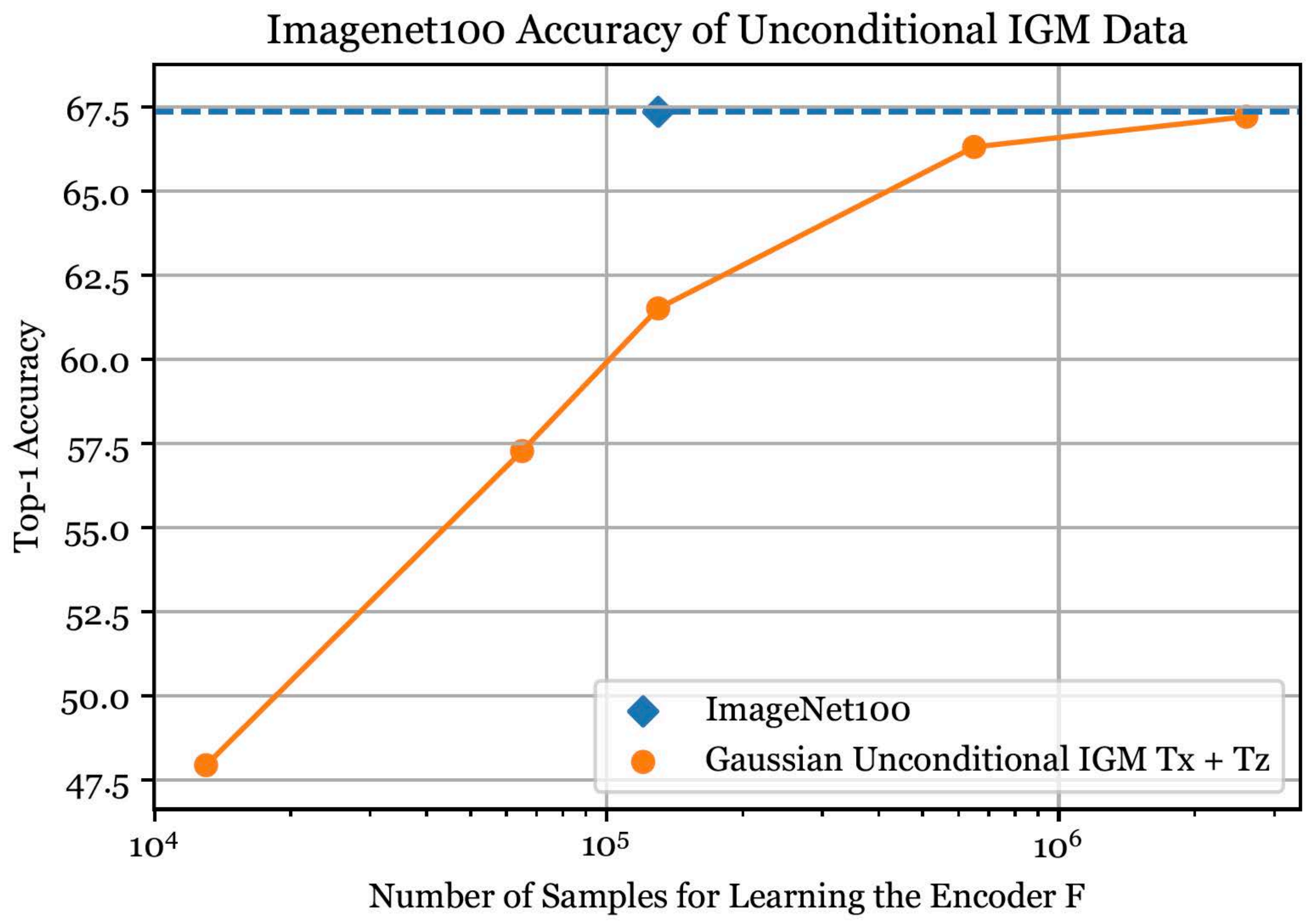
