

Deep Learning for Music Analysis and Generation

Introduction

Music Information Research



Yi-Hsuan Yang Ph.D.

yhyangtw@ntu.edu.tw

Outline

- **Types of music related research/products**
- Fundamentals of music signal processing
- Types of data

Music Information Retrieval (MIR)

- Intelligent ways to *retrieve*: “content-based search”

White Stripes - Hotel Yorba

This is a **brit poppy, alternative** song that is **not calming** and **not mellow**. It features **male vocal, drum set, distorted electric guitar**, a nice **distorted electric guitar solo**, and **screaming, strong vocals**. It is a song **with high energy** and **with an electric texture** that you might like listen to while **driving**.

Miles Davis - Blue in Green

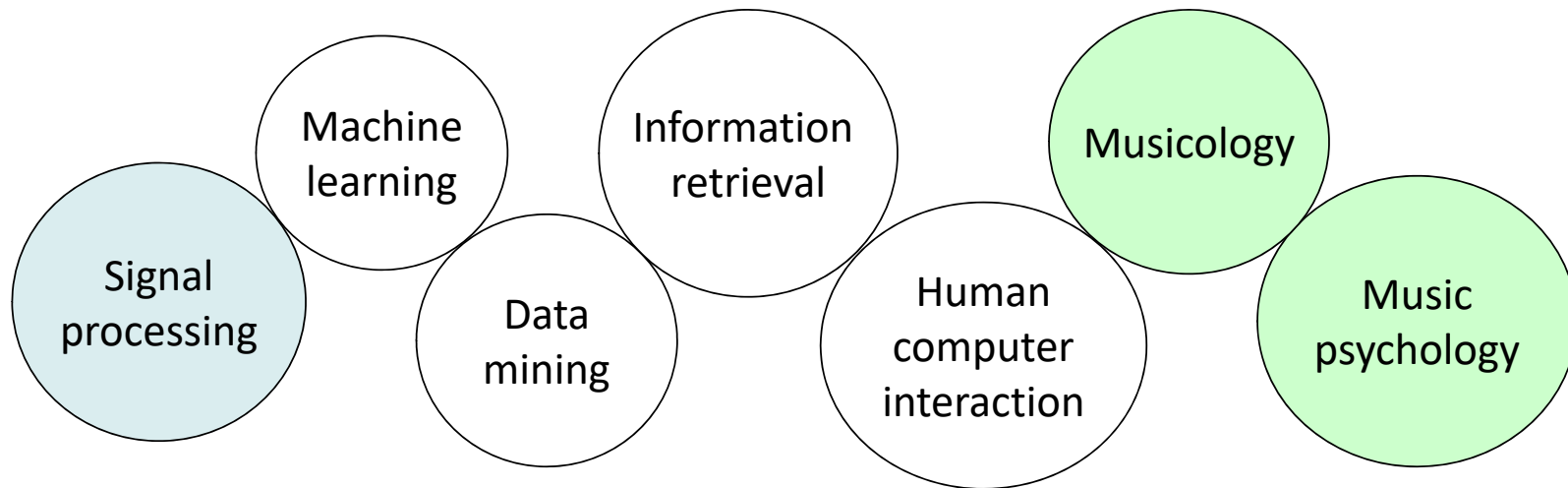
This is a **jazzy, folk** song that is **calming and not arousing**. It features **acoustic guitar, saxophone, piano**, a nice **piano solo**, and **emotional, low-pitched vocals**. It is a **song slow tempo** and with **low energy** that you might like listen to while **reading**.

Depeche Mode - World In My Eyes

This is a **funky, dance pop** song that is **arousing** and **not tender**. It features **male vocal, synthesizer, drum machine**, a nice **male vocal solo**, and **altered with effects, strong vocals**. It is a song with a **synthesized texture** and that was **recorded in studio** that you might like listen to while **at a party**.

Music Information “Research” (MIR)

- Intelligent ways to analyze, retrieve, and create music
- Music + signal processing + machine learning



Types of Music Related Research/Products

- Intelligent ways to analyze, retrieve, and create music

1. Music information **analysis**

music → features

2. Music information **retrieval**

query → music

3. Music **generation**

X → music

Types of Music Related Research/Products

1. Music information “analysis”



automatic page turner

interactive
concert



automatic
Karaoke scoring

1. Music information “analysis”



Types of Music Related Research/Products

2. Music information “retrieval”

- Search
 - through [keywords/labels](#) (genre, instrument, emotion)



