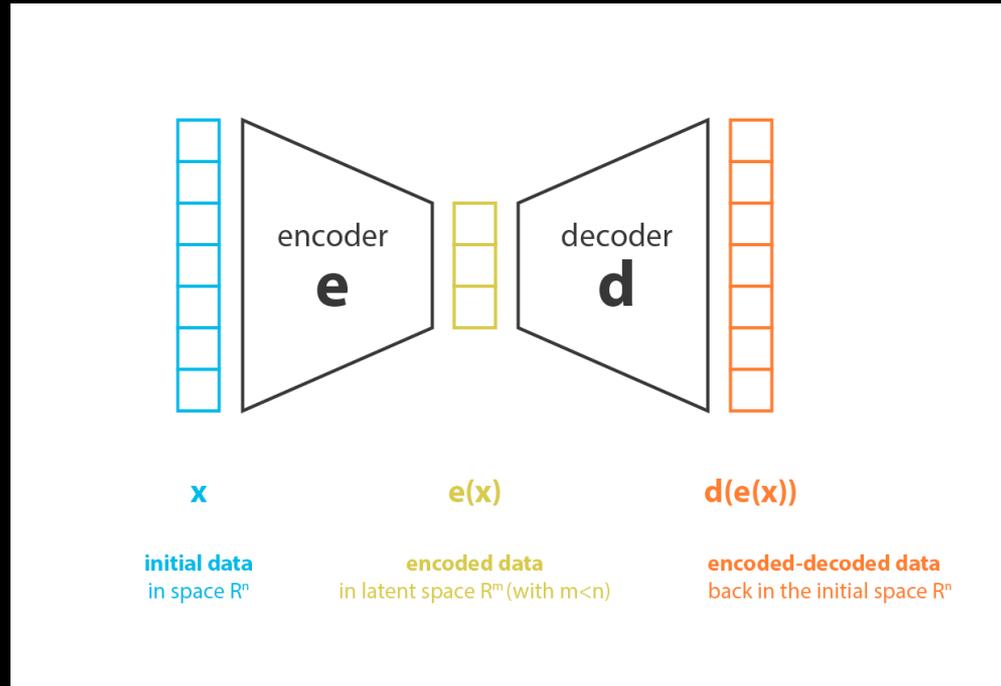
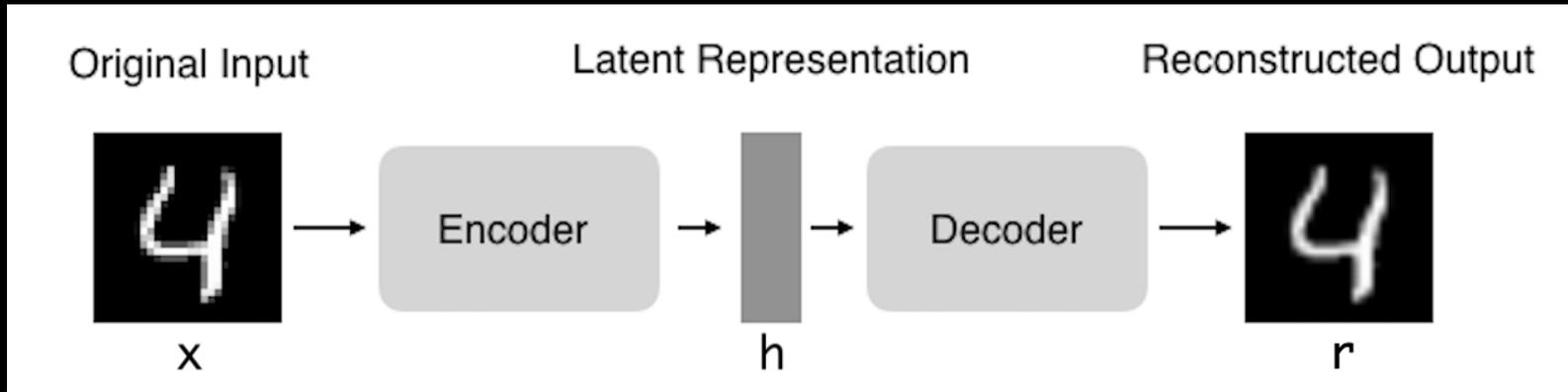


Síntesis de audio a partir de  
interpolaciones en espacios latentes

# Reducción de dimensionalidad

“Autoencoder”





Representación de datos (imágenes, audio, etc) en un espacio latente de baja dimensionalidad

