



Universitat  
Pompeu Fabra  
Barcelona

MTG  
Music Technology  
Group

# LA IA Y EL ECOSISTEMA MUSICAL DESAFÍOS, RIESGOS, OPORTUNIDADES

Sergi Jordà

Music Technology Group - DTIC-UPF

# IA Y CREATIVIDAD 2022



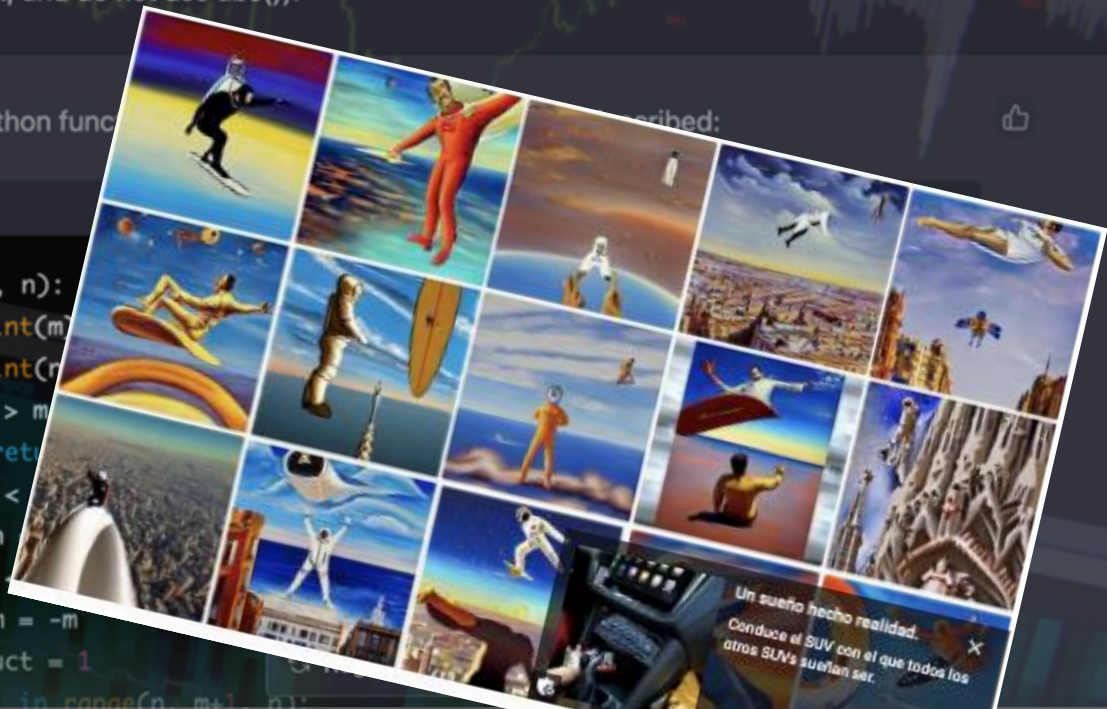
Write a Python function  $f1(m,n)$  that receives two numerical values ( $m$  and  $n$ ) and returns the product of all values from 1 to  $m$  that are multiples of  $n$ . If  $m$  and/or  $n$  are floats (not integers), the function should truncate them before using them. If  $n > m$ , the function will return -1. If  $m$  and/or  $n$  are negative, their absolute values should be used (use your own code for that, and do not use `abs()`).

Here is a Python func

described:



```
def f1(m, n):
    m = int(m)
    n = int(n)
    if n > m:
        return -1
    if n < 0:
        n = -n
    if m < 0:
        m = -m
    product = 1
```



Un sueño hecho realidad.  
Conduce el SUV con el que todos los  
otros SUV's sueñan ser.



Universitat  
Pompeu Fabra  
Barcelona

Unidad de Coordinación Académica  
de Ingenierías y Tecnologías  
de la Información y las Comunicaciones

Escuela de Ingeniería  
Departamento de Tecnologías  
de la Información y las Comunicaciones

Educación ▾

Investigación ▾

Divulgación ▾

Actualidad ▾

Sobre nosotros ▾



# El Departamento de Tecnologías de la Información y las Comunicaciones, reconocido como unidad de excelencia María de Maeztu por segunda vez

La Agencia Estatal de Investigación publicó el viernes, 4 de noviembre, la resolución provisional de concesión de las nuevas acreditaciones y ayudas públicas 2022 de centros de excelencia 'Severo Ochoa' y de unidades de excelencia 'María de Maeztu'.

[Leer la noticia >](#)

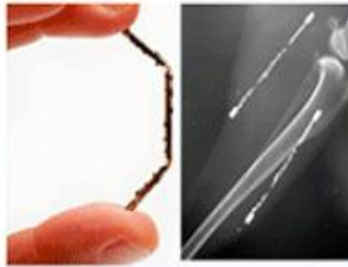
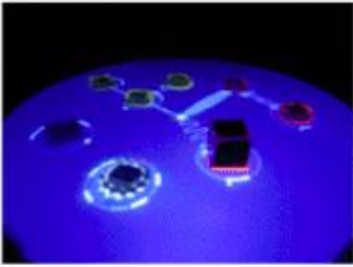
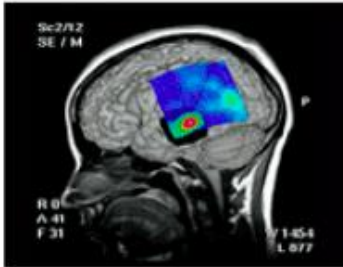
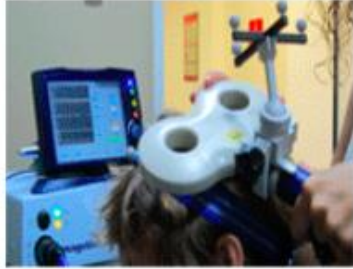


Universitat  
Pompeu Fabra  
Barcelona

MTG  
Music Technology  
Group

# Escuela de Ingenierías

## Departamento de Tecnologías de la Información y las Comunicaciones (DTIC)



# About the MTG

The Music Technology Group (MTG) is a research group part of the Pompeu Fabra University in Barcelona with more than 25 years of expertise developing research projects around sound and music technologies, that has resulted in a solid knowledge base and a significant portfolio of technologies.

The MTG is experienced in achieving exploitation objectives with international impact: it has created 4 spin-off companies and it has a number of active license agreements and projects with the industry.

**75**

MTG current  
researchers, developers  
and students

**+100**

Datasets / software  
tools

**+90**

Companies with which  
we have collaborated

**+100**

Research projects (on  
going and finished)

# What we do



## Education

The MTG has extensive experience in educating researchers and professionals in the various topics related to Music Technology. Apart from a PhD program, we coordinate a master program with 20 new students every year coming from all over the world.



## Research

Most of our research is carried out within projects funded by public or private sources. The research results into publications and software & datasets, while also emphasizing technology transfer and outreach initiatives.



## Technology transfer

We transfer our know-how and results by licensing technologies and providing services with them, through specific industrial collaborations. We create solutions customized to the needs of companies, building upon our technologies and using the expertise of our researchers and developers. Special areas of technology transfer include automatic sound/music description, tools to support music education, and musical interfaces for music creation.