Types of Music Related Research/Products

2. Music information “retrieval”

- Search
 - through [keywords/labels](#) (genre, instrument, emotion)

musical event localization



J.-Y. Liu and Y.-H. Yang, "Event localization in music auto-tagging," *MM* 2016

Types of Music Related Research/Products

2. Music information “retrieval”

- Search
 - through keywords/labels (genre, instrument, emotion)
 - through **audio examples** (humming, audio recording)



Types of Music Related Research/Products

2. Music information “retrieval”

- Search
 - through keywords/labels (genre, instrument, emotion)
 - through **audio examples** (humming, audio recording)



Justin Salamon @justin_salamon · 12 小時

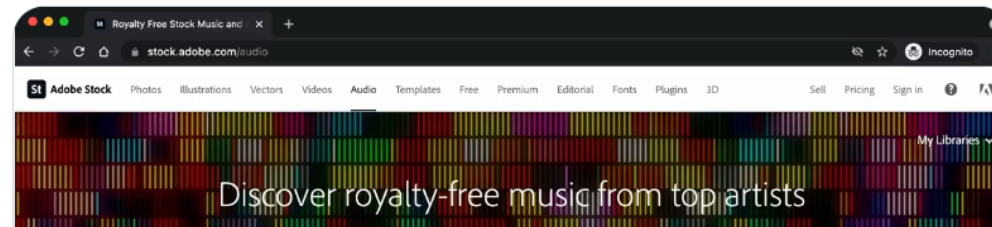


VERY excited about the announcement of Find Similar Audio at [#AdobeMAX](#) 🎵 today!

Find similar music just by uploading an example 🎵🤖



You can try it out here: stock.adobe.com/audio

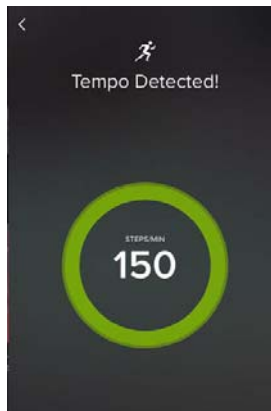


Types of Music Related Research/Products

2. Music information “retrieval”

- **Match**

- to match 1) a video clip, 2) a photo slideshow, 3) a song lyrics, or 4) a given context
- cross-domain retrieval



史無前例！Spotify推出Running功能幫你
搭配跑步歌單！

TEXT/EVA | 2015-05-27

Google Challenge: Automatic Music Video Generation

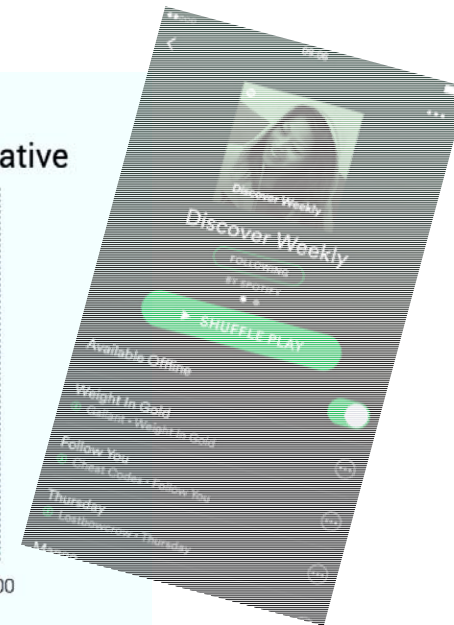
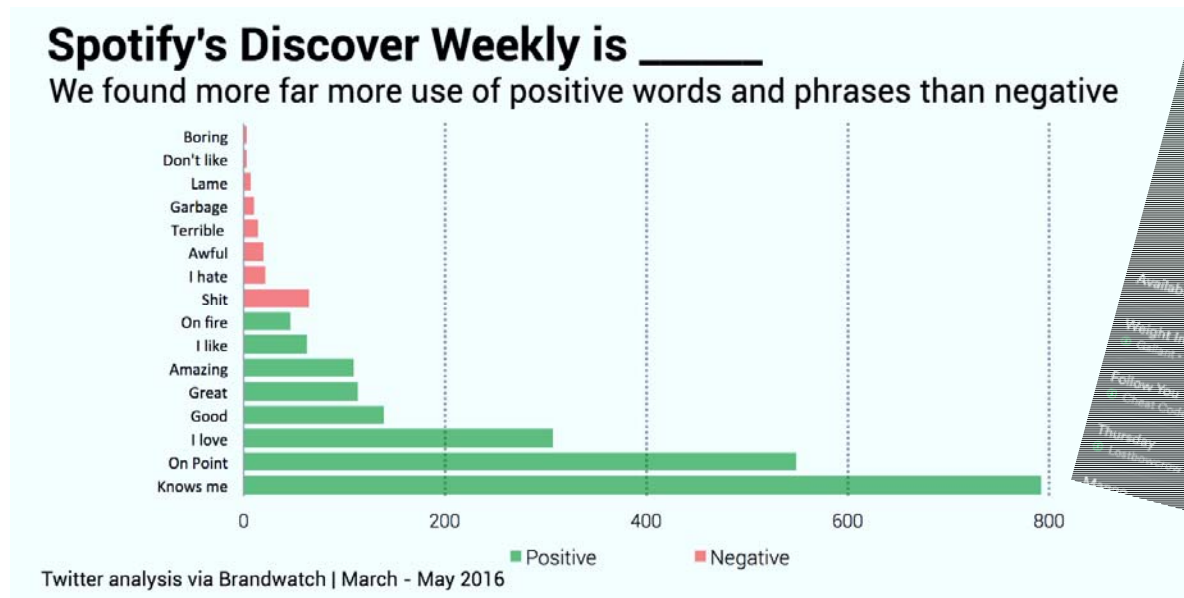
Visual and audio are the two key elements of an entertaining video. A befitting soundtrack makes a video compelling and likewise Lady Gaga's music videos greatly enhance her songs.



