

Using Generative AI in Music Composition & Performance



TCF2024 - Jeremy dePrisco





Jeremy dePrisco **Digital Media Specialist**

Background:

Healthcare IT
Project management
Software development
Music production

Instruments:

Bass **Guitar**
Keyboard **Vocals**



AGENDA

- Opening track
- Overview of music business roles
- Where music is used
- Customer base
- Song anatomy
- Where we've been
- Where we're heading
- Johnny Cash Lyric AI shoot out
- Why AI Music?
- 2023 \$\$\$
- Discussion/Questions

CAVEATS:

Overview

Non-technical

Lots of questions

No definitive answers



Creation		Administration	Exploitation	Ancillary
(Primary)	(Secondary)	(Business)	(Sales/Promotion)	(Misc)
Lyricist	Recording Engineer	A&R Representative	Streaming companies	Live Sound Engineer
Composer	Mixing Engineer	Record Label Executive	Radio DJ	Lighting Technician
Singer/Songwriter	Mastering Engineer	Artist Manager	Music Publisher	Roadie
Music Producer		Booking Agent	Music Publicist	Studio Manager
Session Musician		Tour Manager	Music Marketing Manager	Music Video Director
Sound Designer		Music Lawyer	Music Journalist	Web Developer
DJ/Remixer		Music Licensing Manager	Music Supervisor	Instrument Technician
		Live Event / Venue Manager		Recording software developers
		Performing Rights Organizations		Music hardware manufacturers
				Music Therapist
				Music Teacher
	<i>Process</i>		<i>End Result</i>	Advocate
				Choreographer
				Enthnomusicologist
				CD Manufacturing Plant
				Vinyl Manufacturing Plant

Jeremy's Roles

Jeremy's Roles				
Creation		Administration	Exploitation	Ancillary
(Primary)	(Secondary)	(Business)	(Sales/Promotion)	(Misc)
Lyricist	Recording Engineer	A&R Representative	Streaming companies	Live Sound Engineer
Composer	Mixing Engineer	Record Label Executive	Radio DJ	Lighting Technician
Singer/Songwrite	Mastering Engineer	Artist Manager	Music Publisher	Roadie
Music Producer		Booking Agent	Music Publicist	Studio Manager
Session Musician		Tour Manager	Music Marketing Manager	Music Video Director
Sound Designer		Music Lawyer	Music Journalist	Web Developer
DJ/Remixer		Music Licensing Manager	Music Supervisor	Instrument Technician
		Live Event / Venue Manager		Recording software developers
		Performing Rights Organizations		Music hardware manufacturers
				Music Therapist
				Music Teacher
	<i>Process</i>		<i>End Result</i>	Advocate
				Choreographer
				Enthnomusicologist
				CD Manufacturing Plant
				Vinyl Manufacturing Plant

Video Games



Commercials

For wealthy adults with type 2 diabetes.

