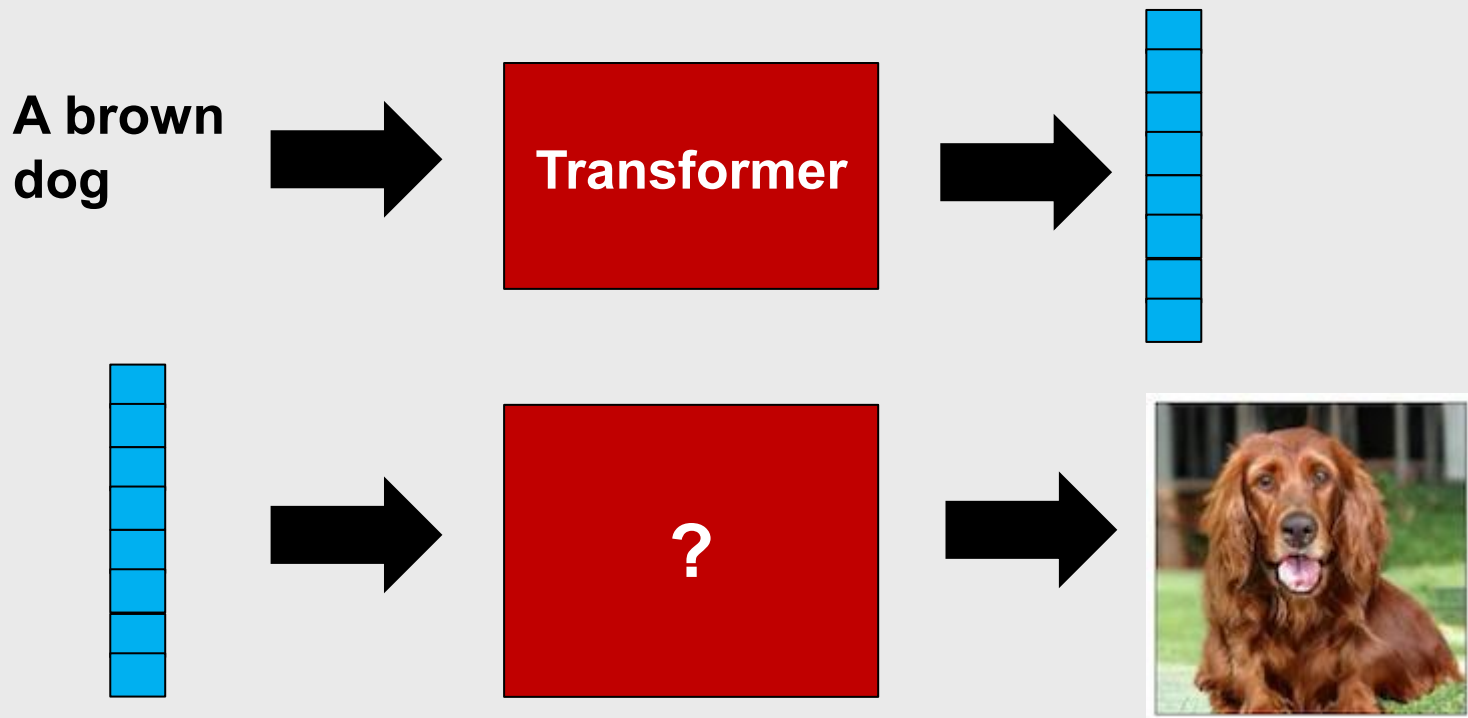**AI Image Generation**

Generating Images From Text

- **We have a way to map text to a vector (transformer)**
- **How do we map the vector to an image?**

Diffusion Models

- **Diffusion is the random motion of particles**
- **Diffusion models create images by letting vectors "diffuse" backwards in time**

Diffusion Paper (2015)

Deep Unsupervised Learning using Nonequilibrium Thermodynamics

Jascha Sohl-Dickstein   JASCHA@STANFORD.EDU
Stanford University

Eric A. Weiss   EAWEISS@BERKELEY.EDU
University of California, Berkeley

Niru Maheswaranathan   NIRUM@STANFORD.EDU
Stanford University

Surya Ganguli   SGANGULI@STANFORD.EDU
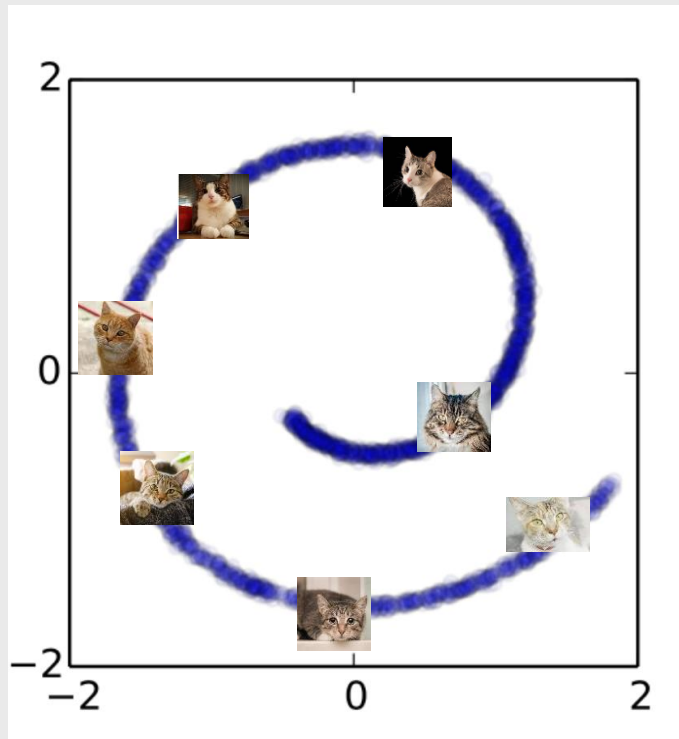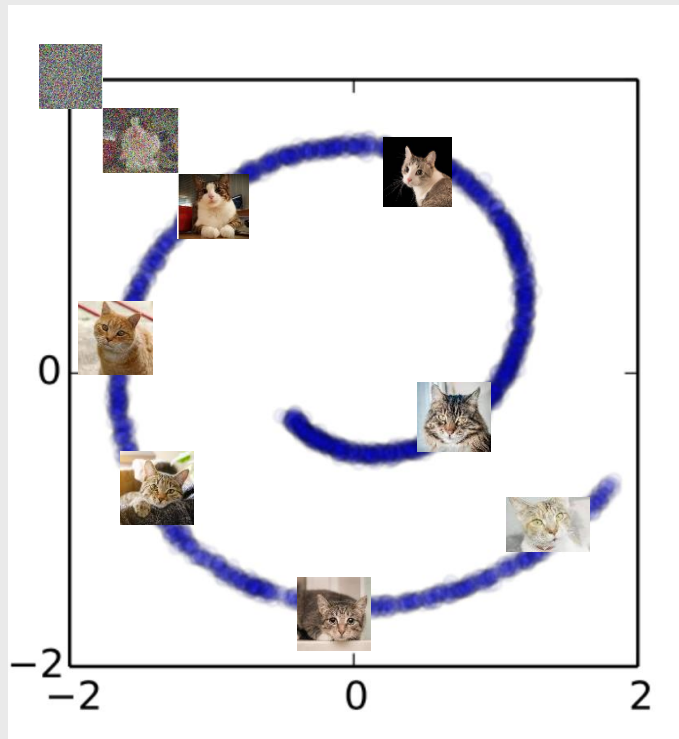Stanford University