Types of Music Related Research/Products

2. Music information “retrieval”

- **Discover**
 - recommendation: diversity, serendipity, explanations



Types of Music Related Research/Products

2. Music information “retrieval”

- **Discover**
 - recommendation: diversity, serendipity, explanations

2千萬首歌想聽哪一首？KKBOX運用機器學習隨時告訴你

隨時找出使用者想聽的歌，才能黏住使用者願意留下來聽歌，單靠人工還不夠，KKBOX依靠機器學習技術即時修正推薦模式，給你現在就會想聽的歌

文/ 李建興 | 2015-07-06 發表

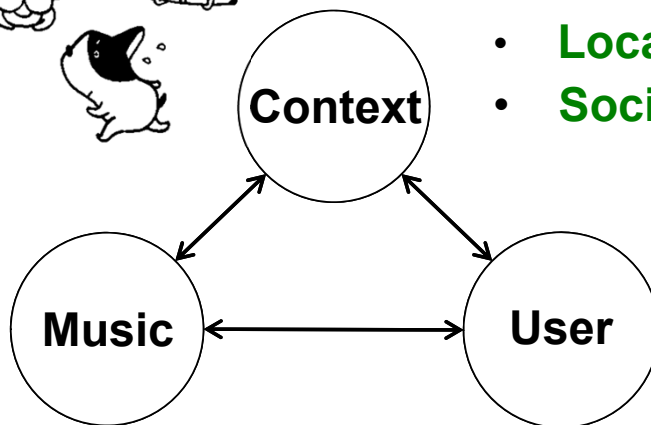
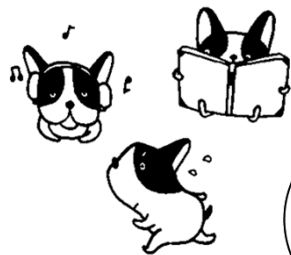
讚 2.9 萬 按讚加入iThome粉絲團 讚 832 分享 G+ 8