# Industrial partners

Some of our industrial partners:



NATIVE INSTRUMENTS

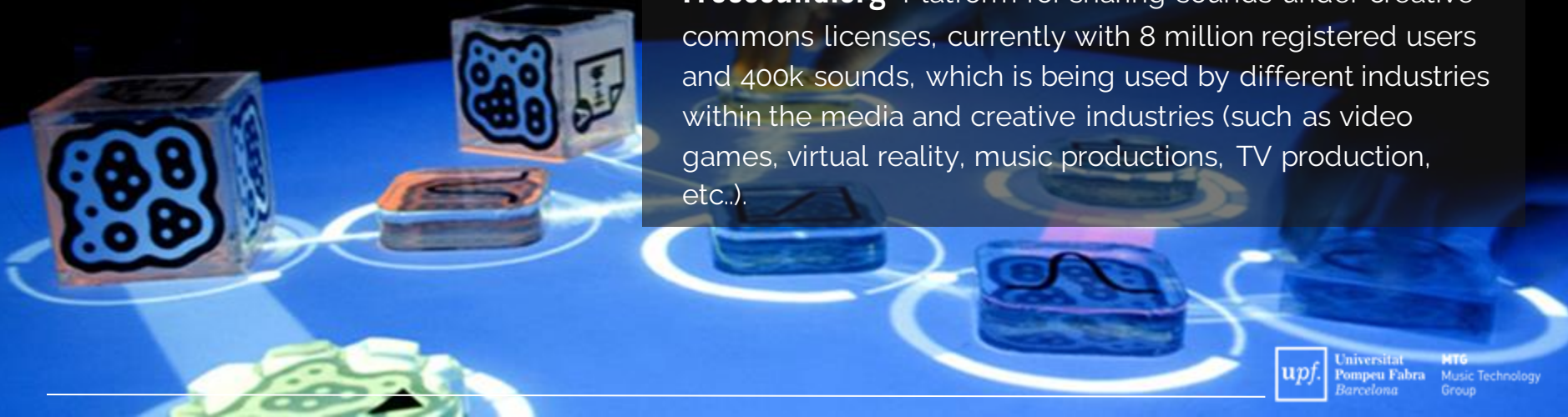


# Success stories

**Reactable:** Electronic musical instrument with a tabletop tangible user interface that became very popular after many musicians used it in their shows such as Björk or Coldplay.

**Vocaloid:** Singing voice synthesizer software developed in collaboration with Yamaha. It has been specially successful in Japan, due to the big impact of virtual singer Hatsune Miku.

**Freesound.org:** Platform for sharing sounds under creative commons licenses, currently with 8 million registered users and 400k sounds, which is being used by different industries within the media and creative industries (such as video games, virtual reality, music productions, TV production, etc..).





# Technologies

The MTG develops technologies for a wide variety of industrial applications in sectors such as media, video games, music, automotive, mobile...

Some of the solutions that our technologies offer are:

## **Synthesis:**

Analyze, transform and synthesize sounds using spectral modeling approaches

**Similarity:** Analyze audio and compute features to find similar sounds or music tracks

**Beat tracking:** Estimate beat positions and tempo (BPM) of a song

**Audio fingerprinting:** Extract fingerprints from any audio source using the Chromaprint algorithm

**Onset detection:** Detect onsets (and transients) in an audio signal

**Loudness metering:** Use various loudness meters including algorithms compliant with the EBU R128 broadcasting standard

**Melody extraction**  
Estimate pitch in monophonic and polyphonic audio

**Key detection:** Find a key of a music piece

**Classification:** Classify sounds or music based on computed audio features

**Spectral analysis:** Analyze spectral shape of an audio signal

**Mood detection:** Find if a song is happy, sad, aggressive or relaxed

**Segmentation:** Split audio into homogeneous segments that sound alike

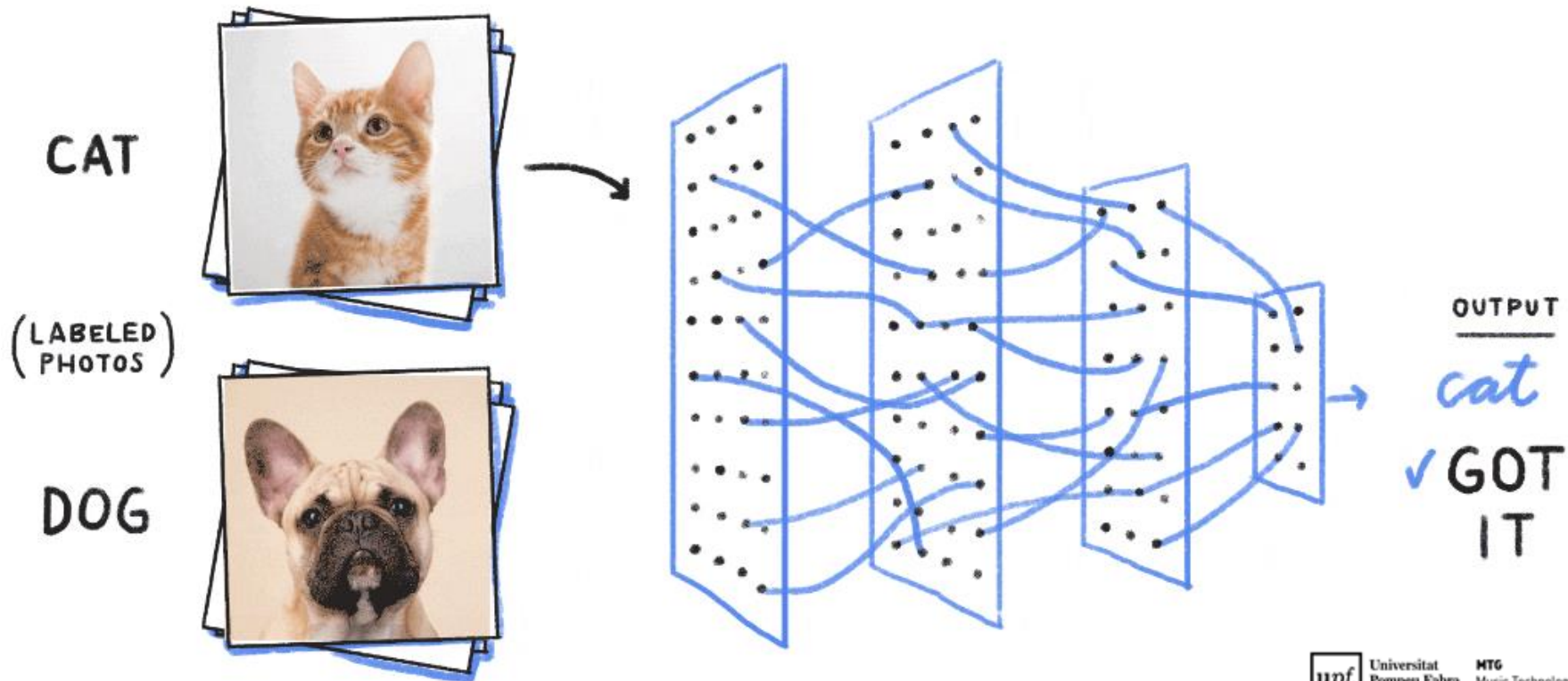
# Antecedentes de la IA musical

- Ada Lovelace (1842) *"Suponiendo que las relaciones fundamentales de los sonidos en la ciencia de la armonía y de la composición musical fueran susceptibles de formalizaciones numéricas, la máquina podría componer piezas musicales elaboradas y científicas de cualquier grado de complejidad o extensión"*
- 1957 ILLIAC Suite, primera pieza musical compuesta por ordenador por Lejaren Hiller (imagen fondo)
- Aplicaciones experimentales durante las décadas siguientes, con sistemas expertos, basados en reglas y probabilísticos (e.g. Markov), hasta la explosión del Deep Learning (DL) y las Redes Neuronales (RN) ca. 2015-2017