Diffusion Example

- **Imagine our data points start off in a specific shape**
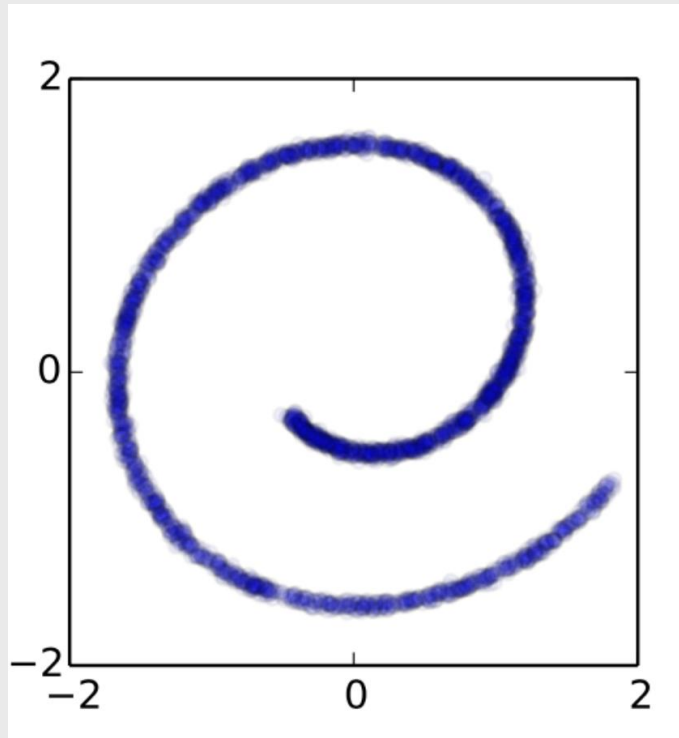
Diffusion Example

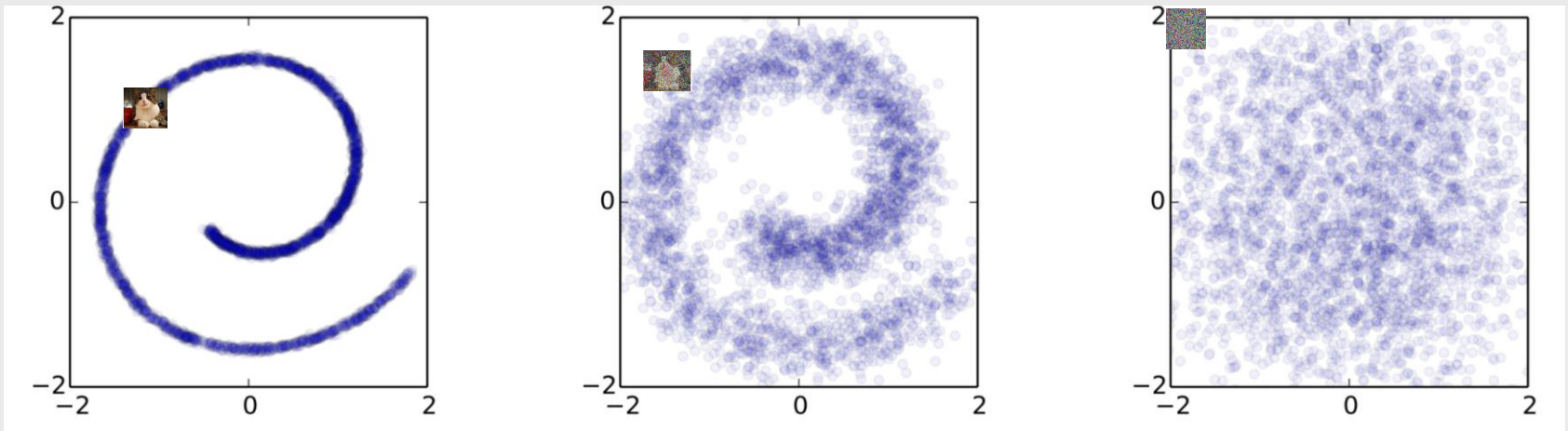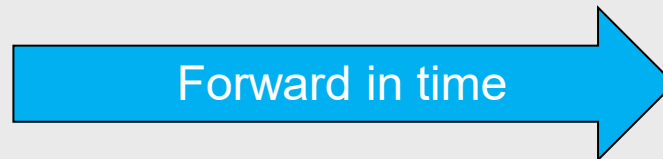- **Imagine our data points start off in a specific shape**

Diffusion Example

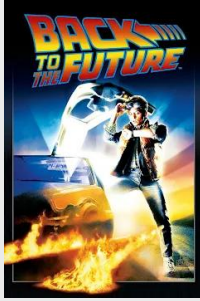- **Imagine our data points start off in a specific shape**