Movies



Streaming



Theatre



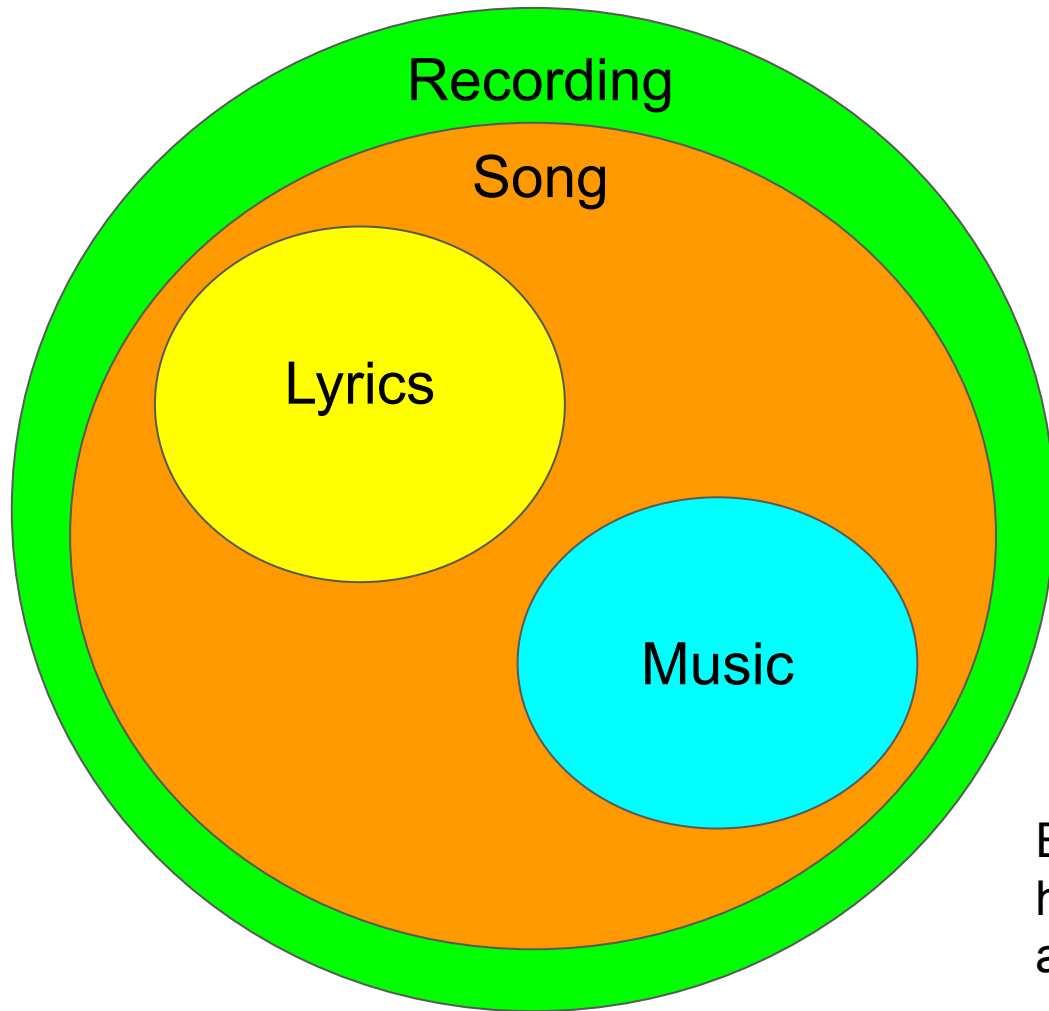
Podcasts

"**Magic**" is a 1974 song by Scottish **pop rock** band **Pilot**, and was the first hit single for the group. It was written by band members **David Paton** and **Billy Lyall** for their debut album, *From the Album of the Same Name*.

Customers Who Benefit from the Music Industry

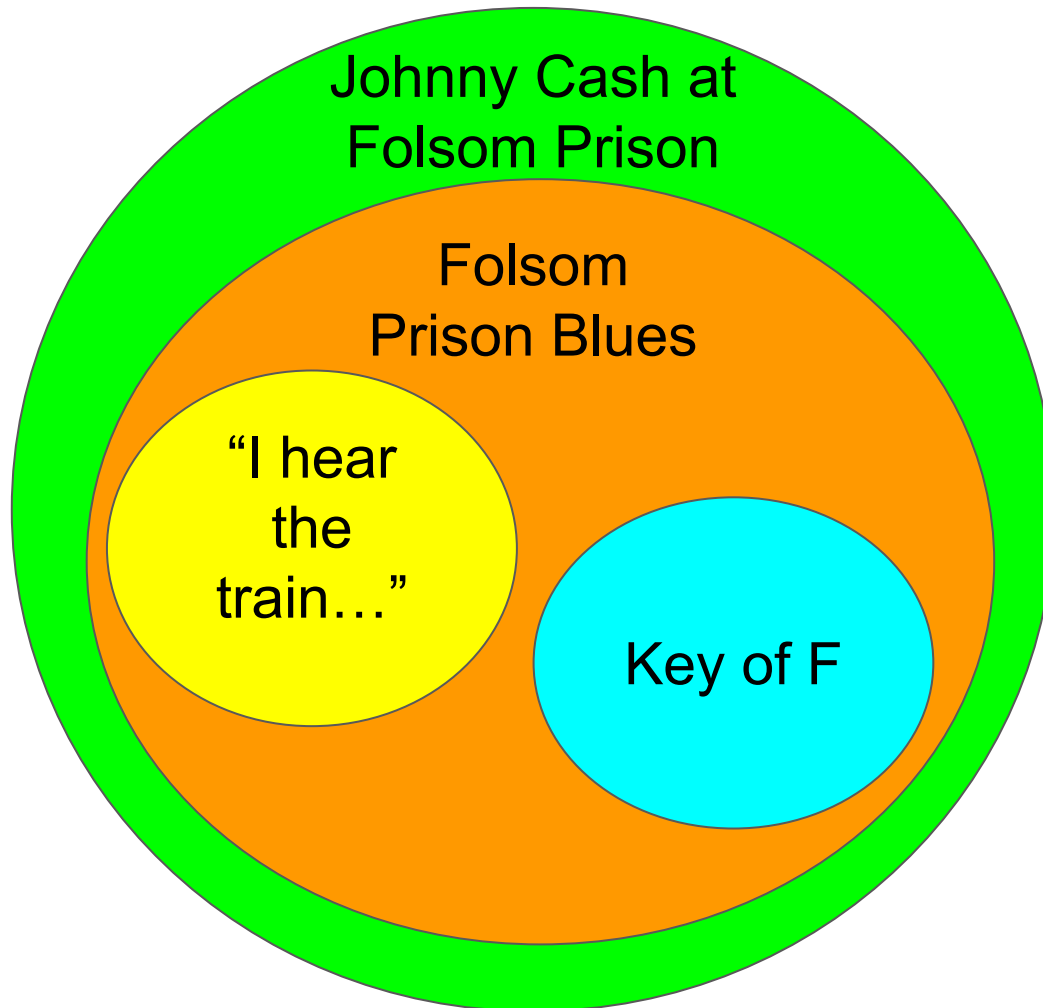
Direct Beneficiaries	Indirect Beneficiaries	Additional Beneficiaries
Music Fans	Technology Companies	Educational institutions incorporate music into their curriculum.
Musicians and Singers	Telecommunication Companies	Hospitals and healthcare facilities using music therapy for patients.
Songwriters and Composers	Marketing and Advertising Agencies	Communities and individuals experiencing cultural enrichment and emotional well-being through music.
Record Labels and Distributors	Film and Television Studios	
Concert and Live Event Promoters	Radio and Broadcasting Stations	
Music Venue Owners and Operators	Social Media Platforms	
Music Instrument Manufacturers and Retailers	Travel and Tourism Industry	
Audio Engineering and Production Professionals	Merchandise Retailers	
Music Teachers and Instructors	Food and Beverage Vendors at Music Events	
Music Journalists and Critics	Cleaning and Sanitation Companies	