# Redes Neuronales para la creación musical !

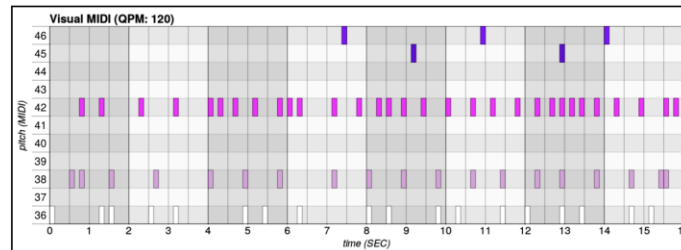
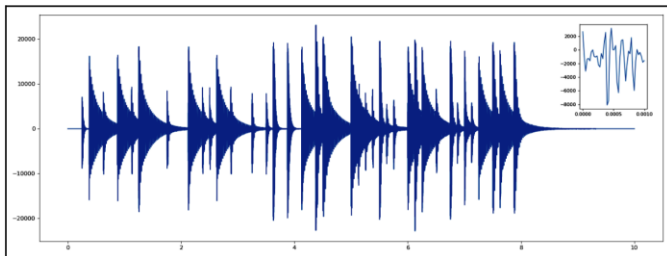


# Como funciona una red neuronal?



# Generación musical con Redes Neuronales

## Audio vs. Simbólico (MIDI)



**Generación simbólica (MIDI)** mucho más extendida que de **audio**

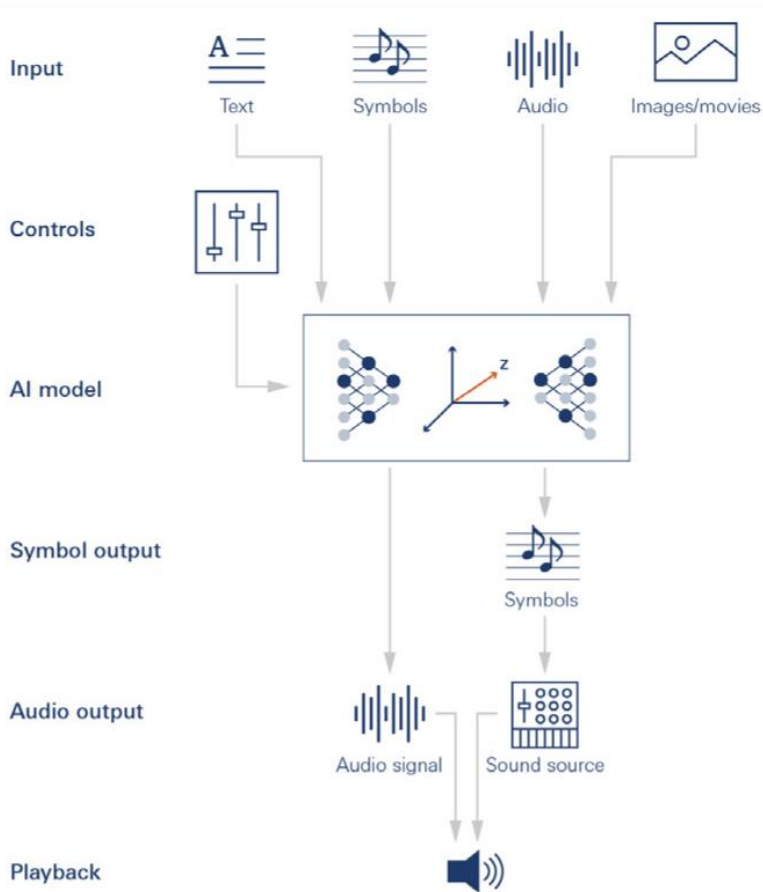
### Ventajas del MIDI

- Datos de menor volumen (~1/1000) y más fáciles de interpretar (~1/1000)
- Voces separadas en el fichero ==> para generar acompañamientos

### Inconvenientes del MIDI

- Mucha música NO disponible en este formato
- Muchos ficheros de mala calidad
- Sonido menos realista / no voces

# Generación musical con Redes Neuronales



## Entradas y Salidas

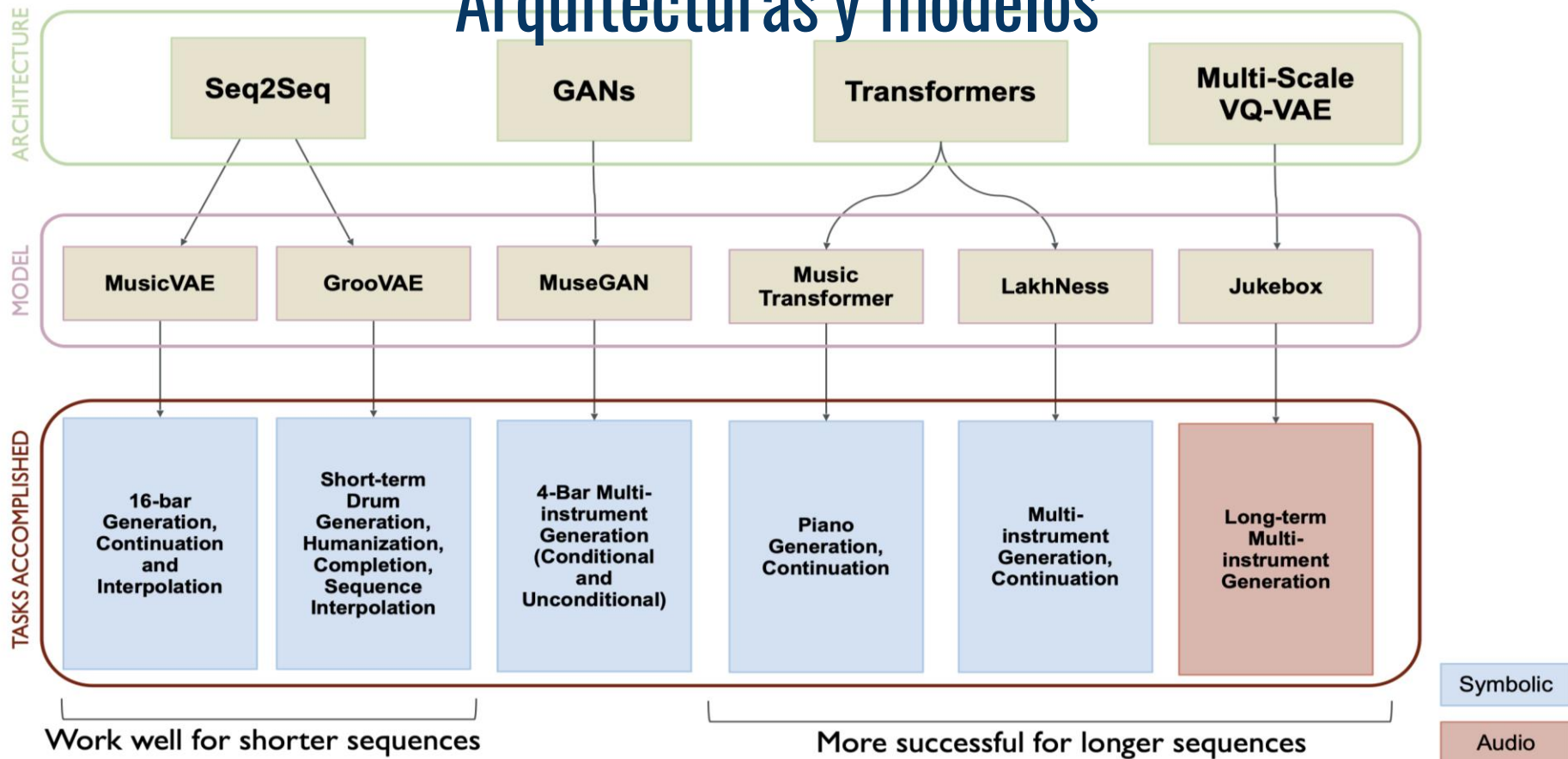
### Los inputs pueden ser variados

- Descripción textual de la música deseada (e.g. tonalidad, orquestación, genero, paisajes sonoros, etc.)
- Fragmento (MIDI) a complementar (e.g. acompañamiento para una melodía) o a continuar (e.g. interpretado en vivo desde un teclado)
- Fragmentos-melodías (MIDI) a combinar/interpolar, transferencia de estilos ...
- Generación de banda sonora para un video ...