Types of Music Related Research/Products

2. Music information “retrieval”

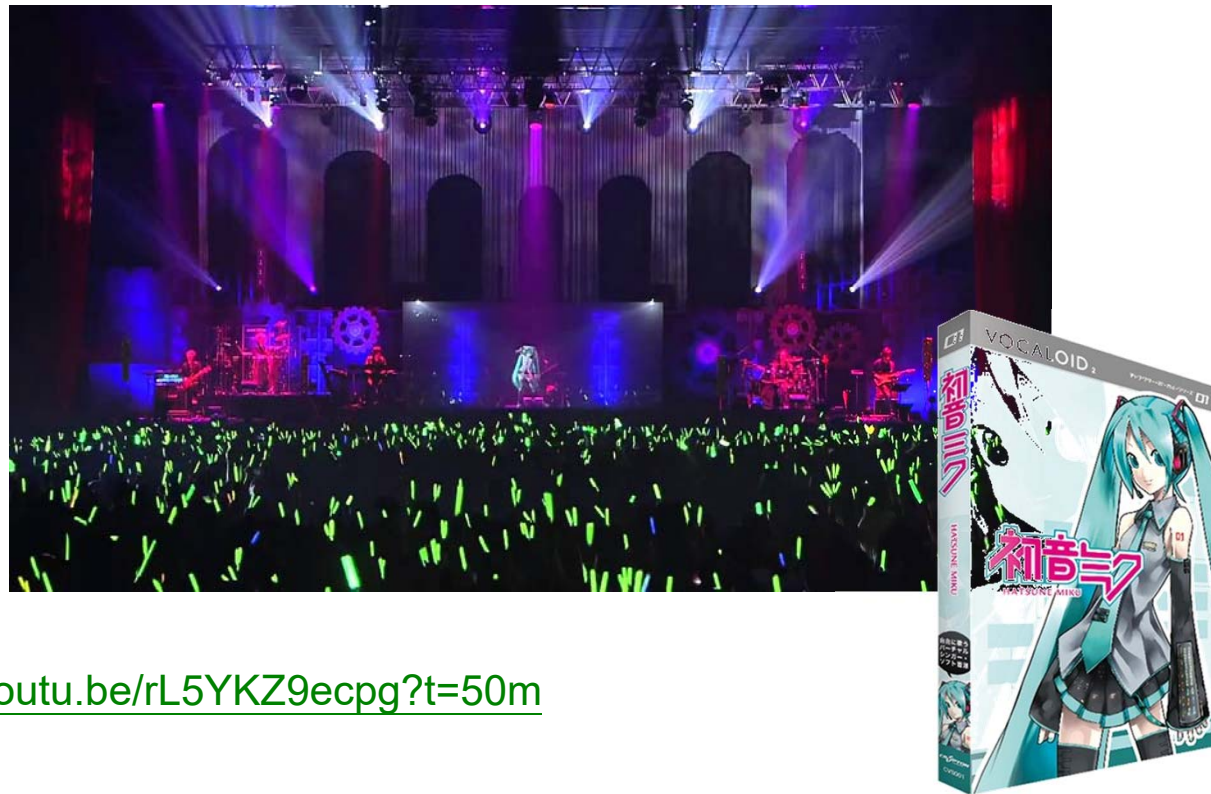
- **Discover**
 - **Context-aware Music Recommendation**



- **Activity**: driving, studying, working, walking
 - **Mood**: happy, sad, angry, relaxed
 - **Location**: home, work, public place
 - **Social company**: alone, w/ friends, w/ strangers
-
- age
 - gender
 - personality
 - cultural background
 - musical background

Types of Music Related Research/Products

3. Music creation



<https://youtu.be/rL5YKZ9ecpg?t=50m>

Types of Music Related Research/Products

1. Music information analysis

- Education, data visualization

2. Music information retrieval

- **Search:** through keywords (genre, instrument, emotion) or audio examples (humming or audio recording)
- **Match:** cross domain retrieval
- **Discover:** recommendation

