3. History of generative music

Generative Music AI

Overview

- 5 eras of generative music

- Notable examples

Pre-computer era (1700 - 1956)

Pre-computer era (1700 - 1956)

Academic era (1957 - 2009)

Pre-computer era (1700 - 1956)

Academic era (1957 - 2009)

First startup wave (2010 - 2016)

```
┌─────────────────────────────────────┐
│                                     │
│     Pre-computer era (1700 - 1956)  │
│                                     │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│                                     │
│      Academic era (1957 - 2009)     │
│                                     │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│                                     │
│    First startup wave (2010 - 2016) │
│                                     │
└─────────────────────────────────────┘
                  │
                  ▼
┌─────────────────────────────────────┐
│                                     │
│   Big tech experiments (2017 - 2022)│
│                                     │
└─────────────────────────────────────┘
```

```
Pre-computer era (1700 - 1956)
            │
            ▼
Academic era (1957 - 2009)
            │
            ▼
First startup wave (2010 - 2016)
            │
            ▼
Big tech experiments (2017 - 2022)
            │
            ▼
Music AI hype (2023 - ?)
```

Pre-computer era (1700 - 1956)

- Manual algorithms

- Composers lead the charge

- Randomness

- Re-combination

- Algorithms derive various musical parameters

Mozart Dice Game (1787)

- Random recombination

- 176 pre-composed bars

- Roll dice to select bars

Mode de valeurs et d'intensités (Messiaen, 1949)

- Parametrization of dynamics, articulation, pitch, duration

- Algorithm to select musical elements

Mode de valeurs et d'intensités (Messiaen, 1949)

Pre-computer era (1700 - 1956)

Academic era (1957 - 2009)

Academic era (1957 - 2009)

- Research activity
  - Lots of experimentation
  - Incremental advancements
  - Scattered community

Academic era (1957 - 2009)

- Research activity
  - Lots of experimentation
  - Incremental advancements
  - Scattered community

- Musical output
  - Full-piece generation is rare
  - No focus on audio production quality
  - Score generation (symbolic)
  - Classical music

Illiac Suite (Hiller & Isaacson, 1957)

- First computer-generated piece

- String quartet

- Four movements

- Different generative techniques

Illiac Suite (Hiller & Isaacson, 1957)

Lejaren Hiller - *Illiac Suite for String Quartet* (1956)

First experiment: presto, andante, allegro

Experiments in Musical Intelligence (Cope, 1981)

- Invented because of composer's block

- Generate full piece

- Recombination approach
    a. Analyse corpus
    b. Extract signatures
    c. Re-combine

Experiments in Musical Intelligence (Cope, 1981)
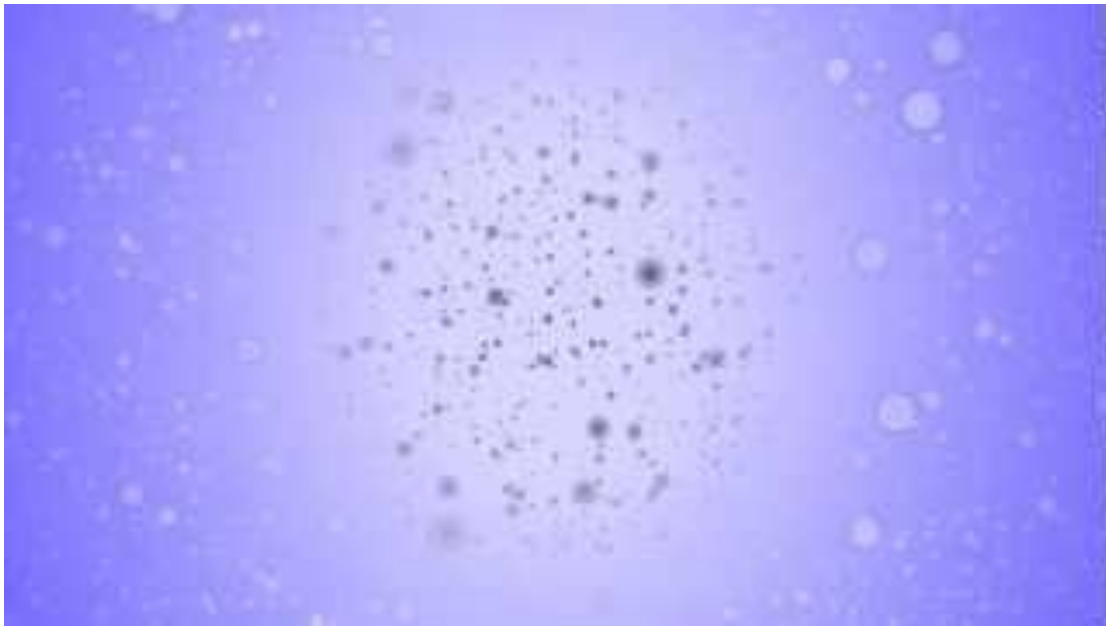
Pre-computer era (1700 - 1956)