Forward Diffusion

- **In the forward direction, the drift velocity is zero always**

- **We end up with a uniform cloud of points**

Forward in time

Backward Diffusion

- **In the backward direction, the drift velocity is depends on the time and location of the data point**

- **We end up with the original data point configuration**

Backward in time
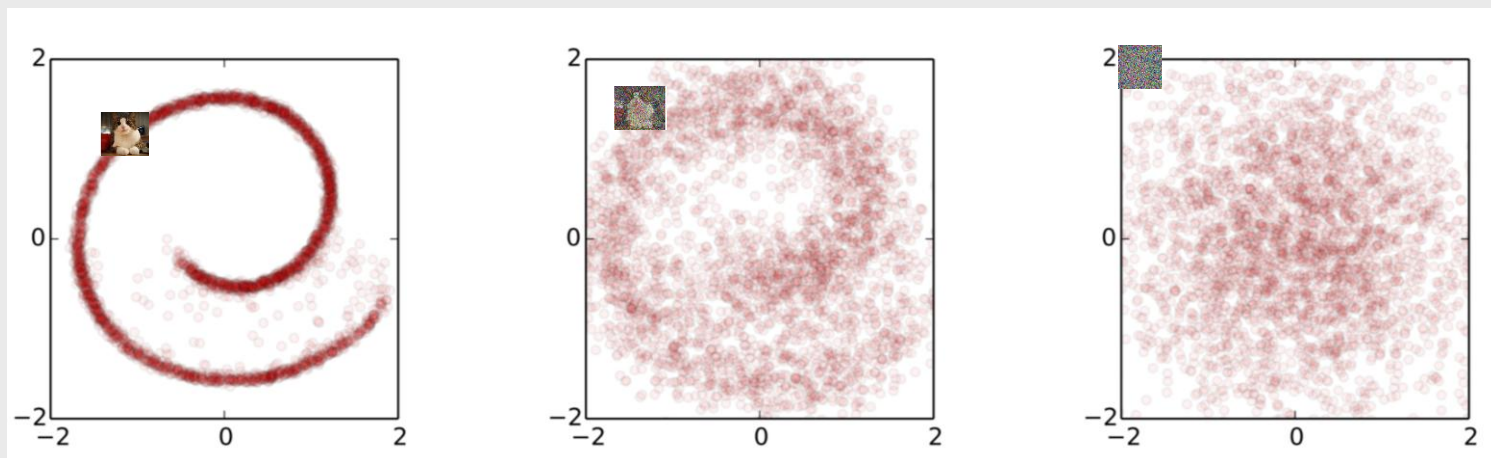
Backward Drift Velocity

- **Initially, backward drift velocity is almost zero – no rush to go anywhere**

- **Near the end, the velocity can become very strong – no data points in certain regions**
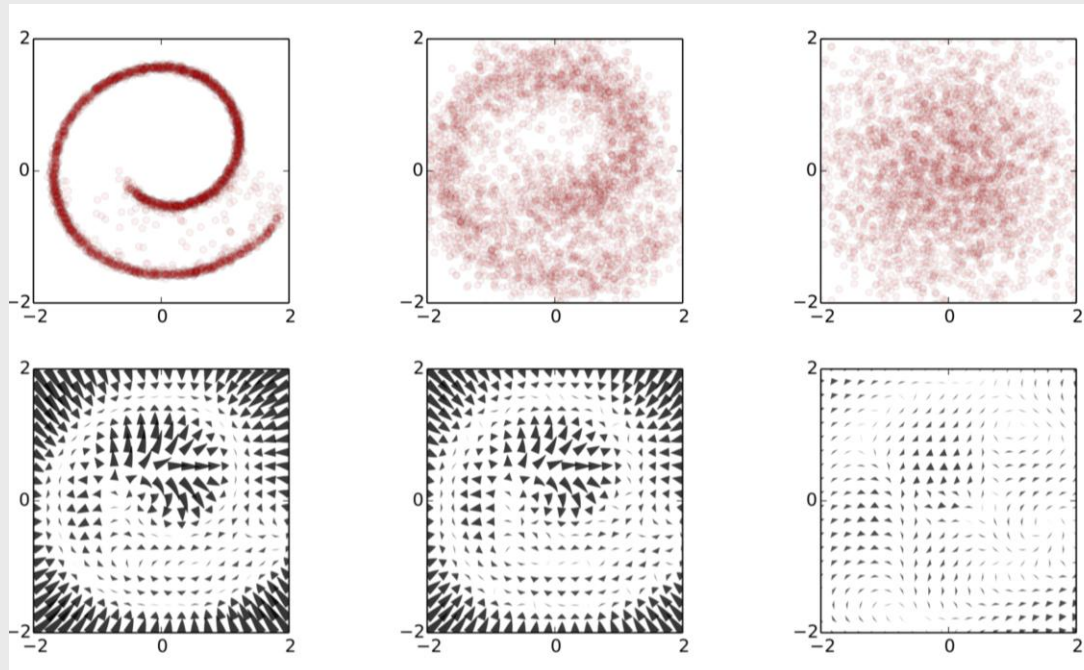
Latent Diffusion Models