-> Modelos generativos

# Imágenes



interpolación en el espacio  
de píxeles



interpolación en el espacio  
latente

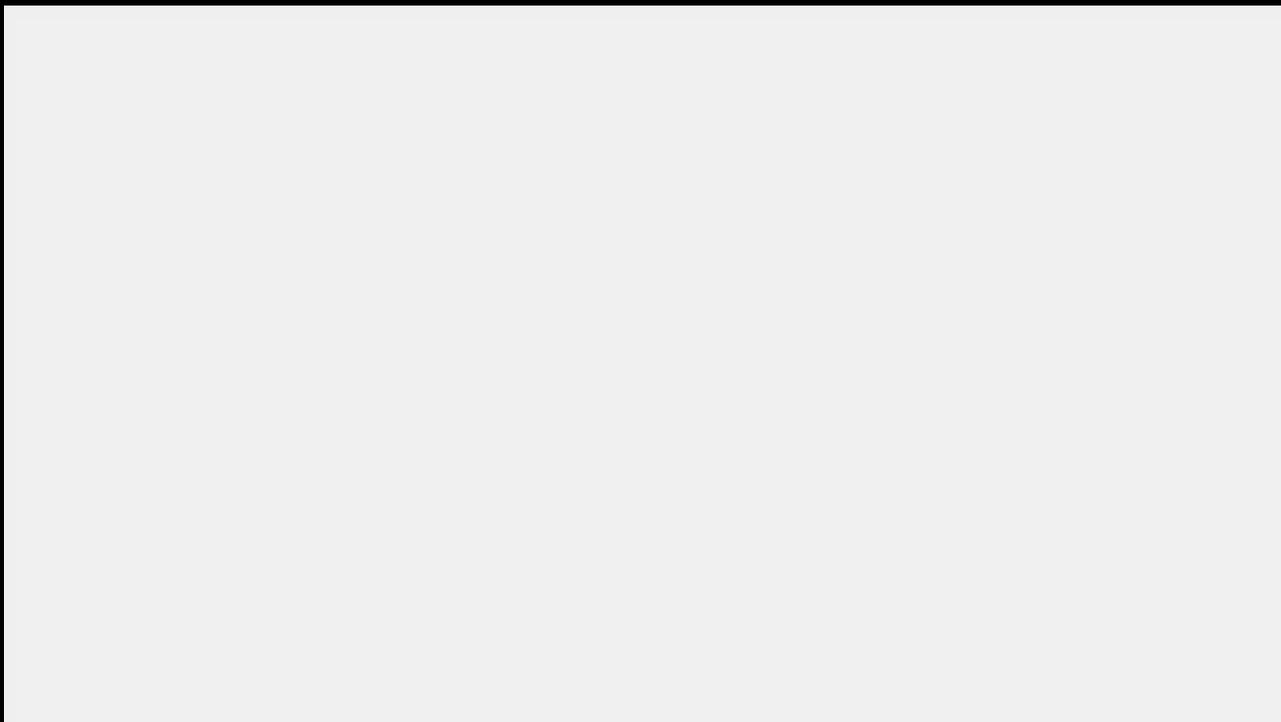
# Audios



interpolación en el espacio de audio



interpolación en el espacio latente



Interpolación latente de audio

# Algunos modelos

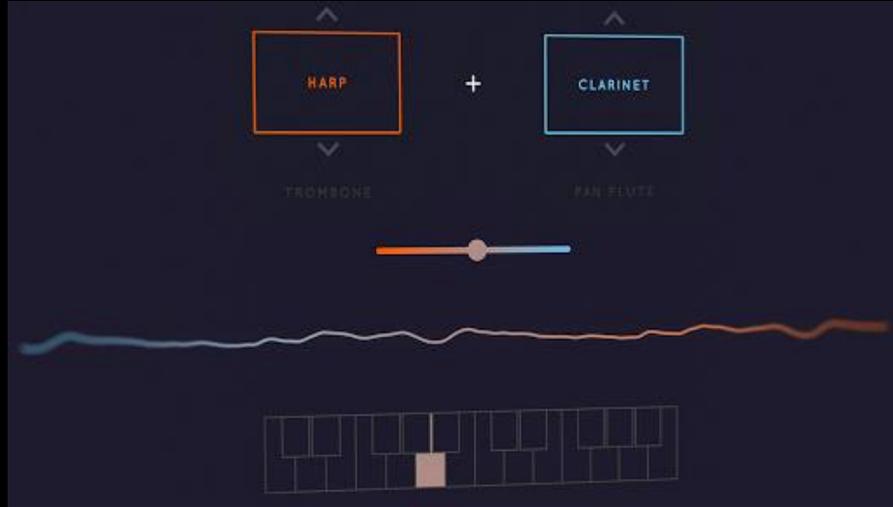
NSynth (2017) – Wavenet Autoencoder, Raw Audio

GANSynth (2019) – GAN, MEL-Spec. + IF

TiFGAN (2019) – GAN, Mag. Spec., PGHI

Latent Timbre Synthesis (2020) – VAE, CQT Spec.

# NSynth: Neural Audio Synthesis



Sound Maker

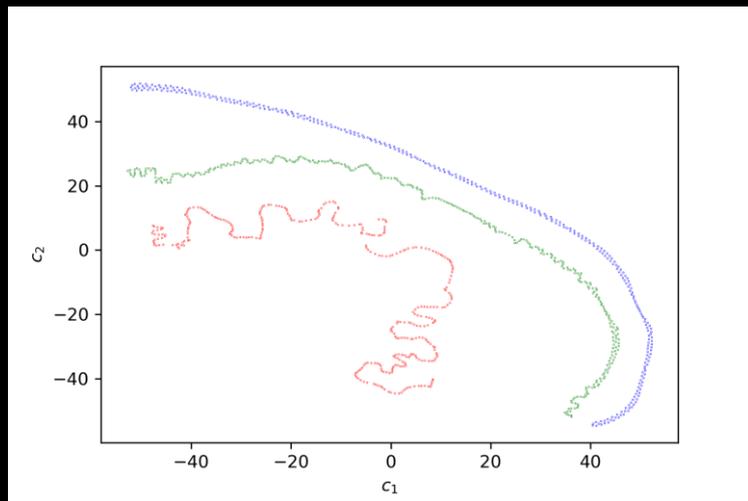


NSynth Super (RasPi + OF)



NSynth: Neural Audio Synthesis  
Ableton Instrument

# Latent Timbre Synthesis



Visualización del espacio latente



GUI Max



Audio 1



Audio 2



Interpolación (0.5)

# Modelo en desarrollo (LSS)

TwoStageVAE

Mag. Spec + PGHI

Sintetiza a 6.5-7x en PC Core I3 (aprox. 150 ms por cada segundo de audio)

Corre en Raspberry PI 4 a 1.7x



Artemiev



Ry Cooder



Fade 1



Morph 1



Morph 2