Academic era (1957 - 2009)

First startup wave (2010 - 2016)

First startup wave (2010 - 2016)

- Product focus

- Full-piece generation

- High-quality music datasets

- Machine learning

First startup wave (2010 - 2016)

Melodrive (2016)

- Real-time video game music generation

- Music adapts to emotional context

- Unity SDK for indie game devs

```
┌─────────────────────────────────────┐
│                                     │
│      Pre-computer era (1700 - 1956) │
│                                     │
└─────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────┐
│                                     │
│      Academic era (1957 - 2009)     │
│                                     │
└─────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────┐
│                                     │
│      First startup wave (2010 - 2016)│
│                                     │
└─────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────┐
│                                     │
│      Big tech experiments (2017 - 2022)│
│                                     │
└─────────────────────────────────────┘
```

Big tech experiments (2016 - 2022)

- Deep learning focus

- Massive datasets

- Massive computational power

- No commercial end goal

Big tech experiments (2016 - 2022)
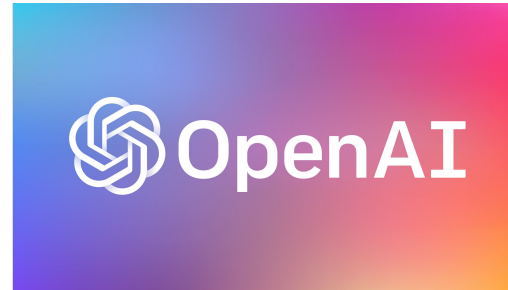
AWS DeepComposer (Amazon, 2019)

Jukebox (OpenAI, 2020)

- Raw-audio generation

- Advanced Deep Learning

- Full piece + lead vocals

- Performance details

[Open Source Research](#) (The Sound of AI, 2020)

- Community-driven research project

- 150+ people

- 2 years

- Voice-to-sound synthesizer

- [Paper](#) published at AIMC 2022

generative music in 2023

generative music before 2023

```
┌─────────────────────────────────────┐
│                                     │
│    Pre-computer era (1700 - 1956)   │
│                                     │
└─────────────────────────────────────┘
                   │
                   ▼
┌─────────────────────────────────────┐
│                                     │
│     Academic era (1957 - 2009)      │
│                                     │
└─────────────────────────────────────┘
                   │
                   ▼
┌─────────────────────────────────────┐
│                                     │
│    First startup wave (2010 - 2016) │
│                                     │
└─────────────────────────────────────┘
                   │
                   ▼
┌─────────────────────────────────────┐
│                                     │
│  Big tech experiments (2017 - 2022) │
│                                     │
└─────────────────────────────────────┘
                   │
                   ▼
┌─────────────────────────────────────┐
│                                     │
│      Music AI hype (2023 - ?)       │
│                                     │
└─────────────────────────────────────┘
```

Music AI hype (2023 - ?)