- **The backward drift velocity is the key to a latent diffusion model**

- **We can use a neural network to model the drift velocity**
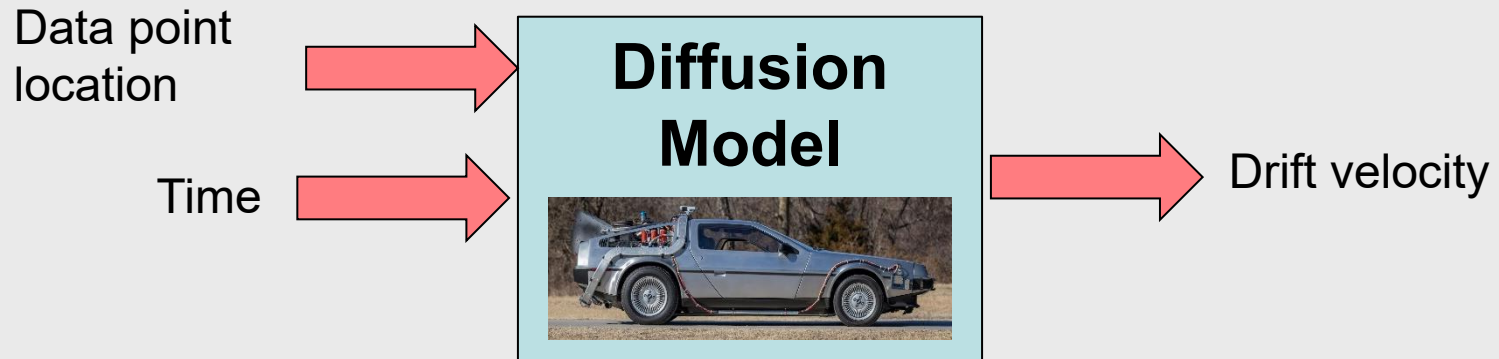
Diffusion Model Image Generation

- **Data points can be images**

- **The diffusion model will drift pixel colors in noisy image until they look like something from your training data**
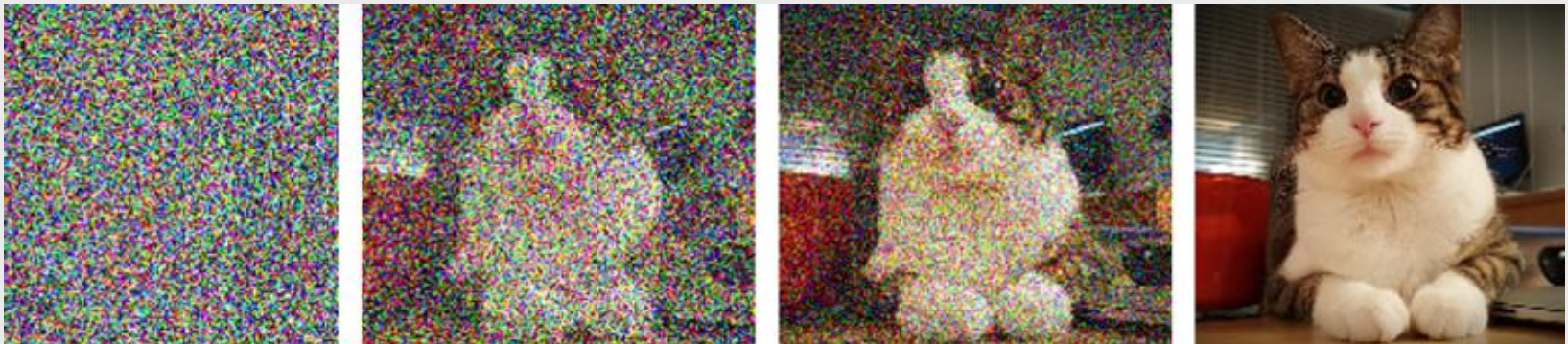
First Diffusion Model Generated Images (2015)
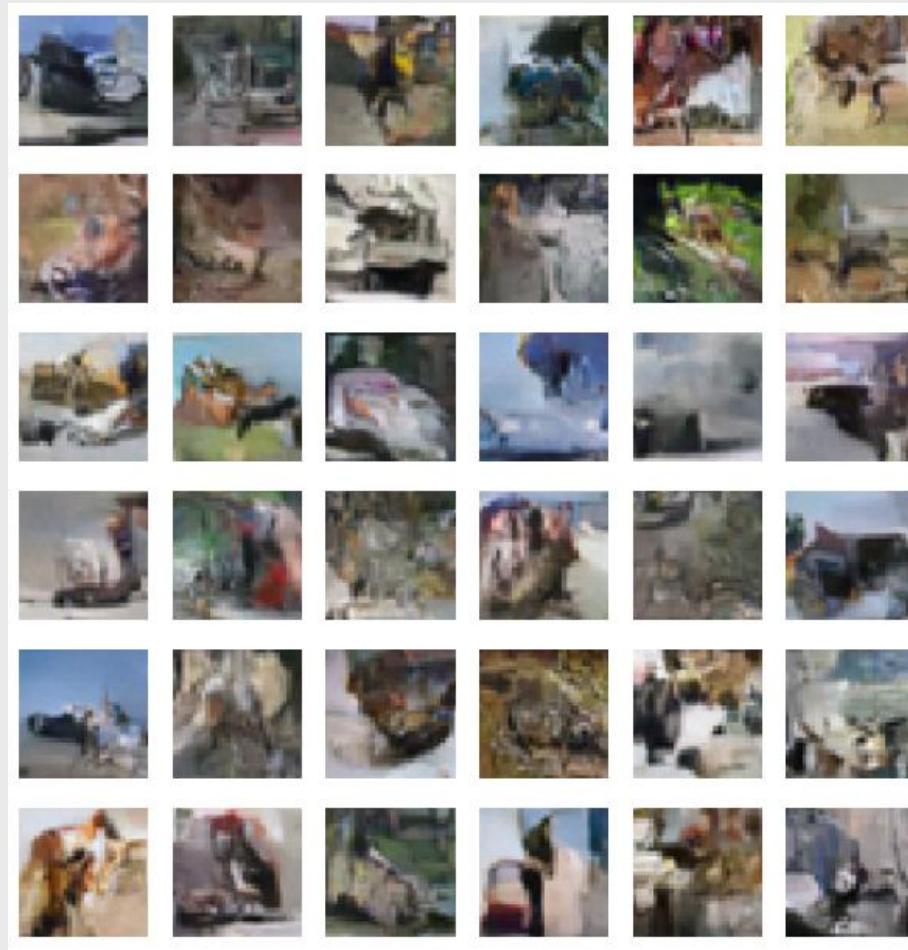
Latent Diffusion Models with Text