3. Music creation

Conclusion: Music Information Research

- Intelligent ways to analyze, retrieve, and create music

1. Music information **analysis**

music \rightarrow features

- machine learning
- signal processing
- musicology

2. Music information **retrieval**

query \rightarrow music

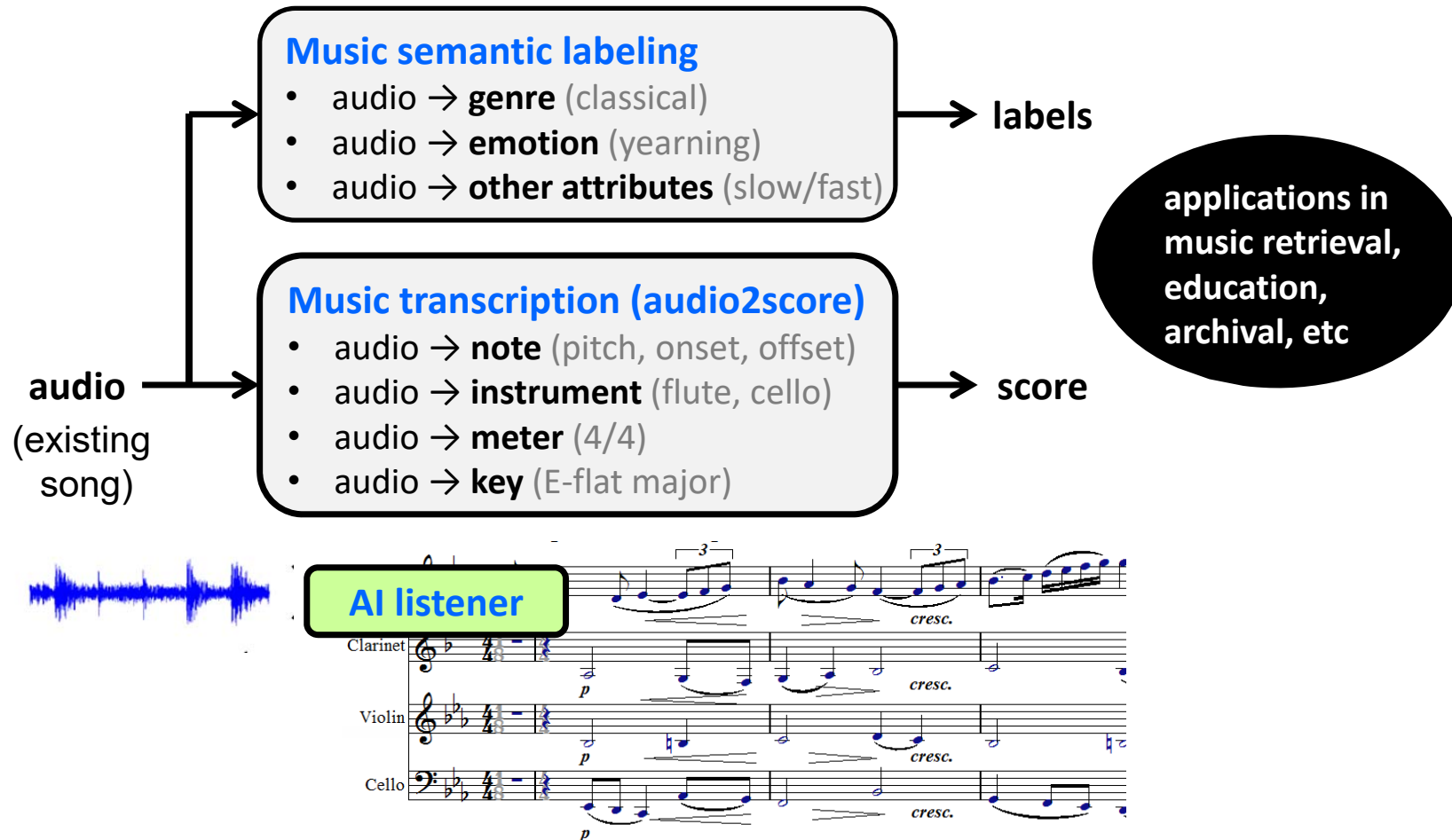
- machine learning
- info. Retrieval
- UI design