Song Anatomy

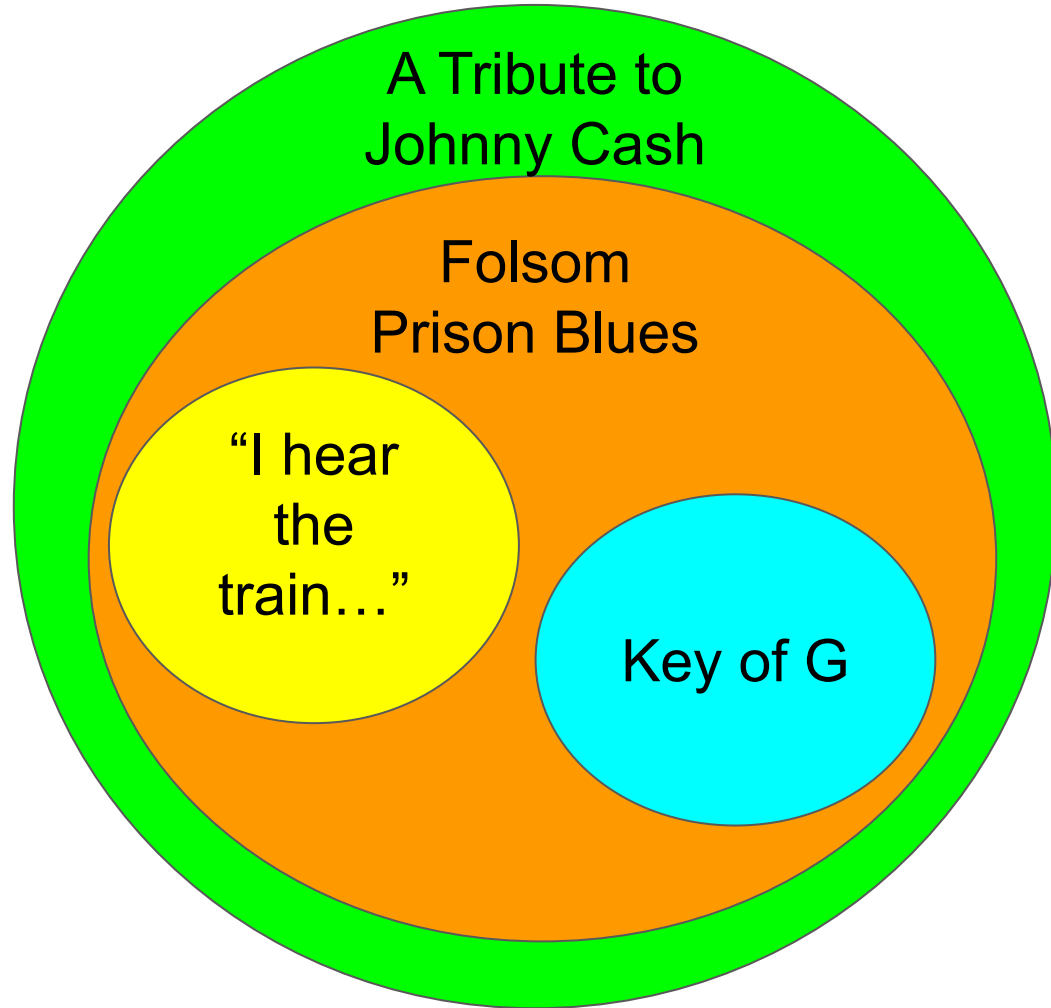


Each element could
have metadata
attached

1968



2002



Characteristics of music

- **Mood**
- **Rhythm**
- **Melody**
- **Harmony**
- **Dissonance**
- **Texture**
- **Dynamics**
- **Tempo**
- **Timbre**
- **Form**
- **Expression**
- **Style**
- **Narrative**
- **Technique**
- **Context**

Where we've been...



