

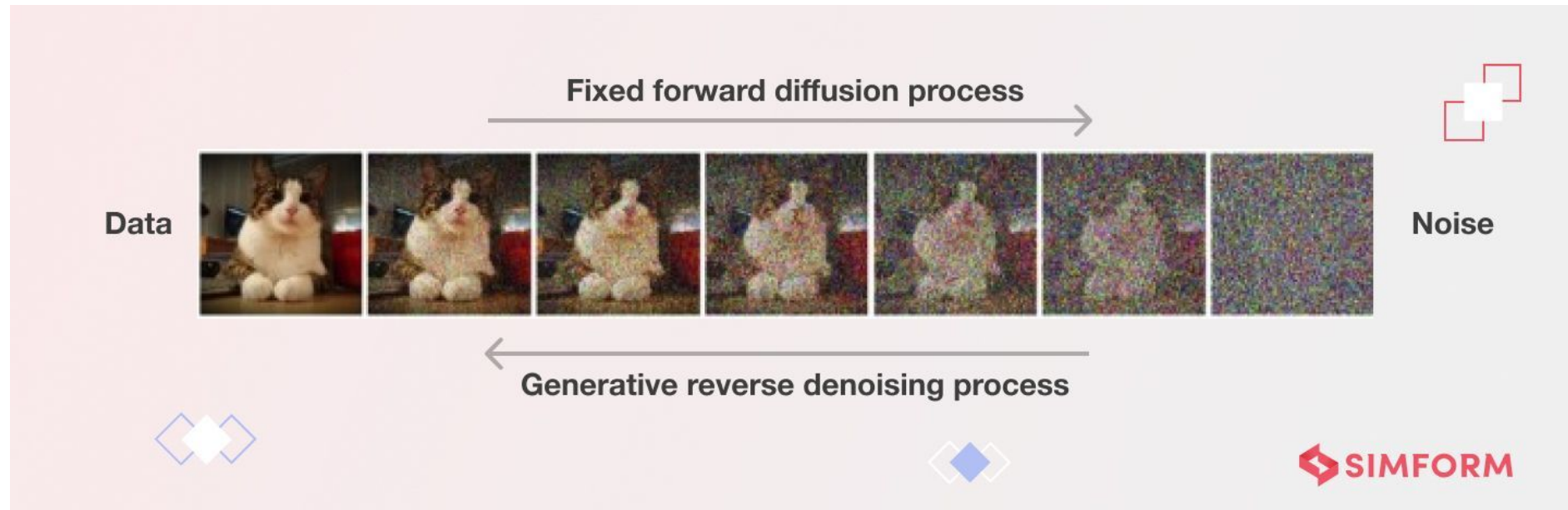
# AI Engineering

Clase 12 - Introducción a Imágenes

# Agenda

- Que son y cómo funcionan los modelos de imagenes
- Prompts de Imagenes
- APIs y tools
- Image analysis -> GPT 4o

# Diffusion



## ¿Cómo Funciona en Términos Generales?

- **Entrenamiento:** Se entrena el modelo con millones de pares texto–imagen.
- **Interpretación del Prompt:** El modelo “lee” la descripción y predice cómo debería lucir la imagen.
- **Proceso de Difusión** (en modelos como Stable Diffusion):
  - Comienza con ruido aleatorio y va refinando la imagen en sucesivas etapas.
- **Resultado:** Una imagen que coincide (en mayor o menor medida) con el texto introducido.

# Black Box

# Prompt Engineering

La calidad del prompt afecta directamente la calidad de la imagen.

Componentes Clave de un Prompt:

- **Tema/Objeto:** “Un perro labrador...”
- **Estilo:** “...en estilo de acuarela, realista, anime...”
- **Ambiente/Atmósfera:** “iluminación tenue, al atardecer...”
- **Detalles Adicionales:** “en un paisaje futurista, con tonos pastel...”



# Segmind Demo

# Enviando Imágenes a GPT-4

- ¿Por qué codificar en Base64?
  - Muchos endpoints requieren los datos en formato de texto.
  - Base64 convierte los bytes de la imagen en un string seguro para transmisiones por HTTP/JSON.

## Ejemplo de Código

```
import base64
def imagen_a_base64(ruta_imagen):
    with open(ruta_imagen, "rb") as f:
        return base64.b64encode(f.read()).decode('utf-8')
base64_img = imagen_a_base64("mi_imagen.png") print(base64_img[:100] + "...")
```





# DEMO

# PREGUNTAS

# Práctica