**Output:** Generación **simbólica (MIDI)** mucho más extendida que de **audio** (que necesita volumen de datos y computación muy superior). Entrenamiento con MIDI datasets (~1-100K)

# Generación musical con Redes Neuronales

## Arquitecturas y modelos



# Algunos Ejemplos





PopMAG: Pop Music Accompaniment Generation

Arxiv: <https://arxiv.org/abs/2008.07703>

# Generación simbólica (MIDI)

Audio

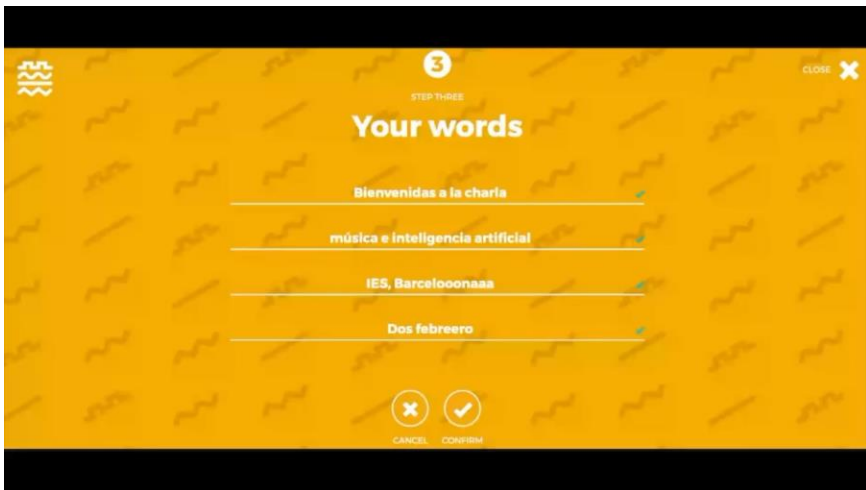
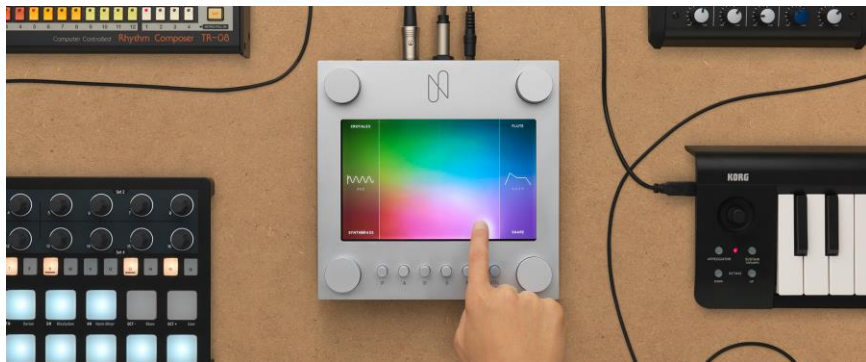
/	Melody (Input)	PopMAG	Ground-Truth
Sample 1			
Sample 2			
Sample 3			
Sample 4			

PopMAG (Microsoft Research 2020)

Genera acompañamiento para una melodía dada



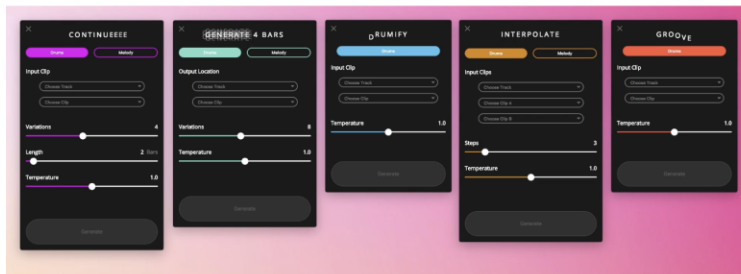
# Diseño sonoro / Síntesis / Producción



## Magenta Studio (v1.0)

Magenta Studio is a collection of music plugins built on Magenta's open source tools and models. They use cutting-edge machine learning techniques for music generation.

These tools are available both as standalone applications and as plugins for [Ableton Live](#). To find out more information, choose one of the links below:



**Nsynth Super Neural Synthesizer (Google 2018)**  
**Magenta Studio (Google 2018)**  
**VoctroLabs 2022**

# Generación de audio: Jukebox (OpenAI 2020)

## Jukebox

We're introducing Jukebox, a neural net that generates music, including rudimentary singing, as raw audio in a variety of genres and artist styles. We're releasing the model weights and code, along with a tool to explore the generated samples.

- [▶ Country, in the style of Alan Jackson – Jukebox](#)
- [▶ Rock, in the style of Elvis Presley – Jukebox](#)
- [▶ Pop, in the style of Katy Perry – Jukebox](#)
- [▶ Blues Rock, in the style of Joe Bonamassa – Jukebox](#)
- [▶ Heavy Metal, in the style of Rage – Jukebox](#)
- [▶ Classic Pop, in the style of Frank Sinatra – Jukebox](#)

April 30, 2020

12 minute read, 10 day listen



# Generación de audio: Jukebox (OpenAI 2020)

## Jukebox

We're introducing Jukebox, a neural net that generates music, including rudimentary singing, as raw audio in a variety of genres and artist styles. We're releasing the model weights and code, along with a tool to explore the generated samples.

Se le puede pedir (a partir de texto) que combine artistas, estilos, canciones, etc.

Lo más cercano al "Deepfake musical"

Entrenado con 1,2M de canciones

Horas de computación con potentes GPUs para 1 minuto de audio lo-fi

Todavía muy pocos ejemplos similares (generación audio)



April 30, 2020

12 minute read, 10 day listen

# Google MusicLM (27 Jan 23)

Escucharemos 2 ejemplos, **música para un videojuego** y un tema **reggae** que responden a las descripciones siguientes:

- *Banda sonora de un juego de arcade. Ritmo rápido y optimista, con un pegadizo riff de guitarra eléctrica. La música es repetitiva y fácil de recordar, pero con sonidos inesperados, como golpes de platillos o redoble de tambores*
- *Canción reggae de tempo lento, bajo y batería. Con guitarra eléctrica sostenida, bongos, y voces relajadas y muy expresivas*

A continuación veremos como el sistema admite también esbozos sonoros como entrada:

- Escucharemos un fragmento de "Bella Ciao" tarareado por el usuario, y le pediremos dos versiones, una electrónica y una de jazz con saxo

# Google MusicLM (27 Jan 23)

## MusicLM: Generating Music From Text

| paper |

*Andrea Agostinelli, Timo I. Denk, Zalán Borsos, Jesse Engel, Mauro Verzetti, Antoine Caillon, Qingqing Huang, Aren Jansen, Adam Roberts, Marco Tagliasacchi, Matt Sharifi, Neil Zeghidour, Christian Frank*

Google Research

**Abstract** We introduce MusicLM, a model generating high-fidelity music from text descriptions such as *"a calming violin melody backed by a distorted guitar riff"*. MusicLM casts the process of conditional music generation as a hierarchical sequence-to-sequence modeling task, and it generates music at 24 kHz that remains consistent over several minutes. Our experiments show that MusicLM outperforms previous systems both in audio quality and adherence to the text description. Moreover, we demonstrate that MusicLM can be conditioned on both text and a melody in that it can transform whistled and hummed melodies according to the style described in a text caption. To support future research, we publicly release MusicCaps, a dataset composed of 5.5k music-text pairs, with rich text descriptions provided by human experts.