Algorithmic composition (hundreds of years)

https://en.wikipedia.org/wiki/Algorithmic_composition

Aleatoric music (late 15th century to present)

https://en.wikipedia.org/wiki/Aleatoric_music

Turing Machine (1936 Alan Turing)

https://en.wikipedia.org/wiki/Turing_machine

Generative music (1990s Brian Eno)

https://en.wikipedia.org/wiki/Generative_music

Euclidean music (Godfried Toussaint 2004-05)

https://en.wikipedia.org/wiki/Euclidean_rhythm





AudioCraft consists of three models:

MusicGen

- Trained with Meta-owned and specifically licensed
- Generates music from text-based user inputs

AudioGen

- Trained on **public** sound effects
- Generates audio from text-based user inputs
- Generate environmental sounds and sound effects (dog barking, cars honking, etc)

EnCodec decoder

- allows for higher quality music generation
- fewer artifacts

Source: [AudioCraft: Generative AI for audio](#)



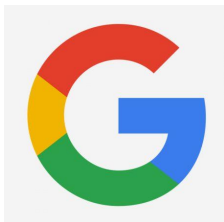
OpenAI

MuseNet (2019)

- a deep neural network
- can generate 4-minute musical compositions
- 10 different instruments
- Can combine styles from country to Mozart to the Beatles.
- **not explicitly programmed with “human” understanding of music**
- AI **“discovered” patterns of harmony, rhythm, and style** by learning to predict the next token in hundreds of thousands of MIDI files.
- Unsupervised technology (GPT-2), a large-scale transformer model trained to predict the next token in a sequence, whether audio or text.

Jukebox (2020) - neural net that generates music, including **rudimentary singing**, as raw audio in a variety of genres and artist styles.

Where we're heading...



[Google's MusicFX](#)

Prompt based music

Example: "Sound that creates a soothing and adventurous atmosphere."

[DeepMind Partnerships with YouTube:](#)

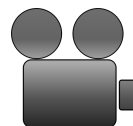
Lyria - a GenAI model for music

Dream Track - AI tunes - Text to sound track for shorts.



[Adobe's Project Music GenAI Control](#)

create audio tracks using text prompts and even transform your audio based on reference melodies. Transform time-consuming workflows with pixel-level control over music.



Where we're heading_

[Aiva](#) (2016) - targets individuals and businesses aiming to craft soundtracks for different media, allowing for customization and versatility.

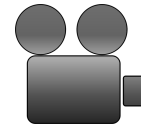
[WavTool](#) - The World's First Text-to-Music AI DAW Uses GPT-4

[MelodyStudio](#) - AI-based system that comes up with original melody ideas. The melodies are made specifically for the lyrics provided.

[Bandlab Songstarter](#) - free AI-powered idea generator that acts as your starting point to create new music – all in one easy-to-use interface.

[EZ Drummer](#) - Virtual drummer that responds to your playing in a number of styles.

[Oslyn](#) - Digital Bandmate for Guitarists - a mobile and web app that performs real time analysis of your guitar performance and adds ambient pads.



Loudly

- Advanced AI-driven music creation platform
- Extensive training set of 10 million songs
- A sound bank containing 200,000 human-generated recordings.
- Empowers users to produce **royalty-free music** tailored to their needs
- Uses simple natural-language prompts
- Specify the style, tempo, mood, and even individual instruments.
- Embraces the concept of "music as code"
- Enables deep interaction with music at a micro-level

Notably, the platform holds the copyrights to all the music it's been trained on, ensuring no concerns of copyright infringements or artists feeling their original works have been misused. (Bernard Marr - Forbes)

Login required. <https://www.loudly.com/>

Stable Audio (Stability AI)











- Latent diffusion model architecture for audio
- Conditioned on text metadata as well as audio file duration and start time
- Allows for control over the content and length of the generated audio.
- Works with a heavily downsampled latent representation of audio.
- Allows for much faster inference times compared to raw audio.
- Can render 95 seconds of stereo audio at a 44.1 kHz sample rate in less than one second on an NVIDIA A100 GPU.