- Music industry is on fire

- Big tech explores commercial opportunity

- Scalable technology

- Really massive music datasets

- New startup wave
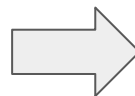
GENERATIVE AI

imgflip.com
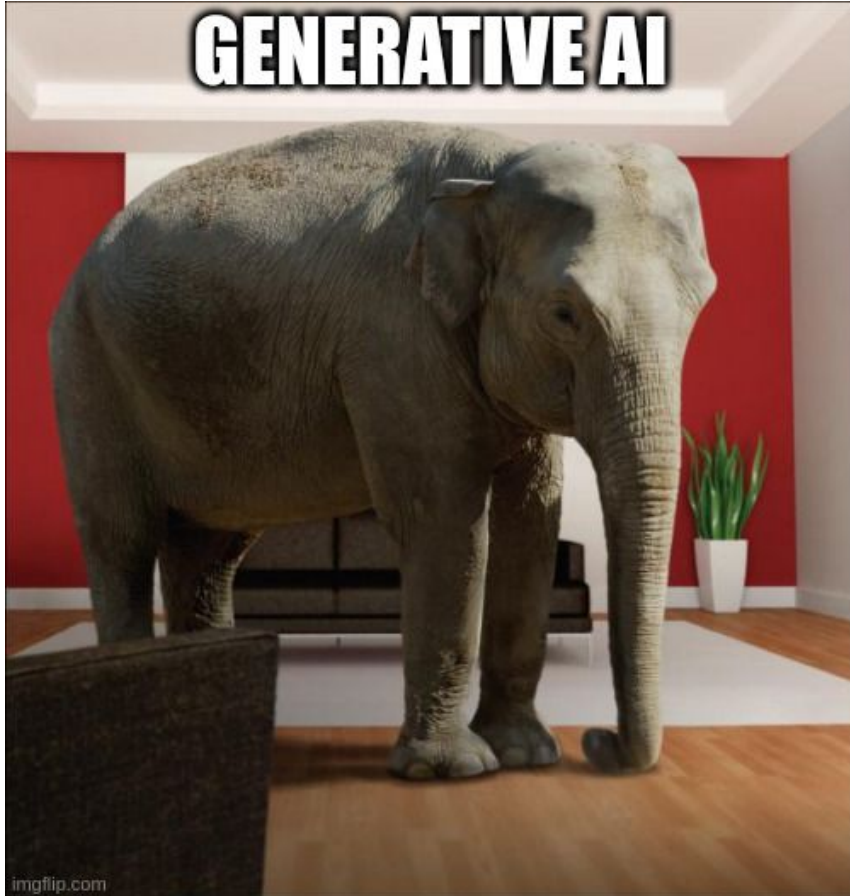
Text-to-image generation

Text-to-music generation

A fusion of reggaeton and electronic dance music, with a spacey, otherworldly sound. → Music model →
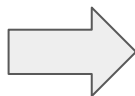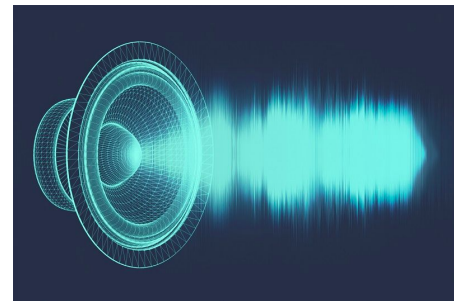
MusicLM (Google, 2023)
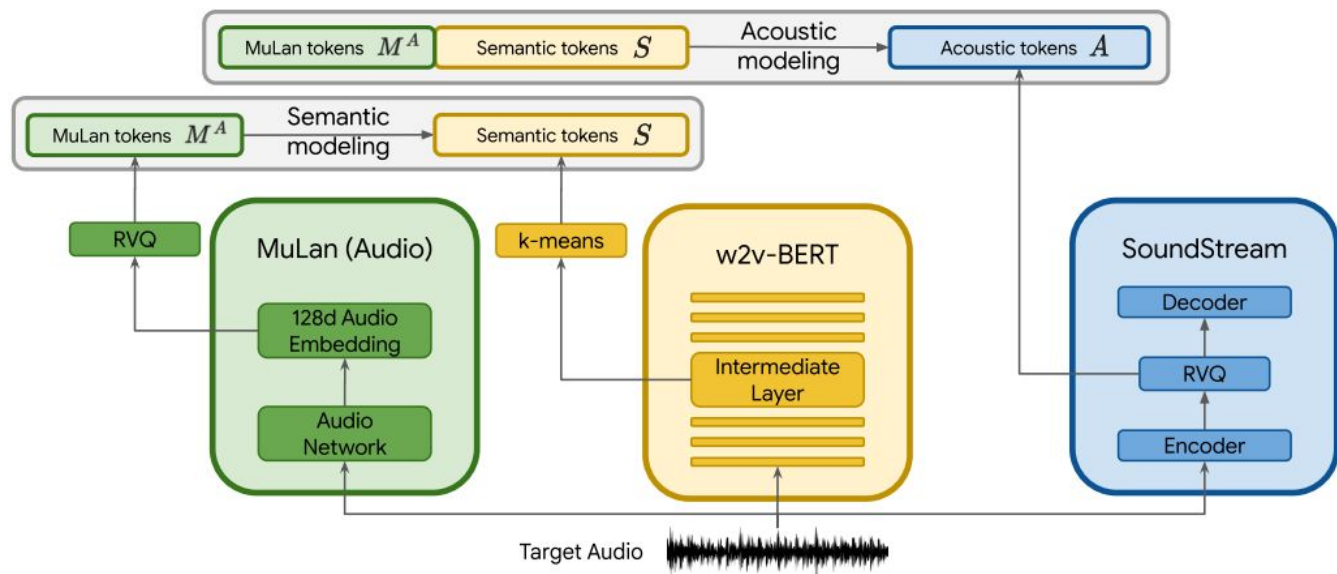
[MusicGen](#) (Meta, 2023)