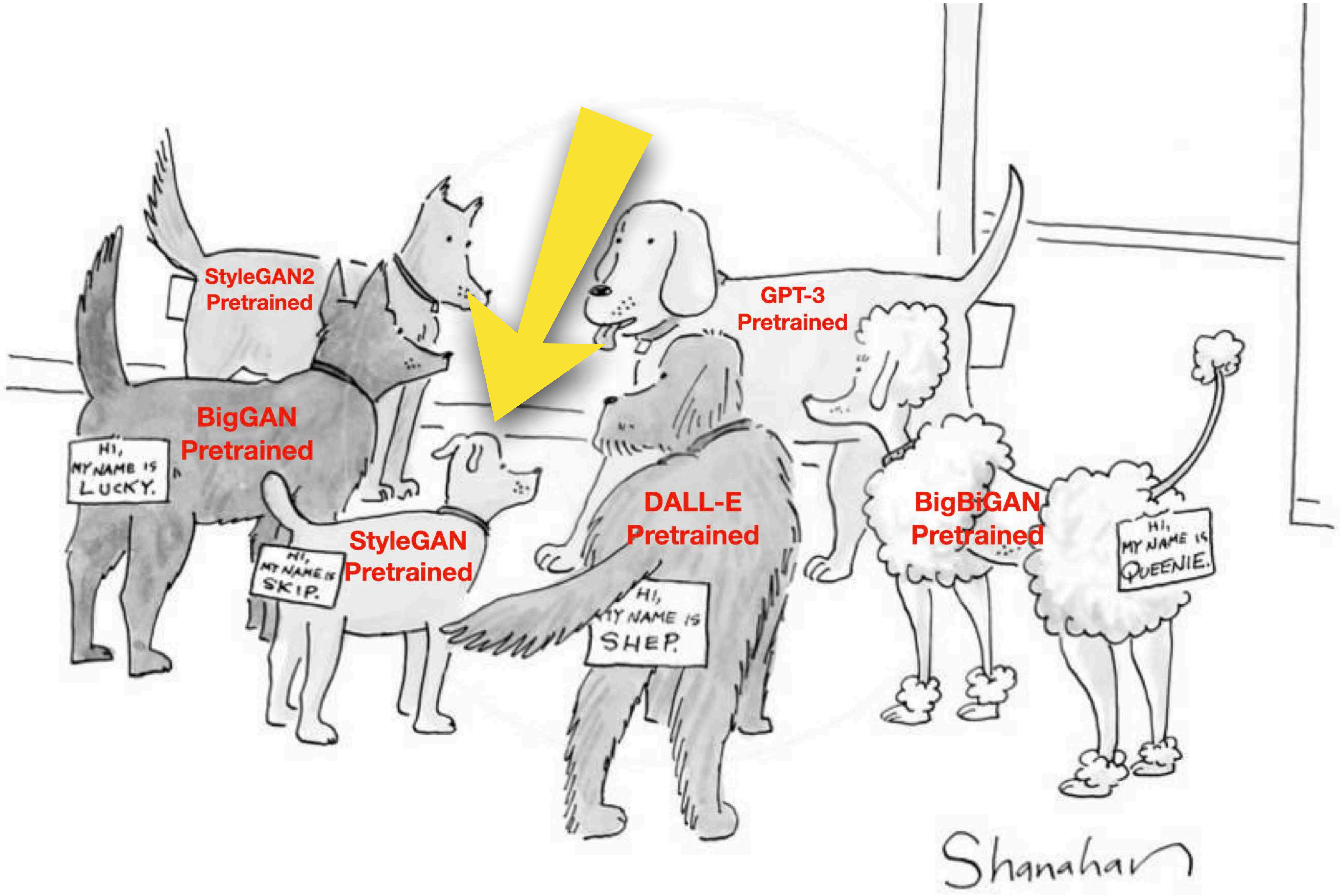