<https://stableaudio.com/> (login required)

Research paper:

<https://stability.ai/research/stable-audio-efficient-timing-latent-diffusion>

Stable Audio (Stability AI)

Free 	Pro 	Studio 	Max 
It's free. Get started!	\$11.99 a month excluding VAT/sales taxes	\$29.99 a month excluding VAT/sales taxes	\$89.99 a month excluding VAT/sales taxes
Monthly track generations 20	Monthly track generations 500	Monthly track generations 1,300	Monthly track generations 4,500
Track duration Up to 45 seconds	Track duration Up to 90 seconds	Track duration Up to 90 seconds	Track duration Up to 90 seconds
License  <input type="radio"/> Personal license	License  <input checked="" type="radio"/> Creator license	License  <input checked="" type="radio"/> Creator license	License  <input checked="" type="radio"/> Creator license
Early feature access At a later date	Early feature access <input checked="" type="checkbox"/> First to experience	Early feature access <input checked="" type="checkbox"/> First to experience	Early feature access <input checked="" type="checkbox"/> First to experience

NOT AN ENDORSEMENT

Enterprise subscription - companies can customize their usage and price. ***Those using the free version cannot commercially use audio they make with Stable Audio.***

Article: <https://www.theverge.com/2023/9/13/23871635/stability-ai-generative-audio-model-platform>

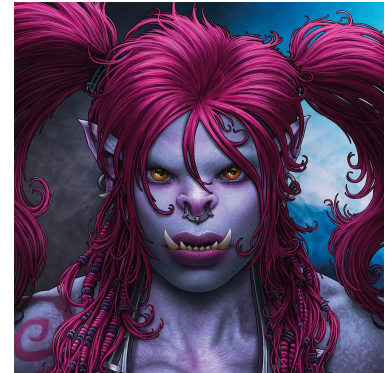
23 AI-Generated Old School Death Metal Bands That Don't Actually Exist

(from <https://lambgoat.com/>)



The Frostbite Orckings

- Company: **Musical Bits** - Bingen, Germany
- Founded in late 2021 by **Stefan Schmidt**, a singer, guitar player, piano player, and composer
- Stefan has toured internationally with his metal a-cappella band VAN CANTO while working a day job in Artificial Intelligence.
- Based on “Creativity Layers”
- Stacks creative parts of composition like arrangement, rhythm, instrumental interplay, harmony, melodies, lyrics and text, on top of each other.
- Each layer has its own set of custom built and trained AI models.
- Started with metal because that was their domain expertise, but branching to other genres.
- “You need the different domain experts from a musician side as well to make a meaningful system.”



Article:

<https://www.forbes.com/sites/charliefink/2023/09/18/ai-generated-music-with-real-human-fans/?sh=79bc419f5fd3>



LYRICS

**Folsom Prison Blues
by Johnny Cash**

ChatGPT vs BARD

Write a parody of Folsom Prison Blues by Johnny Cash - ChatGPT vs Bard

ChatGPT

(Verse 1)

I hear the code a-comin', it's rollin' 'round the bend
And I ain't seen the sunlight since I don't know when
I'm stuck in server prison, and time keeps **dragin'** on
But that code keeps a-rollin', on and on

(Chorus)

**When I was just a baby, my mama told me, "Son
Always be a good dev, don't ever code for fun"
But I hacked a man in Reno, just to watch him cry
When I hear that error message, I hang my head and sigh**

(Verse 2)

**I bet there's rich folks codin' in a fancy office chair
They're probably **drinkin' lattes and pretendin' not to care**
Well, I know I had it comin', I know I can't be free
But those bugs keep on **hauntin'**, that's what tortures me**

(Chorus)

**Well, if they freed me from this prison, if that railroad train was mine
I bet I'd move it on a little farther down the line
Far from server prison, that's where I want to stay
And I'd let that lonesome whistle blow my blues away**

(Bridge)

**I hear the clang of keyboards, it's breakin' all my dreams
The hum of data centers, it's louder than it seems
The server room is cold, and the WiFi's always slow
But I'm stuck in this code, with no place to go**

(Verse 3)

**Well, if they freed me from this prison, if that SQL was kind
I'd run away from coding, leave it all behind
But here I am stuck coding, with no end in sight
Just a prisoner of syntax, tryin' to make it right**

(Chorus)

So, I'll keep on **debuggin'**, in this language I disdain
Hoping for redemption, praying for no more pain
But that code keeps a-rollin', and the deadlines draw near
I'm stuck in this coding prison, with a pocket full of fear

Bard

(Verse 1)

I hear the modem humming, my inbox **overflowin'**
Spam and scams and phishing schemes, all endlessly encroaching
Got messages from princes, promises of sudden wealth
But I ain't buying, not tonight, just locked in digital stealth