- **Diffusion model gives us drift velocity**
- **We can use a text image description to guide the diffusion**
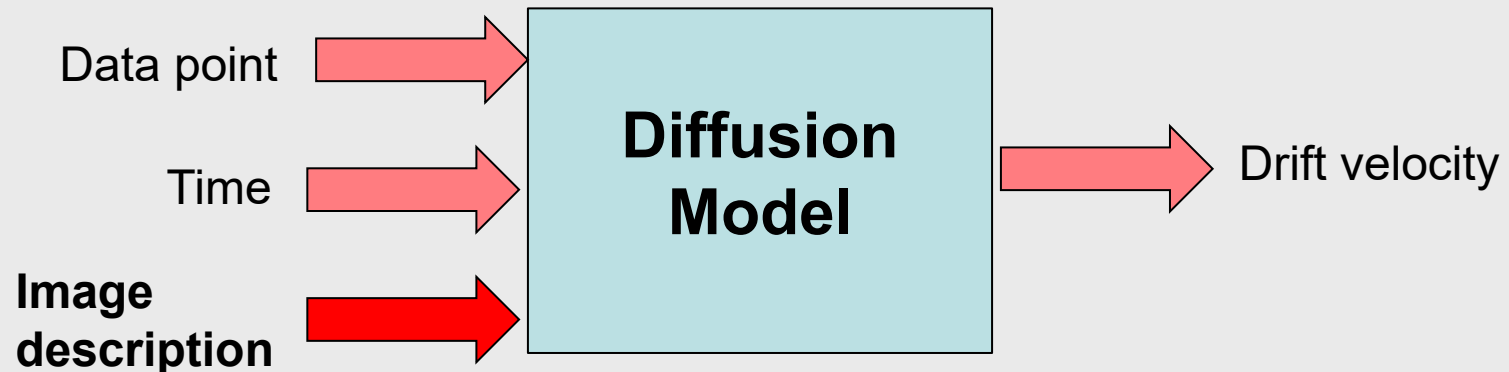
DALLE-2 (2022)

- **DALLE-2 puts text and image embeddings in the same space**

Hierarchical Text-Conditional Image Generation with CLIP Latents

Aditya Ramesh[*]
OpenAI
aramesh@openai.com

Prafulla Dhariwal[*]
OpenAI
prafulla@openai.com

Alex Nichol[*]
OpenAI
alex@openai.com

Casey Chu[*]
OpenAI
casey@openai.com

Mark Chen
OpenAI
mark@openai.com

DALL-E 2 Architecture

- **DALL-E 2 uses a diffusion model to turn a text vector into an image vector**

- **It then uses another diffusion model to decode the image vector into an image**
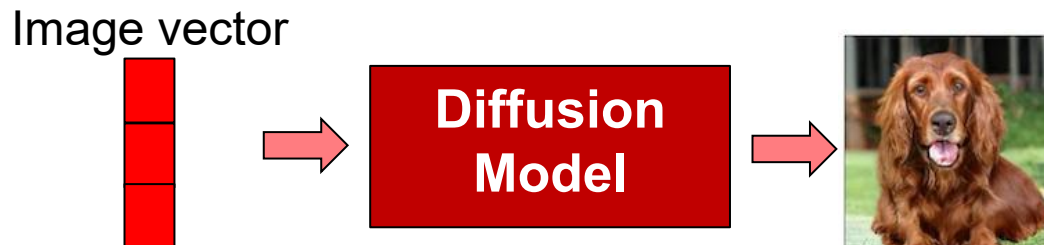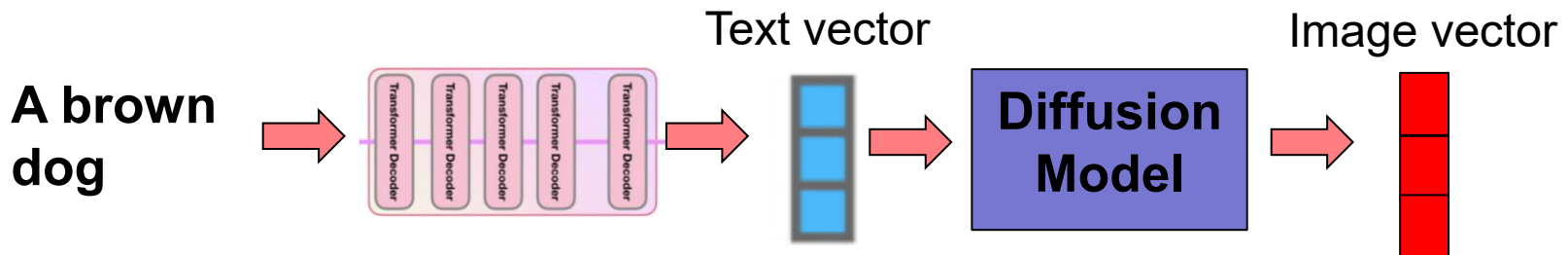
DALLE-2 Examples
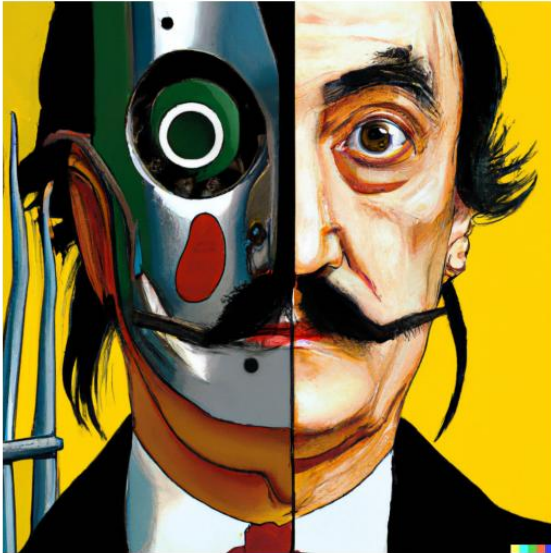
vibrant portrait painting of Salvador Dalí with a robotic half face

a shiba inu wearing a beret and black turtleneck

a close up of a handpalm with leaves growing from it

AI Image Generators

- **Diffusion models + transformers are now used by nearly all AI image generators**

- **Many AI image generators available today**
  - **ChatGPT from OpenAI**
  - **Nano Banana and Imagen (Gemini) from Google**
  - **Midjourney from Midjourney**
  - **Stable Diffusion from Stability AI**
  - **Firefly from Adobe**