Figure 6: Effect of the number of samples for training the representation learner, evaluated using linear transfer to ImageNet100. "Gaussian" refers to the Gaussian views ($T_{\mathbf{z}} = \mathbf{z} + \mathbf{w}_{\texttt{Gauss}}$).

The performance increases with more samples both in the class-conditional and unconditional IGMs, but sub-logarithmically.

Are all types of IGMs the same?



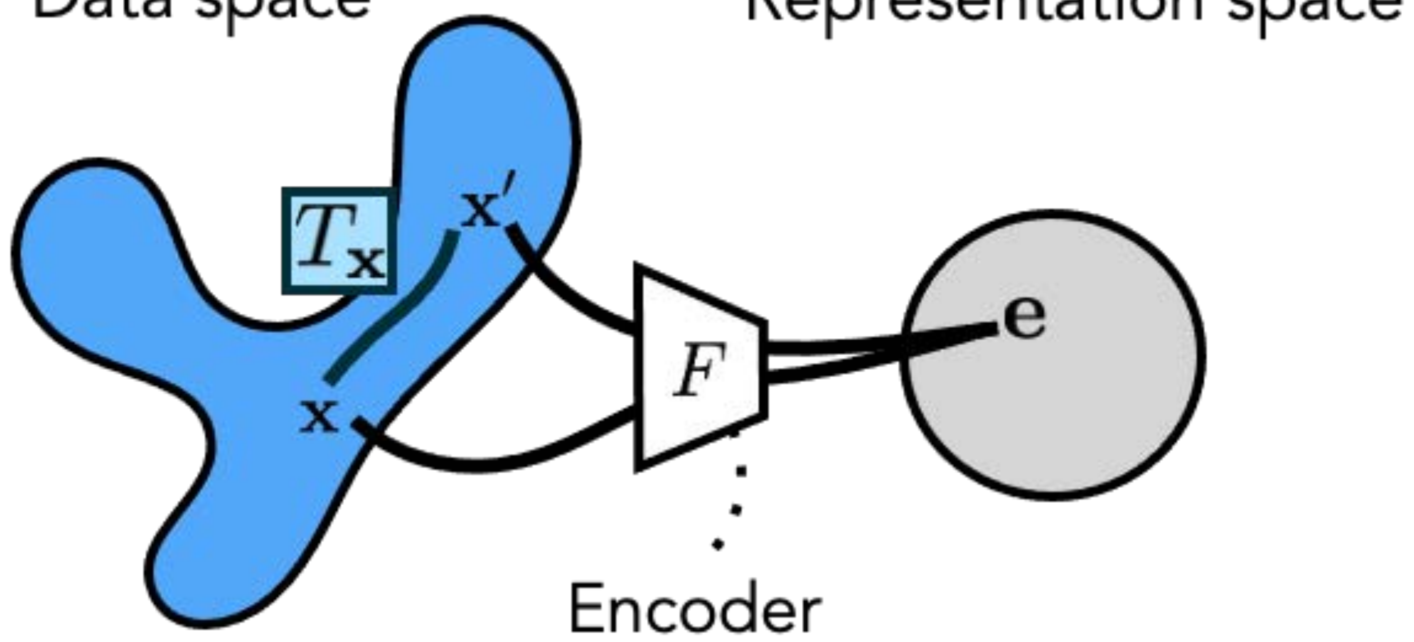
Contrastive learning + Generative modeling

Contrastive learning from real data (SimCLR, etc)