# Dificultades técnicas / Resumen

- Estructura musical: temporal y con larga memoria (coherencia a largo plazo, difícil de conseguir)
- Gran volumen de datos necesario para el entrenamiento
- Datasets MIDI reducidos, de calidad variable, poco anotados
- Oído muy sensible a irregularidades /errores (en el caso de audio)
- Tiempo real (e.g. acompañamiento) impone muchas restricciones



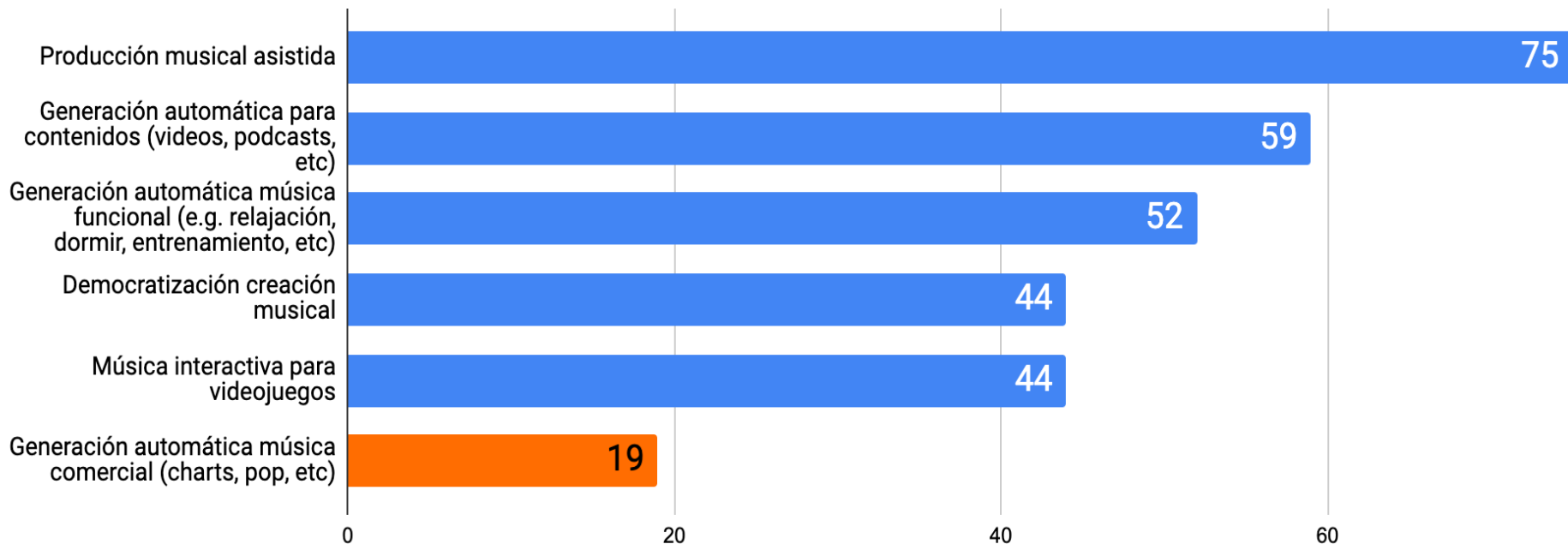


# Mirando al futuro

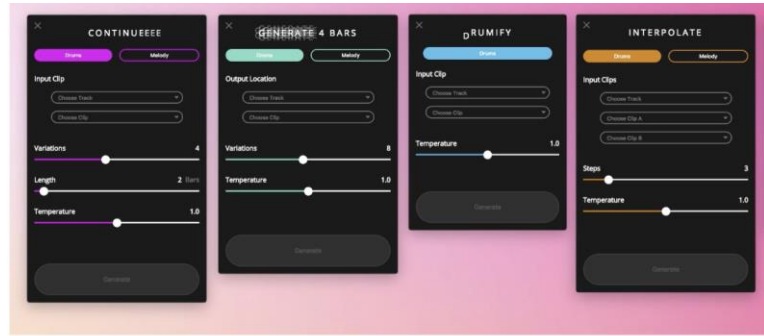
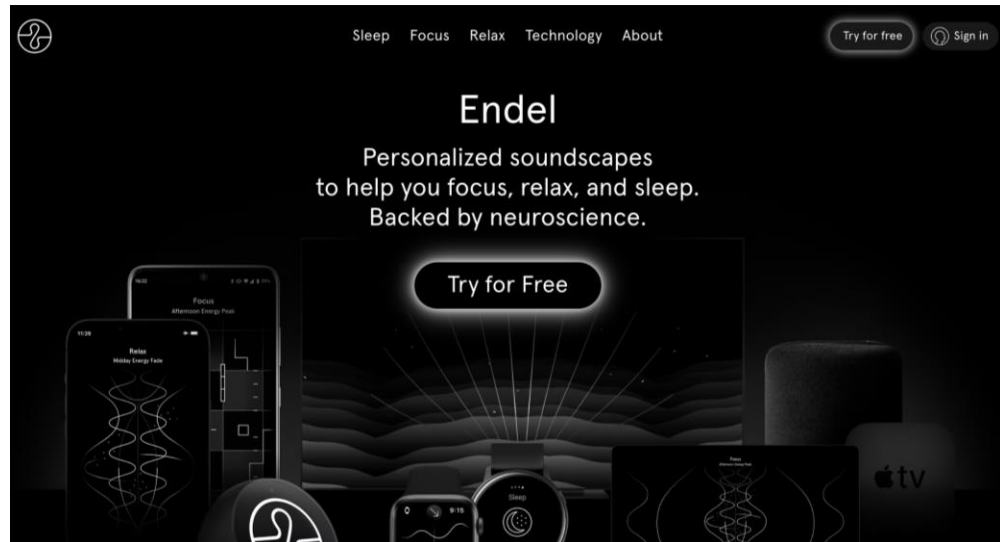
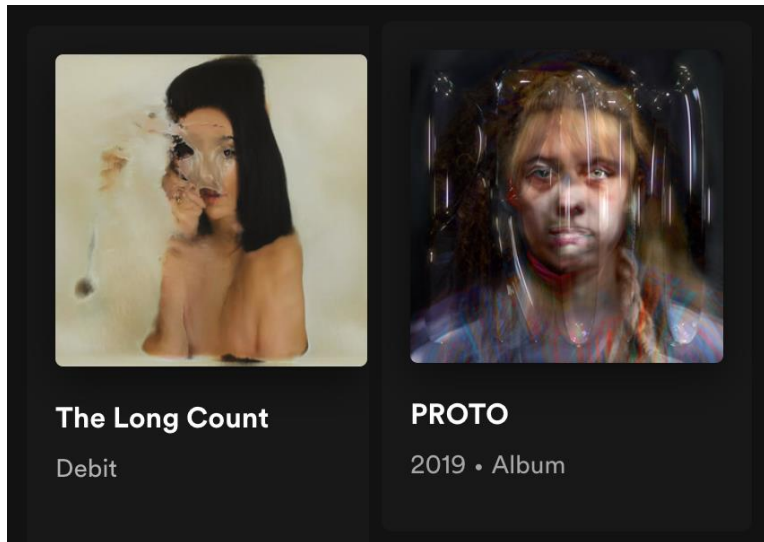
## ¿Hacia donde nos dirigimos?