3. Music **generation**

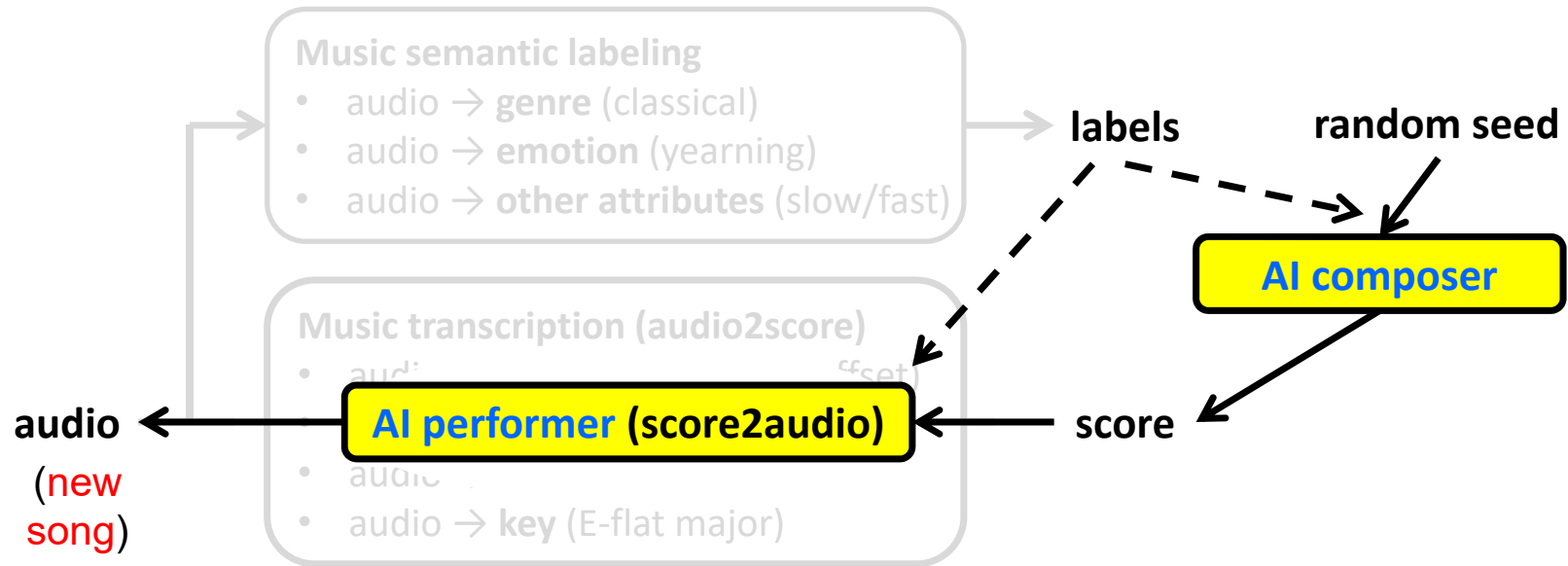
X \rightarrow music

- AI
- musicology

ML in Music: “Music Info Retrieval/Analysis”

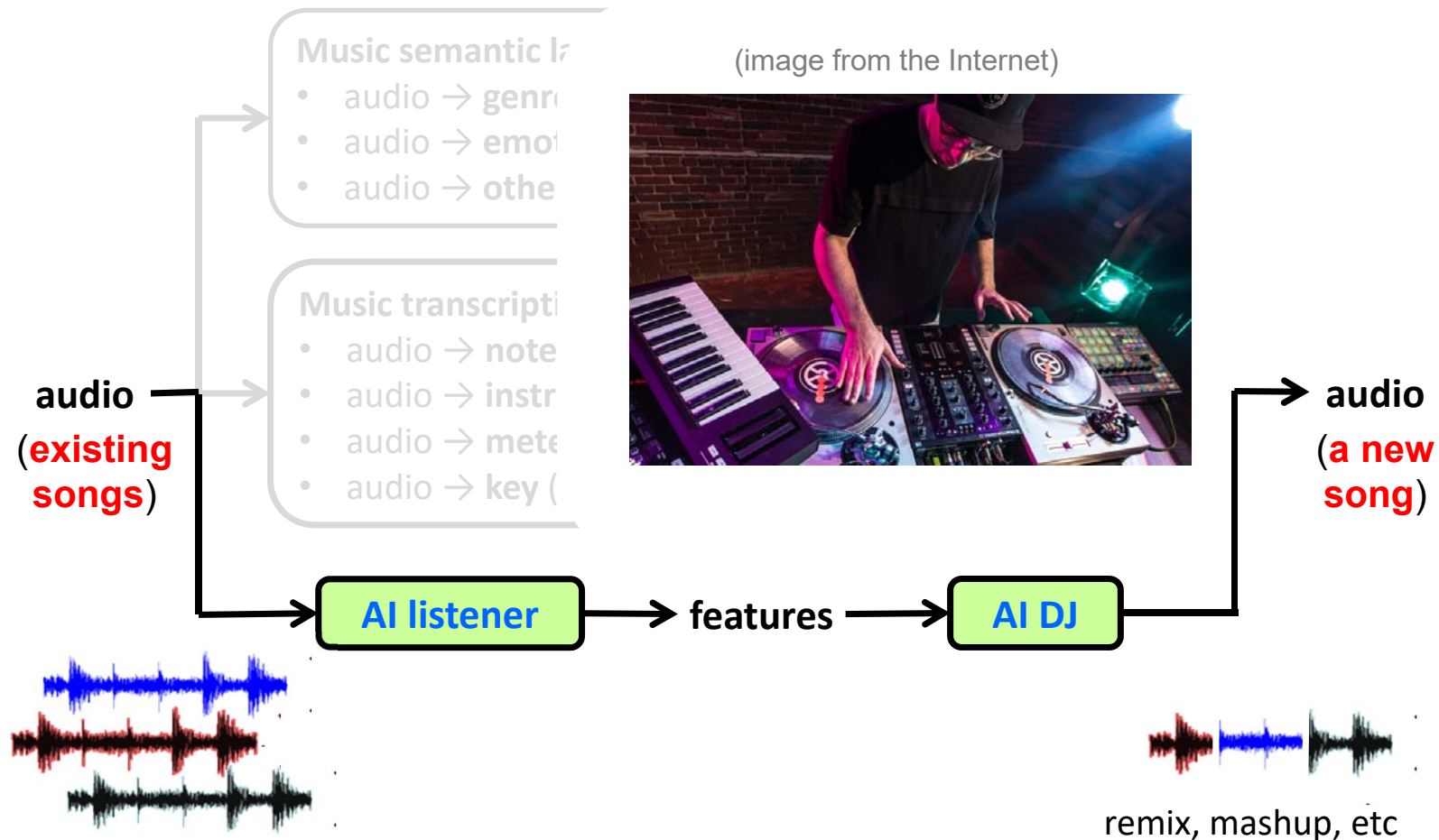


ML in Music: “Music Generation/Synthesis”