Given: $\{\mathbf{x}_i\}_{i=1}^{N}$, $T_\mathbf{x}$

Dataset

Learn: F



Top-1 Accuracy on Stanford Car Classification linear transfer

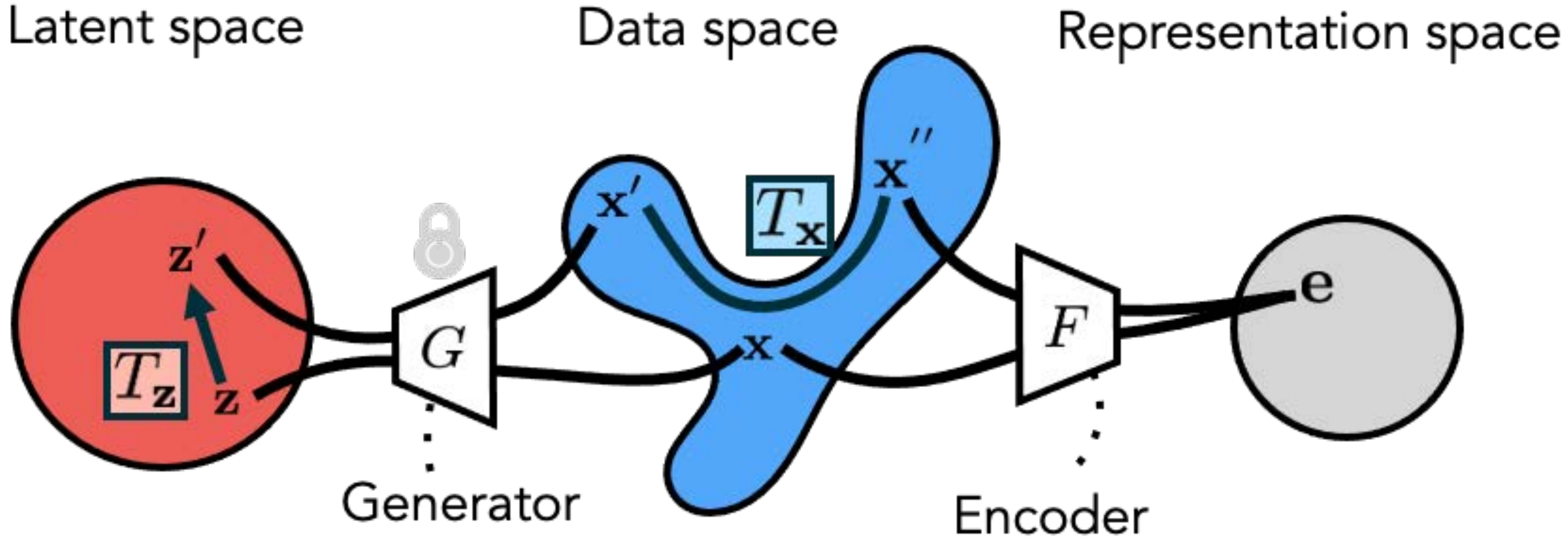
39.7%

Contrastive learning from generated data

Given: G , $T_\mathbf{z}$, $T_\mathbf{x}$

Generator

Learn: F



Only $T_\mathbf{x}$ → 40.9%

$T_\mathbf{z}$ + $T_\mathbf{x}$ → **49.9%**

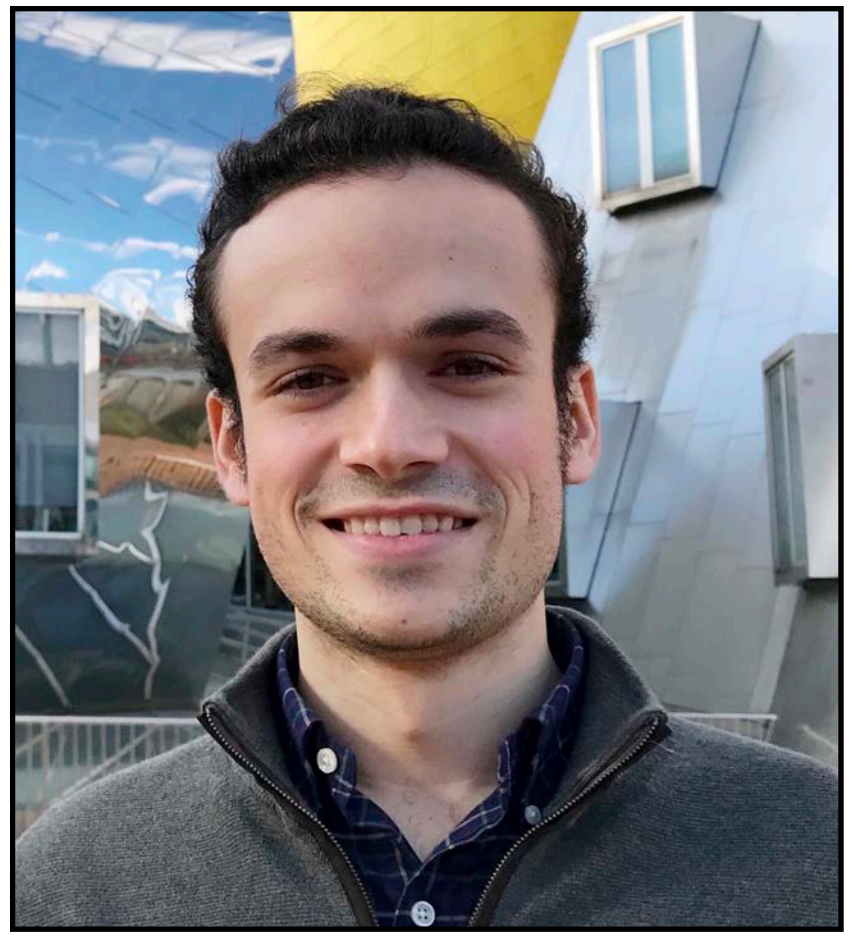
Conclusion

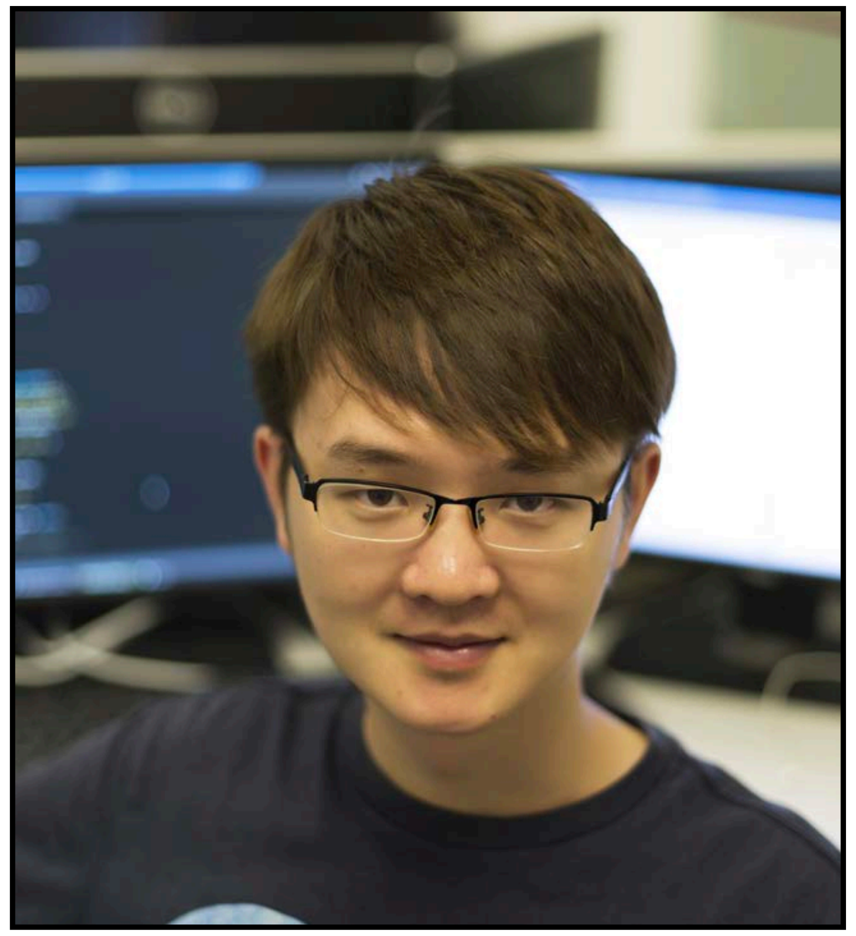
- Contrastive learning algorithms meet steerability
    - Learn “views” of the data in the model’s latent space, followed by those in the image space
- Simple Gaussian transformations worked best among the methods
- Generative models can produce *endless** samples
    - With enough samples, we achieve performance that rivals learning directly from real data
- Learning representations from generative models can outperform real data if the generative model is of **sufficiently quality, e.g. StyleGAN Car**

Find more on the project website!

https://ali-design.github.io/GenRep/


Ali Jahanian


Xavier Puig


Yonglong Tian


Phillip Isola