# Encuesta : (1) Usos de la IA mus. más exitosos en breve?

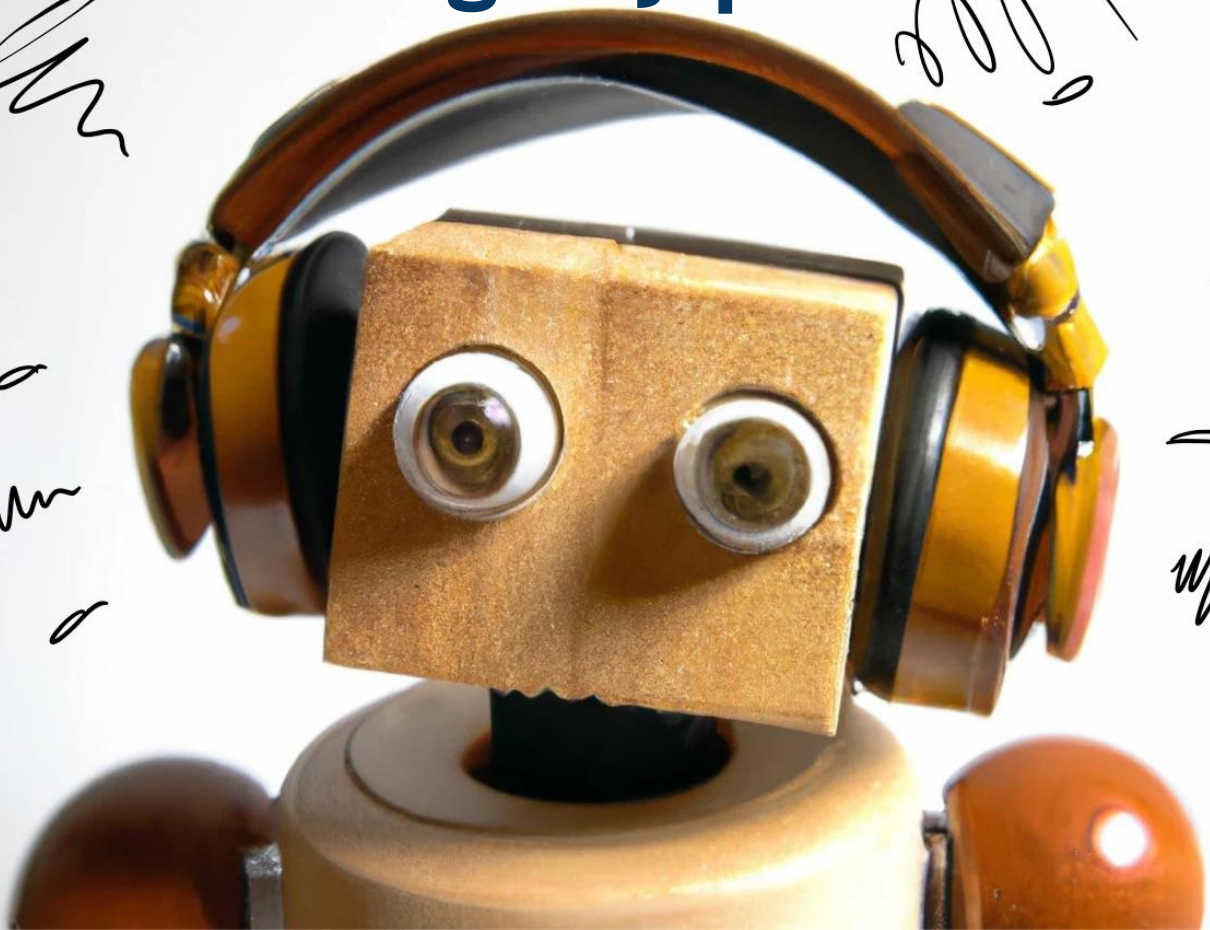


# Aplicaciones actuales



- **Herramientas profesionales** de mezcla, masterizado (y tareas más sistemáticas). Herramientas/plugins de ayuda a la composición
- **Generación de música funcional** (para videos, corporativa, relajación, dormir, training, etc.)
- **Generación interactiva** (videojuegos, training, etc.) menos desarrollado


# Riesgos y problemas




# Democratización de la creación ... ... es esto lo que buscamos?!



 Home



 Create

 Library

Sign In

Sign Up

## Create generative music Share it with the world

-  Make original songs in seconds, even if you've never made music before
-  Submit your songs to streaming platforms and get paid when people listen
-  Join a global community of artists empowered by Boomy AI

Create your song

# THE MUSIC INDUSTRY ISN'T READY FOR THE AI REVOLUTION.

1.7K  
SHARES



NOVEMBER 29, 2022

BY MUSIC BUSINESS WORLDWIDE

"UP UNTIL NOW, NO ONE COMPANY HAS BEEN ABLE TO MASS-CREATE AND MARKET ORIGINAL CATALOG THAT IS OF SUCH A SCALE AND A PREMIUM STANDARD THAT IT CAN COMPETE WITH THE MAJORS IN A MEANINGFUL WAY. IS AI ABOUT TO CHANGE THAT?"

The following MBW op/ed comes from Stef Van Vugt (pictured), the founder and CEO of Fruits Music, a label-cum-playlist company that has racked up tens of billions of plays on [Spotify](#) and other services. Netherlands-headquartered [Fruits Music](#) is home to playlist brands such as Dance Fruits and LoFi Fruits, and is represented by its famous Melon brand identity. You can listen to a popular MBW podcast interview with Stef Van Vugt from earlier this year [through here](#).

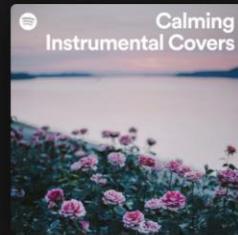
## THE DSP QUESTION

What if western DSPs began populating their own platforms and editorial playlists with AI-generated music created by themselves? This would leave even less room for human creativity to be displayed on these services, and give big tech even more control over what's being consumed.

This already appears to be happening in China. A [recent article](#) in [Music Business Worldwide](#) stated that Tencent Music's streaming services now host over 1,000 songs with AI-generated vocals. All of these tracks have been created by Tencent Music's own AI technology, and are cumulatively amassing millions, if not billions, of plays. If this had happened on western DSPs like [Spotify](#) or [Apple Music](#), I believe the major record companies would have reacted very differently.

**"TENCENT MUSIC'S STREAMING SERVICES NOW HOST OVER 1,000 SONGS WITH AI-GENERATED VOCALS."**

## Relaxing Instrumental Music



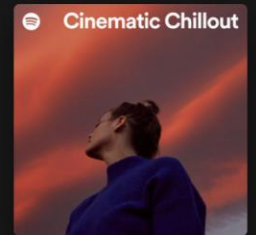
### Calming Instrumental...

Gentle instrumental covers of known songs.