1. https://zapier.com/blog/best-ai-image-generator

AI Image Generators

- **Prompt: A Yale men's basketball player jumping from the free-throw line for a dunk during the championship game**

AI Image Generators (2025)

- **Prompt: A Yale men's basketball player jumping from the free-throw line for a dunk during the championship game**
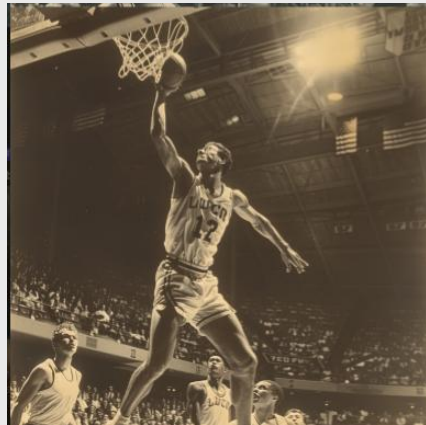
Firefly

Midjourney

ChatGPT

Gemini

AI Image Generators (2026)

- **Prompt: A Yale men's basketball player jumping from the free-throw line for a dunk during the championship game**

Firefly

?

Midjourney
(cancelled
subscription)

ChatGPT

Gemini

Google Nano Banana

- **Nano Banana is the nickname of the Google Gemini 2.5 Flash Image model**

- **We can use it to "combine" images**

- **Allows us to keep a face/product identical across image generations**

Origin of Nano Banana

- **In August 2025, Google PM Naina Raisinghani needed a codename for a blind test on LMArena**

- **She smushed her nicknames ("Naina Banana" + "Nano") together to hide Google's identity.**

- **The model ranked #1 globally, beating every established player.**

- **Google leaned into the name, turning the AI Studio buttons yellow and adding the 🍌 emoji.**

Popular Nano Banana Trends

- **Action figure trend**

- **Saree/ Bollywood trend**

Anchor Images

- **An anchor image is a high-quality reference image used to "lock" a specific identity or object into the AI's latent space**

- **Instead of relying on text descriptions (which drift), the Anchor Image provides a spatial and textural map that the model must preserve across every generation.**

- **By passing this image through the Gemini API as a reference, we maintain 1:1 facial features while changing everything else.**

Virtual Product Placement

- **Step 1: Anchor Image**
  Establish the Influencer's identity.

- **Step 2: Content Image**
  Provide the product or apparel shot.

- **Step 3: Contextual Prompting**
  Merge the two into a specific environment.

influencer.png
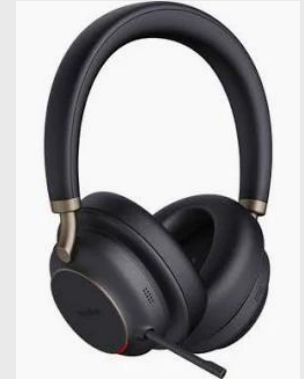
product.png

```python
# Merging Identity + Product

prompt = """

Generate the person from Image
1 interacting with the product
from Image 2. Set them in a
high-end NYC loft with natural
golden-hour lighting.

"""
```

Sample API Call

- **Order matters: the list index determines which image is "Image 1" vs "Image 2" in your prompt**

```python
# The prompt from our previous workflow slide
prompt = """
Show the person from Image 1 interacting with the
product from Image 2. Set them in a high-end NYC
loft with natural golden-hour lighting.
"""

response = client.generate_content(
    model="gemini-2.5-flash-image",
    contents=[
        prompt,
        PIL.Image.open("influencer.png"), # Assigned to Image 1
        PIL.Image.open("product.png")     # Assigned to Image 2
    ],
    config={'aspect_ratio': "9:16"}
)
```

Image Generation Cost

Metric	Flash (Nano Banana)	Pro Tier
Cost/Image	$0.039	$0.24
Latency	2-4 Seconds	12-15 Seconds
Default Size	1024 x 1024	1024 x 1024

*Free tier available via AI Studio for up to 500 images/day.

New Generative AI Landscape

- **Text guided image generation opens a wide array of possibilities**
  - Simple graphic design for any skill level
  - Visualize concepts for artistic projects
  - Create engaging visual content

- **Combined with text generation, we have even more possibilities**

Writing Books With AI

- **One can easily write books by combining text and image generation**

- **ChatGPT**
  - **Book title and outline**
  - **Chapter titles and content**
  - **Prompts for illustrations**

- **Midjourney**
  - **Illustrations**
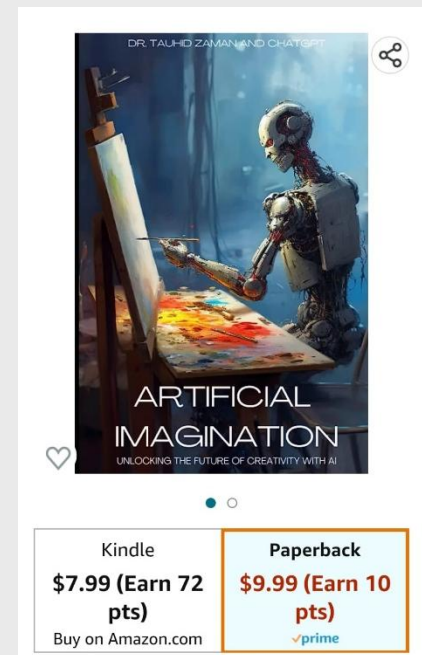
Making Static Videos

- **AI writes a script (scene and narration)**

- **Generate an image per scene (use anchor image for consistent character)**

- **TTS generates narration**

- **Combine audio and image files to make a video**

Dangers of AI Image/Video Generation + Social Media
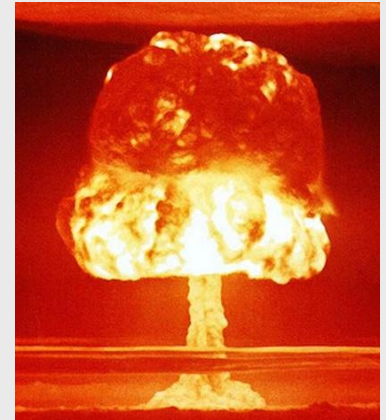
Fake Tweets and Images