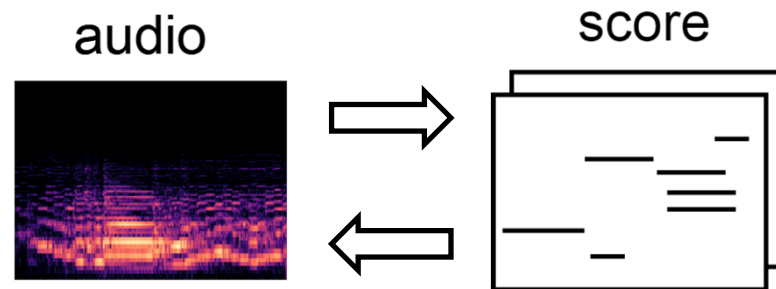


A musical score for four instruments: Flute, Clarinet, Violin, and Cello. The score is written in 4/4 time with a key signature of two flats (B-flat and E-flat). The Flute part features a melodic line with triplets and a crescendo. The Clarinet, Violin, and Cello parts provide harmonic support, each starting with a piano (*p*) dynamic and following a similar crescendo pattern.

ML in Music: “Music Generation/Synthesis”



Music AI Research



- Topics
 - **audio** → **audio**: signal processing
 - **audio** → **score**: transcription
 - **score** → **score**: composition
 - **score** → **audio**: synthesis
 - **audio** → **knowledge**: audio analysis
 - **score** → **knowledge**: symbolic-domain analysis

Outline

- Types of music related research/products
- **Fundamentals of music signal processing**
- Types of data

Fundamentals of Music Signal Processing

- **Pitch**: which notes are played?
- **Rhythm**: how fast?
- **Timbre**: which instrument(s)?

Mozart's *Variationen* (1st phrase)



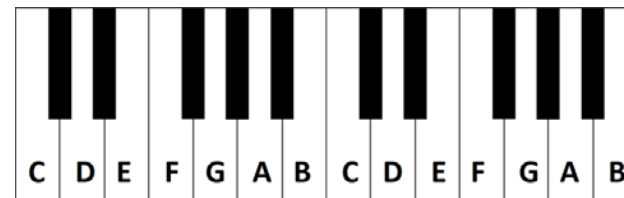
Music Information Analysis

- music → features

melody



1. pitch
2. onset, offset
3. tempo



Music Information Analysis

- music → features



accompaniment

4. chord

演唱 光 良 童 话 作词:光 良 作曲:光 良

1 = G 4/4

(3̣ - - 3̣. 2̣ | 2̣. 1̣ 1̣ - 1̣. 7̣ | 7̣. 6̣ 6̣ - 6̣. 5̣ | 5̣ -) 0̣ 5̣ 1̣ 7̣ |

忘 了 有

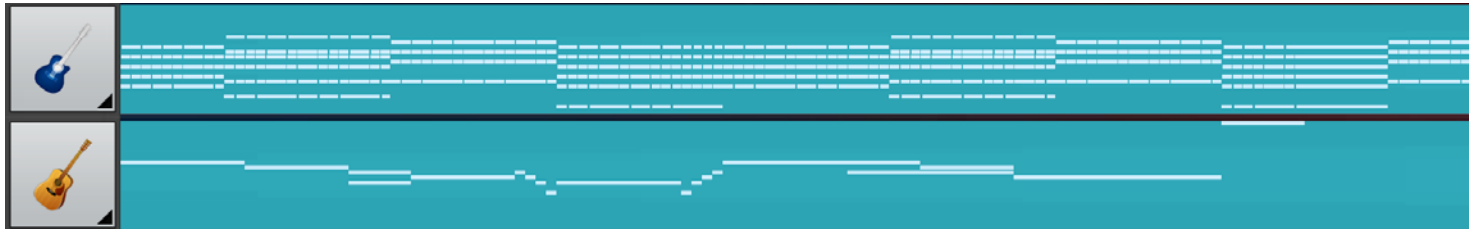
G Em C D

1̣. 5̣ 5̣ 0̣ 5̣ 1̣ 7̣ | 1̣. 5̣ 5̣ 0̣ 5̣ 1̣ 7̣ | 1̣ 0̣ 1̣ 7̣ 6̣ 6̣ 5̣ | 5̣ - 0̣ 5̣ 1̣ 7̣ |