(Chorus)

This ain't San Quentin, ain't Alcatraz, ain't Folsom or no chain
But these walls of firewalls hold me in, another kind of pain
Spam folders piled like cell blocks, passwords rusting in the code
Stuck here in the internet's jail, with nowhere else to go

(Verse 2)

Used to roam the open web, a cowboy of the bytes
Clicking through the endless fields, bathed in neon lights
Downloaded songs and stolen books, movies 'til the dawn
But those carefree days are long gone, the internet's a war zone

(Chorus)

This ain't San Quentin, ain't Alcatraz, ain't Folsom or no chain
But these walls of firewalls hold me in, another kind of pain
Spam folders piled like cell blocks, passwords rusting in the code
Stuck here in the internet's jail, with nowhere else to go

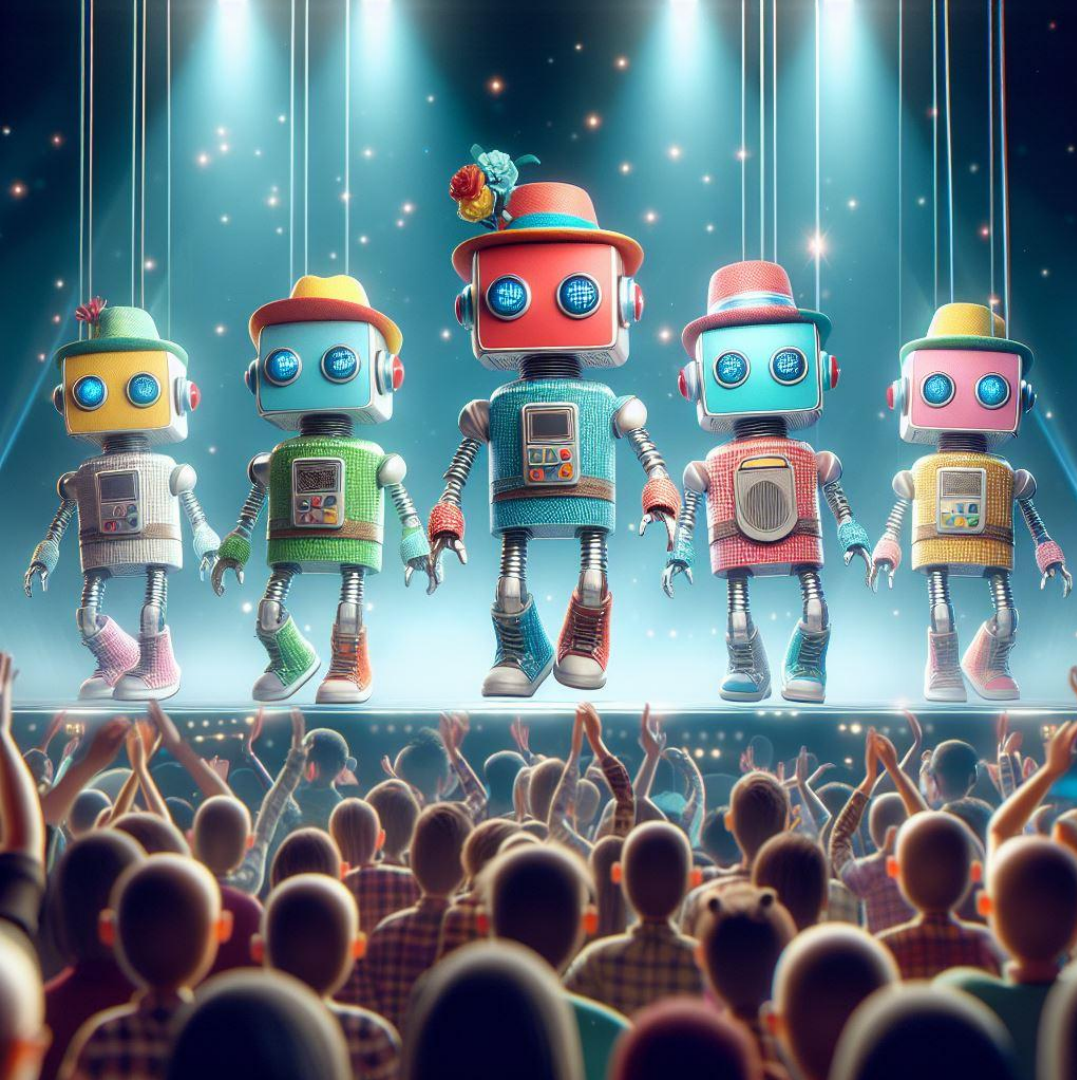
(Bridge)

Phishing lures cast everywhere, tempting clicks like bait
Malware stalks the shadows, spreading digital decay
One wrong move, one foolish trust, and freedom turns to ash
This cyber purgatory, its grip will forever clash

(Chorus)

(Outro)

But maybe someday, through the bars, a signal I can find
Break free from these digital chains, leave this prison far behind
Till then, I'll keep on surfing, through the storm of spam and lies
Hoping for a digital dawn, with freedom in its skies



Why would these groups want to explore AI Music?

