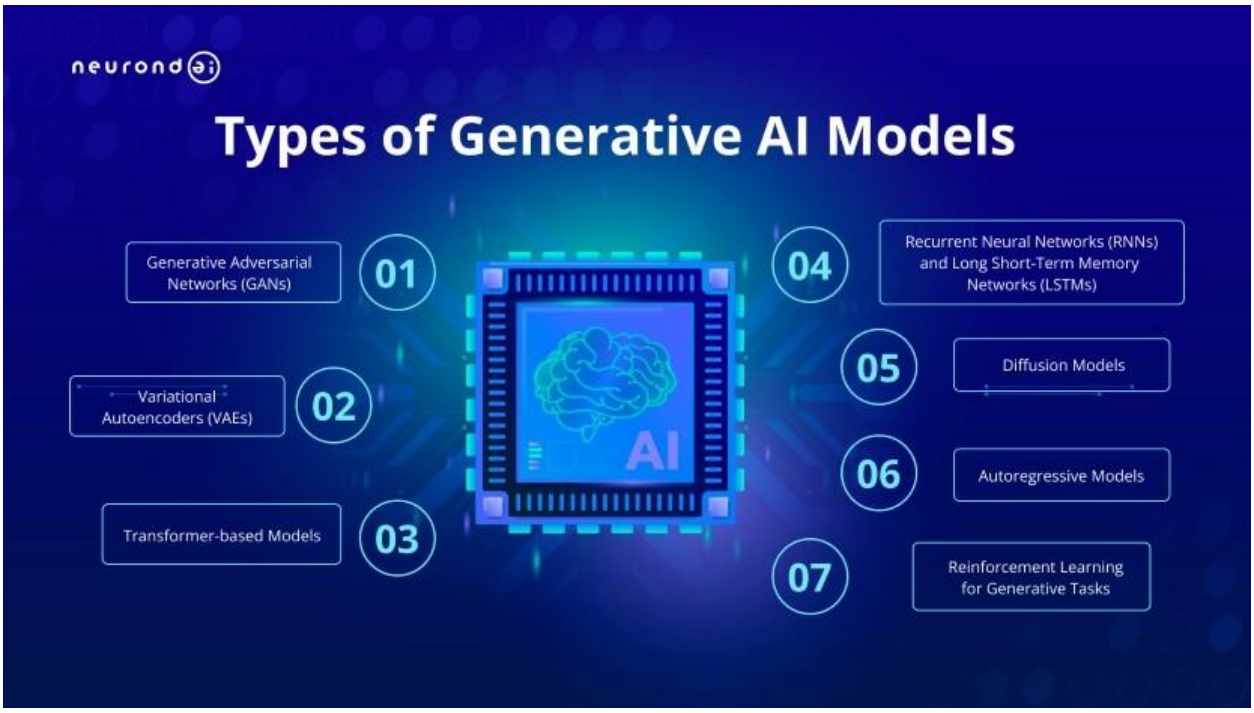


Generative AI refers to a class of artificial intelligence models capable of generating new content such as text, images, audio, video, or code. The **core components** of Generative AI can be divided into several categories, including theoretical foundations, models, data, and training techniques.

🔑 Core Components of Generative AI

1. Model Architectures

These are the backbone structures that define how generative AI learns and creates content.

Model Type	Description	Examples
Autoencoders	Learn to compress (encode) and reconstruct (decode) data	Variational Autoencoders (VAE)
GANs (Generative Adversarial Networks)	Use a generator and a discriminator in a min-max game	StyleGAN, CycleGAN
Autoregressive Models	Predict next token/element based on previous ones	GPT, PixelRNN
Diffusion Models	Gradually add noise to data and then learn to reverse the noise	DALL·E 2, Stable Diffusion

Model Type	Description	Examples
Transformers	Use attention mechanisms for sequence modeling	GPT, BERT (encoder-only), T5 (encoder-decoder)

---

2. Data

Generative AI requires large-scale, high-quality datasets for training.