### Atmospheric Piano

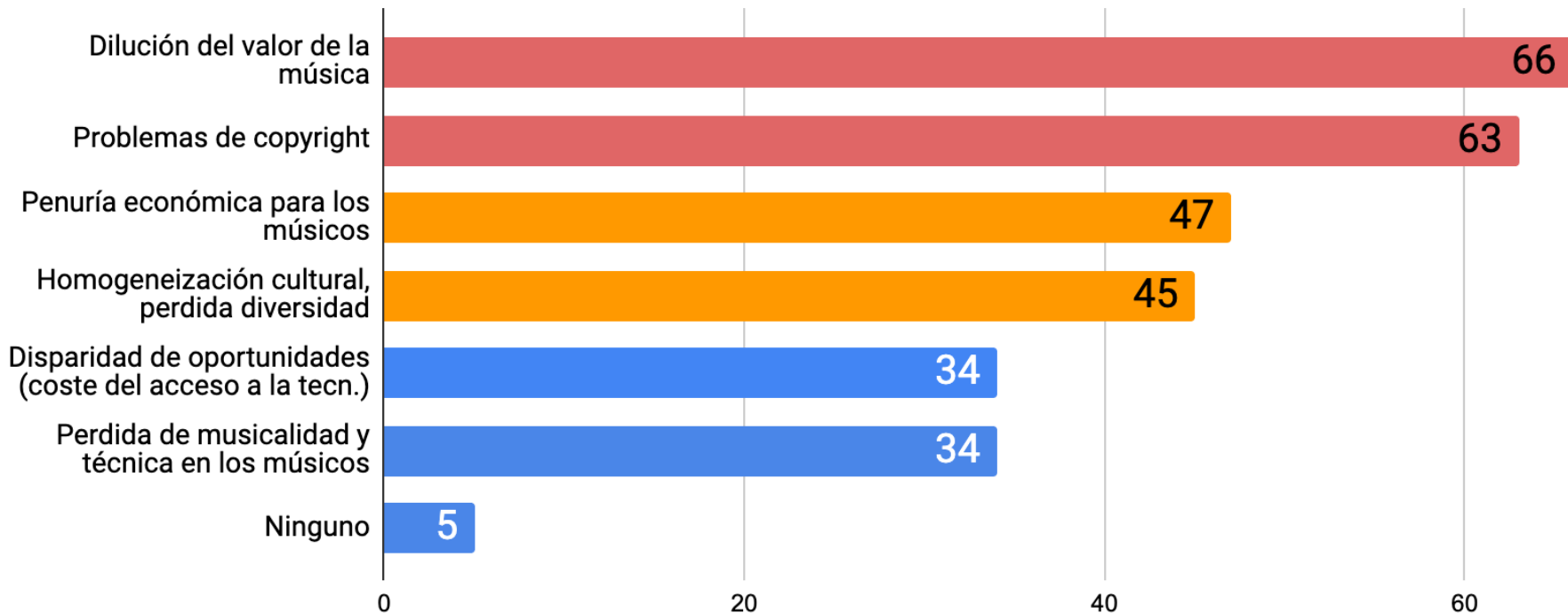
Atmospheric and emotional piano pieces.



### Cinematic Chillout

Relaxing soundtracks to help you escape and...

# Encuesta : (2) ¿Que problemas nos puede traer el uso de la IA en la creación musical?

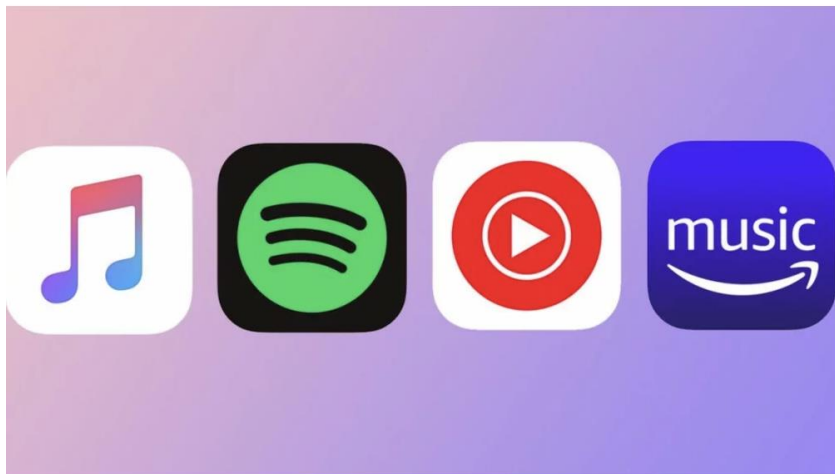


# Precarización Creadores

NEWS

Just 0.4 percent of artists in the UK make a living from streaming plays of their music, according to a new study.

*With the vast majority signed to major labels*



- El modelo *Digital Streaming Platforms* (DSP) ya no es sostenible
- La sobreproducción que se avecina no hará sino empeorar las cosas para las creadores
- El modelo debe cambiar
- **Blockchain?**