Musicians/Producers

- More options for creativity
- Teaching/Learning tool
- Technical problem solving

Non-musicians

- Lower barrier of entry
- Learning tool
- Lower cost (?)

Fans/Listeners

- More variety (but potentially cluttered landscape)

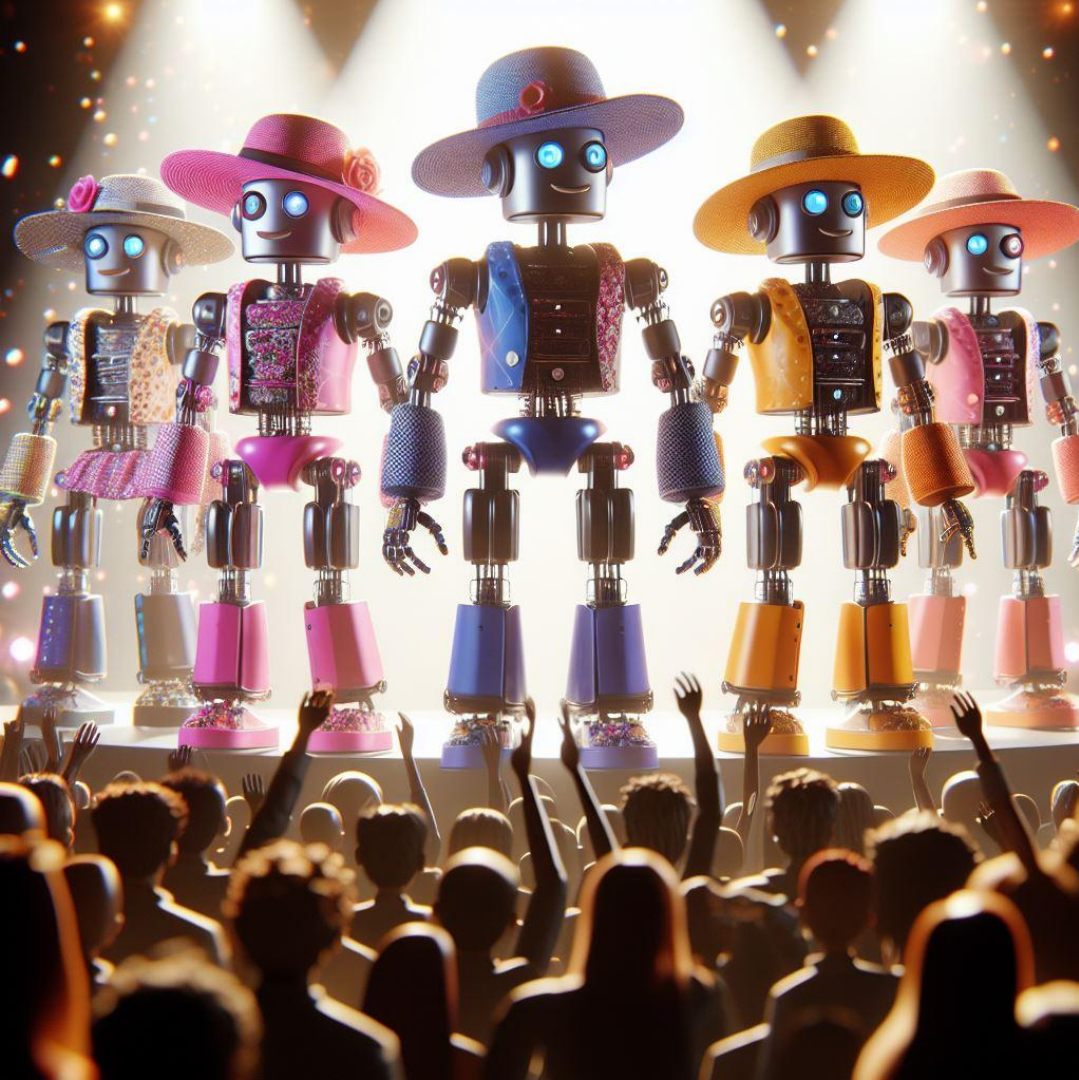


IFPI's Global Music Report 2023 Topline Figures:

- Global music sales up 9% to \$26.2 billion
- Streaming subscription revenues up 10.3% to \$12.7 billion
- Total streaming revenues (including paid and ad-supported) up 11.5% to \$17.5 billion
- Physical revenues up 4% to \$4.6 billion
- Sync income was up 22.3% to \$0.6 billion,
- Performance rights revenues rise 8.6% to \$2.5 billion
- 589 million paid music subscribers
- Streaming's share of global music sales: 67%

SOURCE:

<https://www.billboard.com/pro/ifpi-global-report-2023-music-business-revenue-market-share/>



Questions for discussion

Ethics

Are the models trained ethically?