- **Text data** → Books, websites, chat logs (e.g., for GPT)
- **Image data** → ImageNet, COCO, LAION (e.g., for Stable Diffusion)
- **Audio data** → LibriSpeech, Common Voice (e.g., for text-to-speech)
- **Multimodal data** → Text + Images for models like CLIP, DALL·E

---

3. Latent Space Representation

Most generative models work in a **latent space**, where high-dimensional data (e.g., images) is represented in a lower-dimensional space. This helps in:

- Capturing abstract features
- Generating variations
- Interpolation (blending between data points)

---

4. Loss Functions

Loss functions measure how well the model is generating content.

Model	Common Loss Function
GAN	Adversarial loss (Generator vs. Discriminator)
VAE	Reconstruction + KL divergence loss
Diffusion	Noise prediction loss
Transformer	Cross-entropy loss (for predicting next token)

---

5. Training Techniques

- **Supervised Learning** – Paired data (e.g., image-caption)

- **Unsupervised Learning** – Learn from raw, unlabelled data
- **Reinforcement Learning** – Used in fine-tuning (e.g., RLHF in ChatGPT)
- **Self-Supervised Learning** – Predict parts of the input from other parts

---

6. Sampling & Generation Methods

- **Greedy search / Beam search / Top-k / Top-p sampling** (for text)
- **Noise sampling and denoising** (for diffusion models)
- **Latent vector sampling** (for GANs/VAEs)

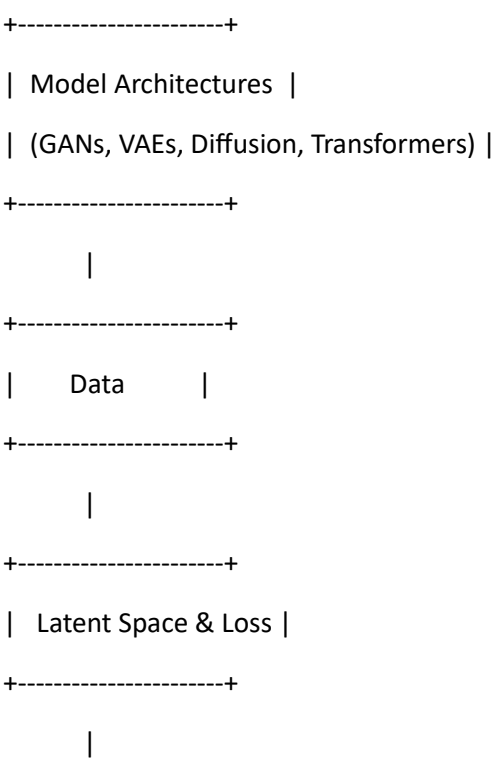
---

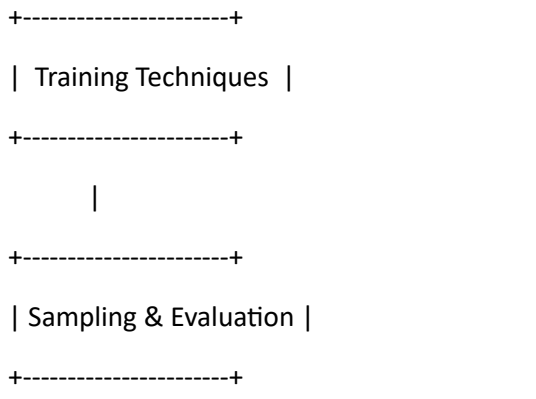
7. Evaluation Metrics

Measuring quality of generative outputs is challenging:

- **Text**: BLEU, ROUGE, perplexity, human evaluation
- **Image**: FID (Fréchet Inception Distance), IS (Inception Score)
- **Audio**: MOS (Mean Opinion Score)

📌 **Summary Diagram**




+----------+
| Training Techniques |
+----------+
|
+----------+
| Sampling & Evaluation |
+----------+