**YES, BLOCKCHAIN MIGHT WELL REVOLUTIONISE MUSIC ROYALTIES. NO, IT'S NOT GOING TO HAPPEN ANYTIME SOON.**

721  
SHARES



🇬🇧 APRIL 12, 2022

BY MUSIC BUSINESS WORLDWIDE

MARTIN GUTTRIDGE-HEWITT  
Wednesday, September 29, 2021 - 13:55

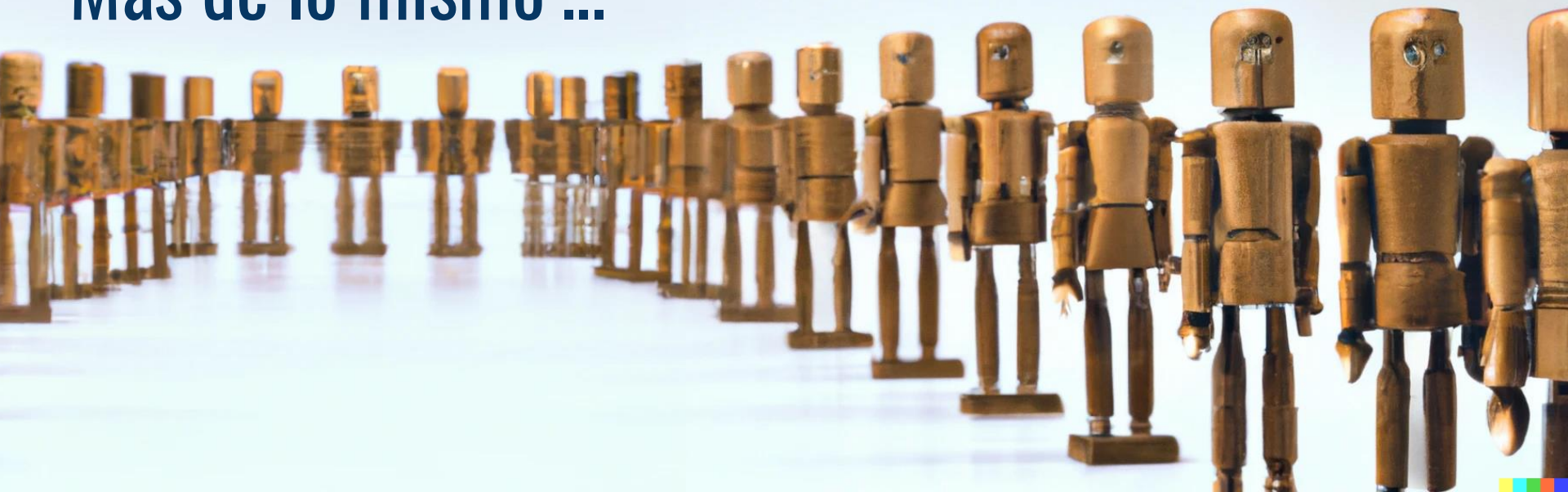


Universitat  
Pompeu Fabra  
Barcelona

MTG  
Music Technology  
Group

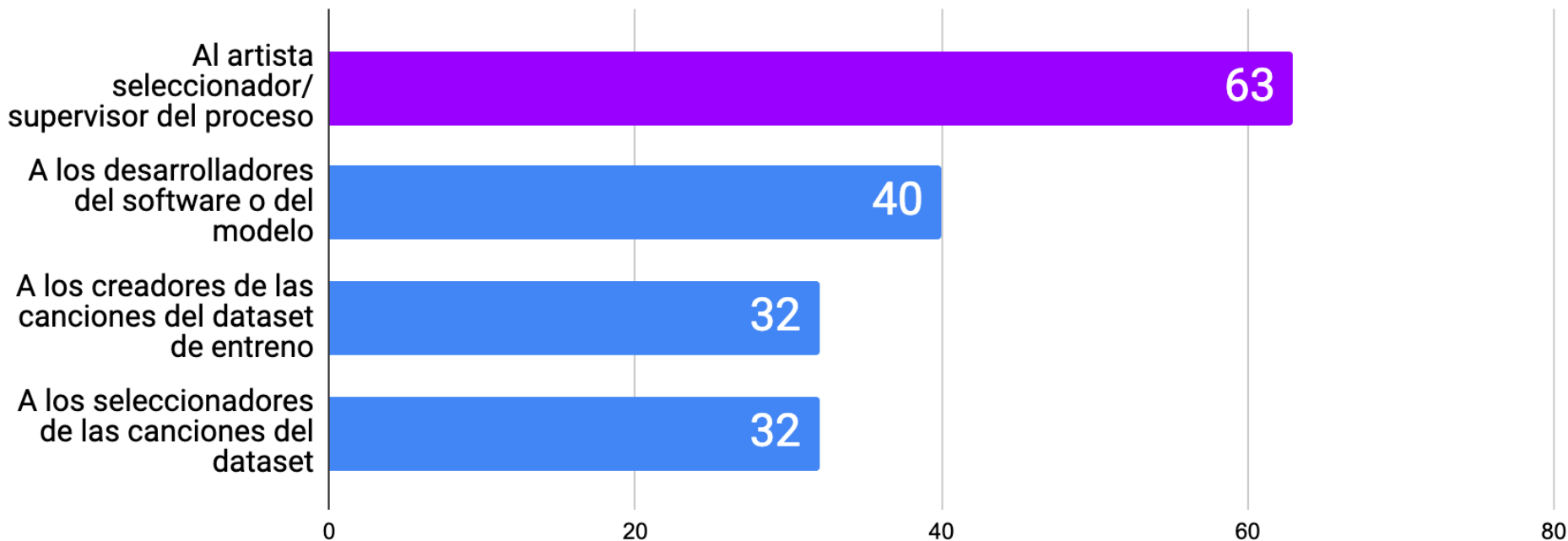


**Homogeneización cultural**  
**Perdida de diversidad**  
**Occidentalización de la música**  
**Más de lo mismo ...**



# Encuesta : (3) Autoría y Copyrights

A quien correspondería la autoría, en un sistema IA que genera música a partir de un dataset de músicas preexistentes?



## Jay Z tries to use copyright strikes to remove deepfaked audio of himself from YouTube



Photo by Alberto E. Rodriguez/Getty Images for The Recording Academy

/ Roc Nation is accusing the creator of 'unlawfully using an AI'

By NICK STATT / @nickstatt  
Apr 29, 2020, 12:38 AM GMT+2 | [0 Comments](#)



### POLÉMICA TECNOLÓGICA

## La comunidad 'anime' se enfurece por la reproducción mediante IA de obra de artistas fallecidos

## Inteligencia artificial generativa: ¿inspiración o plagio?



- Las primeras demandas contra la IA por aprovechar obras humanas para entrenarse avivan el debate sobre su falta de transparencia



Midjourney es una de las herramientas de creación con inteligencia artificial que ha sido demandada por un grupo de artistas (Rafael Enrique/SOPA Images)



MAYTE RIUS  
BARCELONA

23/01/2023 06:00 | Actualizado a  
23/01/2023 10:32



Universitat  
Pompeu Fabra  
Barcelona

MTG  
Music Technology  
Group

Jay Z tries to use copyright strikes to remove deepfaked audio of himself from YouTube



Photo by Alberto E. Rodríguez

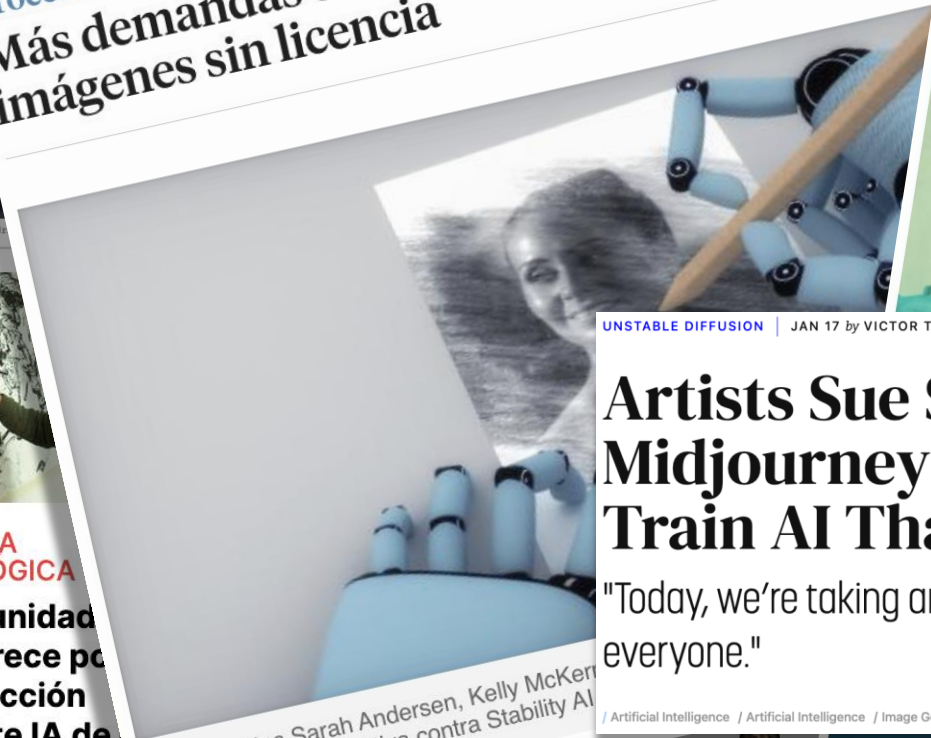


**POLÉMICA TECNOLÓGICA**

La comunidad se enfurece por reproducción mediante IA de obras de artistas fallecidos

Procesos judiciales

Más demandas contra las IA que usan sus imágenes sin licencia



Los artistas Sarah Andersen, Kelly McKernan lanzan una demanda colectiva contra Stability AI

Procedimiento legal

Getty Images asegura que se infringieron derechos de propiedad intelectual



UNSTABLE DIFFUSION | JAN 17 by VICTOR TANGERMANN

Artists Sue Stable Diffusion and Midjourney for Using Their Work to Train AI That Steals Their Jobs

"Today, we're taking another step toward making AI fair and ethical for everyone."

/ Artificial Intelligence / Artificial Intelligence / Image Generator / Stable Diffusion

BA  
23  
23

Imagen generada con la herramienta

# La industria musical deberá reinventarse!

Música terapéutica que se adapte a nuestras constantes biométricas ...

Música personalizada, hecha a medida, para cada momento, para cada uno de nosotros ...



SMARTINGPRO LINE

SMARTING

SMARTFONES

LET'S TALK



THINK. LISTEN. OBSERVE.

## SMARTFONES

We have created a powerful tool for all creative and disruptive minds who want to bring pioneering EEG work to life and apply neuroscience in an everyday setting.



High-quality recording



Sound feature



Smarting Mobi and Smarting Pro compatible



Ayuda

Escucha (inter)activa  
(e.g. compartir escenario VR  
con nuestra banda favorita) ...



Universitat  
Pompeu Fabra  
Barcelona

MTG  
Music Technology  
Group

**Música interactiva !**





# Muchas gracias por la atención !

[sergi.jorda@upf.edu](mailto:sergi.jorda@upf.edu)

Music Technology Group, DTIC-UPF