Is it ethical to use them?

Each platform is different.

Who holds copyright?

It depends.

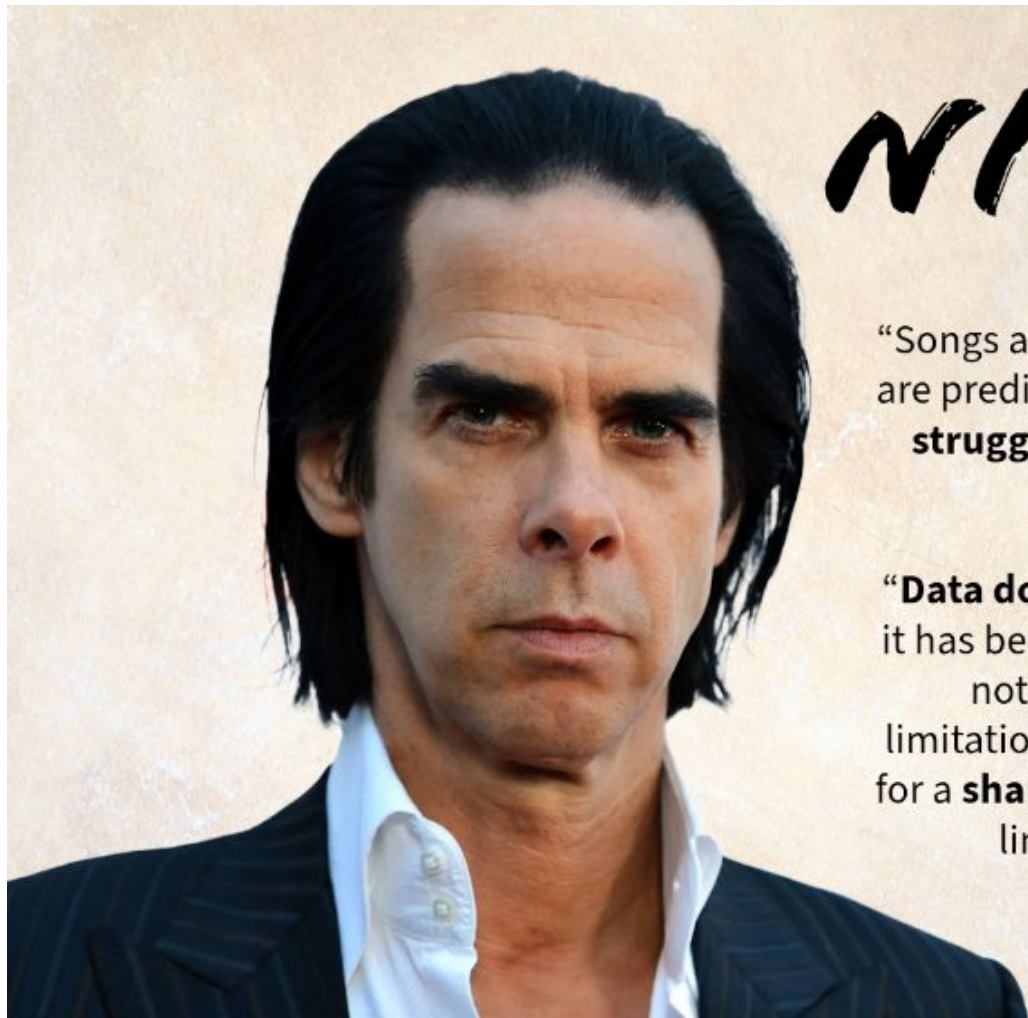
Read the fine print.

What about watermarking?

Early stages

May be thwarted by savvy folks

How many actual musicians are involved in the development process?



NICK CAVE

“Songs arise out of **suffering**, by which I mean they are predicated upon the **complex, internal human struggle of creation** and, well, as far as I know, algorithms don’t feel,”

“**Data doesn’t suffer**. ChatGPT has no inner being, it has been nowhere, it has endured nothing, it has not had the audacity to reach beyond its limitations, and hence it doesn’t have the capacity for a **shared transcendent experience**, as it has no limitations from which to transcend.”

Stay in touch!



Jeremy dePrisco
Digital Media Specialist



jjdeprisco@gmail.com
www.jeremydeprisco.com