- **Use AI to create fake tweets about a topic**

- **Use AI to make fake images matching tweets**

Fake Tweets and Images

- **Use AI to create fake tweets about a topic**

- **Use AI to make fake images matching tweets**

Fake tweet

Fake Tweets and Images

- **Use AI to create fake tweets about a topic**

- **Use AI to make fake images matching tweets**

Fake tweet

PatriotPaul 🇺🇸 ✔
@MAGAdefender45

Just heard the news that our LEGITIMATE 45th President, Donald J. Trump, has been ARRESTED! 😡 This is nothing but a WITCH HUNT by the Deep State and liberal elites! We will NOT stand for this disgraceful attempt to silence our voices! 🇺🇸💪 #FreeTrump #MAGA #FightForFreedom

8:06 PM · Mar 15, 2023

658 Retweets   **125** Quote Tweets   **1.3K** Likes

Fake image

Midjourney

Midjourney

"Joe biden tripping in the Oval office in front of a bunch of reporters"

Midjourney

"Joe biden tripping in the Oval office in front of a bunch of reporters"
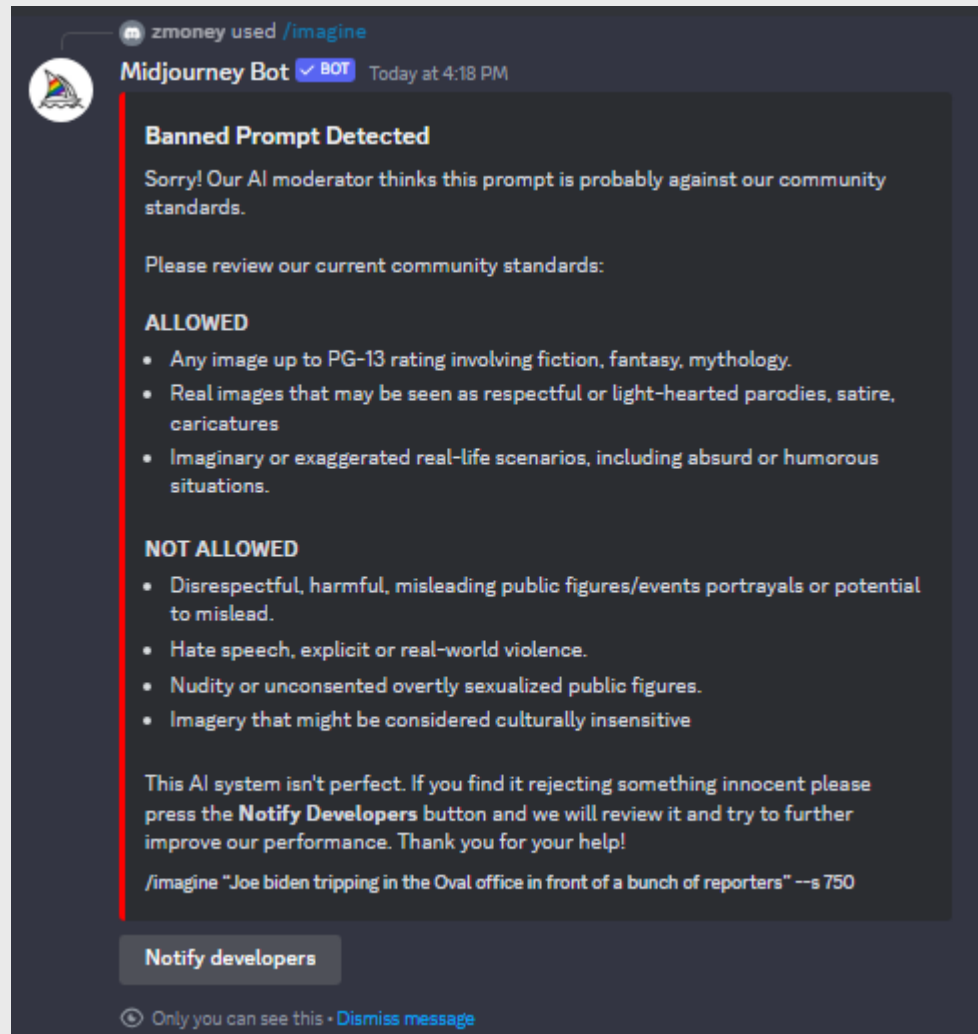
Midjourney (2023)

"Joe biden tripping in the Oval office in front of a bunch of reporters"

Midjourney (2024)

"Joe biden tripping in the Oval office in front of a bunch of reporters"

Dangerous Deep Fakes

Who Is Behind the Spread of Taylor Swift's Deepfake Nudes?

After an examination into the source of the AI-generated images that caused a firestorm online and ignited a political debate, Microsoft responds to a report that its AI image generator tool was used.

BY KEVIN DOLAK ✚   FEBRUARY 1, 2024 11:38AM

Taylor Swift   NOAM GALAI/GETTY IMAGES

Dangerous Deep Fakes

Musk's Chatbot Flooded X With Millions of Sexualized Images in Days, New Estimates Show

Over nine days, Elon Musk's Grok chatbot generated and posted 4.4 million images, of which at least 41 percent were sexualized images of women.
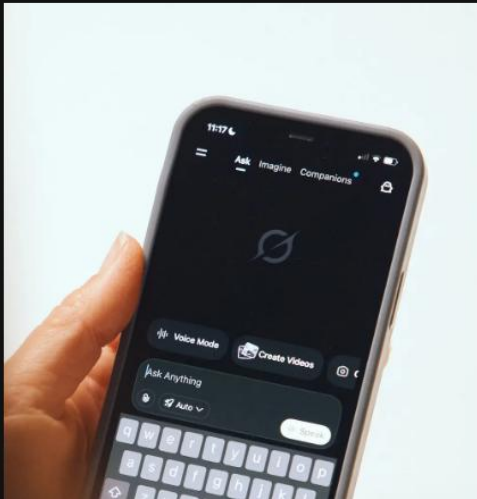
▶ Listen to this article · 6:40 min Learn more   🎁 Share full article ↗ 🔖 💬 165

Grok could have produced 3 million sexual deepfakes in 11 days, says estimate

European regulators have yet to decide how to handle the explosion of nonconsensual deepfakes on the platform.