Simple and Controllable Music Generation

Jade Copet[♠◇]   **Felix Kreuk**[♠◇]   **Itai Gat**   **Tal Remez**   **David Kant**

Gabriel Synnaeve[◇]   **Yossi Adi**[*◇]   **Alexandre Défossez**[◇]

♠: equal contributions, ◇: core team

Meta AI

{jadecopet, felixkreuk, defossez}@meta.com

Abstract

We tackle the task of conditional music generation. We introduce MUSICGEN, a single Language Model (LM) that operates over several streams of compressed discrete music representation, i.e., tokens. Unlike prior work, MUSICGEN is comprised of a single-stage transformer LM together with efficient token interleaving patterns, which eliminates the need for cascading several models, e.g., hierarchically or upsampling. Following this approach, we demonstrate how MUSICGEN can generate high-quality samples, while being conditioned on textual description or melodic features, allowing better controls over the generated output. We conduct extensive empirical evaluation, considering both automatic and human studies, showing the proposed approach is superior to the evaluated baselines on a standard text-to-music benchmark. Through ablation studies, we shed light over the importance of each of the components comprising MUSICGEN. Music samples, code, and models are available at github.com/facebookresearch/audiocraft.

1 Introduction

Text-to-music is the task of generating musical pieces given text descriptions, e.g., "90s rock song with a guitar riff". Generating music is a challenging task as it requires modeling long range sequences. Unlike speech, music requires the use of the full frequency spectrum [Müller, 2015]. That means sampling the signal at a higher rate, i.e., the standard sampling rates of music recordings are 44.1 kHz or 48 kHz vs. 16 kHz for speech. Moreover, music contains harmonies and melodies from different instruments, which create complex structures. Human listeners are highly sensitive to disharmony [Fedorenko et al., 2012, Norman-Haignere et al., 2019], hence generating music does not leave a lot of room for making melodic errors. Lastly, the ability to control the generation process in

Generative audio models launched in Jan '23

- Mousai

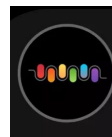- AudioLDM

- SingSong

- RAVE 2

- Riffusion (Dec '22)

- …

Second startup wave

YOUTUBE ANNOUNCES AI MUSIC PRINCIPLES AND LAUNCHES YOUTUBE MUSIC AI INCUBATOR WITH ARTISTS, SONGWRITERS AND PRODUCERS FROM UNIVERSAL MUSIC GROUP

[New York, August 21, 2023] Today YouTube published a first ever set of AI music principles and launched the **YouTube Music AI Incubator**, kicking off with artists, songwriters and producers from **Universal Music Group**. YouTube's three fundamental AI music principles are rooted in its commitment to collaborate with the music industry alongside bold and responsible innovation in the space.

Key takeaways

- 5 eras of GM

Key takeaways

- 5 eras of GM

- It started with manual algorithms

Key takeaways

- 5 eras of GM

- It started with manual algorithms

- It became an academic niche in the 1950s

Key takeaways

- 5 eras of GM

- It started with manual algorithms

- It became an academic niche in the 1950s

- Startups developed GM products from 2010

Key takeaways

- 5 eras of GM

- It started with manual algorithms

- It became an academic niche in the 1950s

- Startups developed GM products from 2010

- Tech giants initially used GM as an AI testground

Key takeaways

- 5 eras of GM

- It started with manual algorithms

- It became an academic niche in the 1950s

- Startups developed GM products from 2010

- Tech giants initially used GM as an AI testground

- In 2023, GM exploded

What next?

Generative music use cases