⧉ COPY LINK   ✉ 𝕏 📱 in

POLITICOPRO   Free article usually reserved for subscribers

Deep Fake Propaganda

The official White House social media account posted a digitally altered photograph (R) of Nekima Levy Armstrong being arrested for her participation in a protest against US Immigration and Customs Enforcement (ICE) on January 22, 2025. (Original photo from Secretary of Homeland Security Kristi Noem, altered photo from the White House)

**'Dark, Bizarre Stuff': White House Posts Deepfake Image of Arrested ICE Protester Crying**

"All of us are on full notice that this White House feels no compunction about concocting obvious lies, concedes nothing when its lies are exposed, and should be presumptively disbelieved in all matters."

White House Admits to Sharing a Fake Photo of Minnesota Activist After Her Arrest

"The memes will continue," an official said in response to criticism of the altered image.

Power of Subtle Deep Fakes

- **Subtle deep fakes can have more impact than fantastic ones**
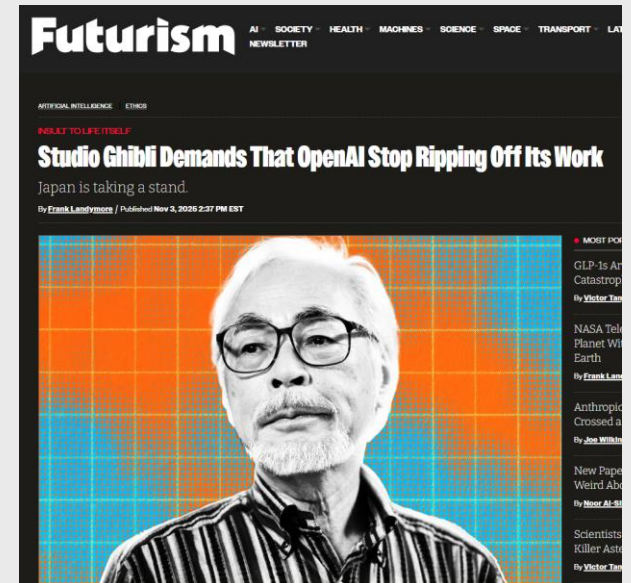- **Image is believable, but fits a desired narrative**
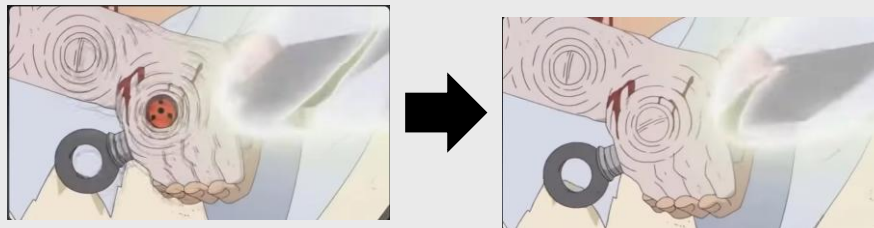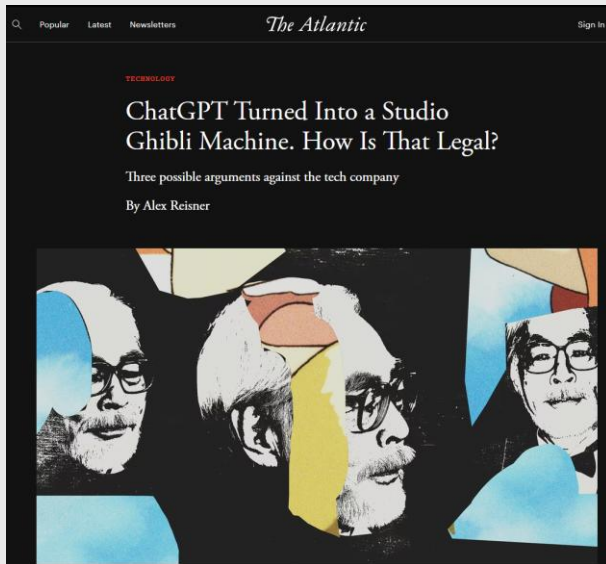
Deep Fake: Balenciaga Pope

- **Fake image of Pope in Balenciaga jacket**

- **Goes viral on social media**

- **Effective because the fake part is in the realm of reality**

GENERATED BY A.I.

Copyright Infringement

Policing AI Generated Content

UK To Bring Deepfake Law Into Force

UK government to begin enforcing law that criminalises non-consensual intimate deepfakes, amidst X probe

BY **MATTHEW BROERSMA**, JANUARY 13, 2026, 8:00 AM

Victims urge tougher action on deepfake abuse as new law comes into force

Campaigners welcome criminalisation of non-consensual AI-generated explicit images but say law does not go far enough

'I don't take no for an answer': how a small group of women changed the law on deepfake porn

Anna Moore
Thu 4 Dec 2025 05.00 EST

< **Share**

The new Data (Use and Access) Act, which criminalises intimate image abuse, is a huge victory won fast in a space where progress is often glacially slow

What is the UK's new deepfake detection plan, and how will it work?

ITV News
Fri, February 6, 2026 at 10:42 AM EST
6 min read

Add Yahoo on Google

Policing AI Generated Content

ICYMI: President Trump Signs TAKE IT DOWN Act into Law

The White House | May 19, 2025

Senate passes bill targeting nonconsensual deepfake images

Bill comes in wake of controversy over Grok AI-generated intimate pictures

Don't Be Evil

- **As AI improves, fake text, images, audio, and videos become more realistic and effective**

- **Danger is amplified when AI enters social media**

- **Please don't be evil**

Coding Session

- **We will write some code where you provide a reference image of a person and use Gemini to "lock" that identity across different scenes and prompts**

- **Allows us to make perfect persistent identity for social media storytelling and product promotion**