多 久 再 没 听 到 你 对 我 说 你 最 爱 的 故 事 我 想 了

Music Information Analysis

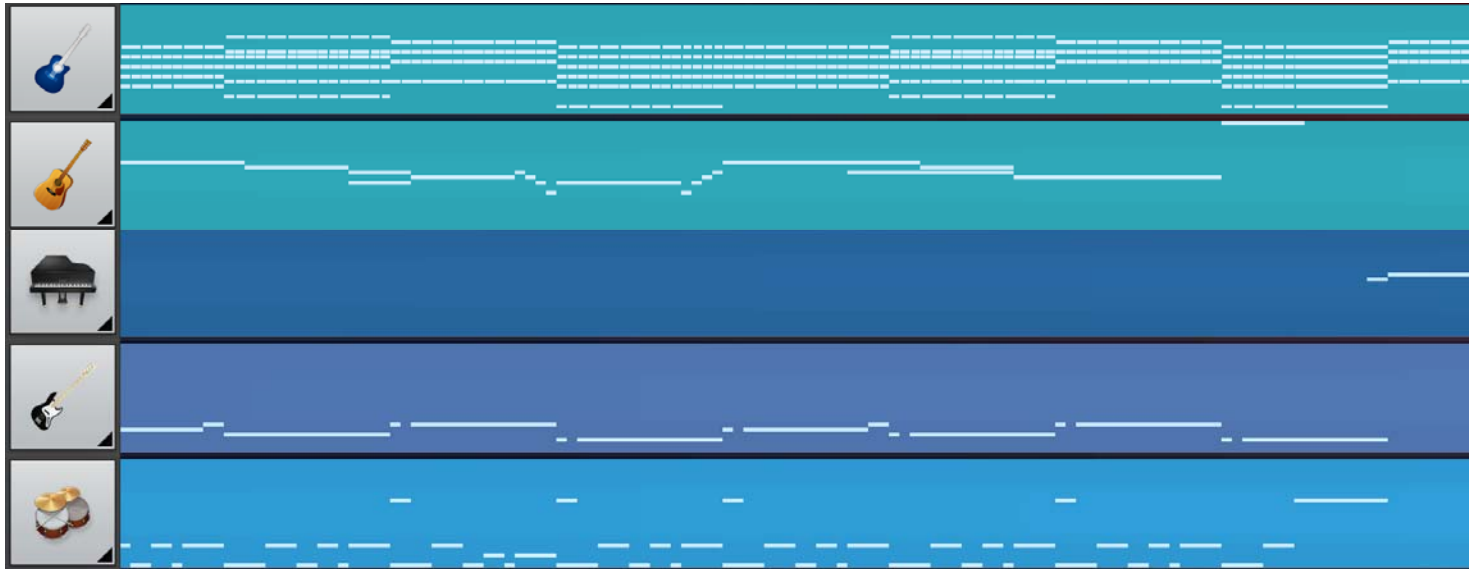
- music → features



5. instruments (**timbre**)

Music Information Analysis

- music → features



6. source separation
7. key, beat, downbeat, meter

Music Information Analysis

- music → features

- 8. Semantic description

- genre: pop, classical, jazz, rock, R&B, “Tai,” “aboriginal”
 - emotion: happy, angry, sad, relaxed
 - usage: at party, working, driving, reading, sleeping, romance
 - theme: lonely, breakup, celebration, in love, friend, battle
 - vocal timbre: aggressive, breathy, duet, emotional, rapping, screaming



genre



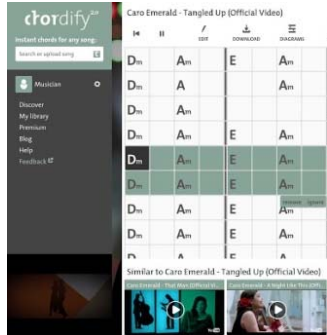


emotion






listening context




Fundamentals of Music Signal Processing

	 <p>page turner</p>	 <p>Karaoke scorer</p>	 <p>chord recognizer</p>
Pitch	♪ ♪ ♪	♪ ♪ ♪	♪ ♪ ♪
Rhythm		♪	
Timbre		♪ ♪	

Fundamentals of Music Signal Processing

	 <p>content ID</p>	 <p>instrument classifier</p>	 <p>Spotify running</p>
Pitch	♪♪♪	♪	
Rhythm			♪♪♪
Timbre		♪♪♪	♪

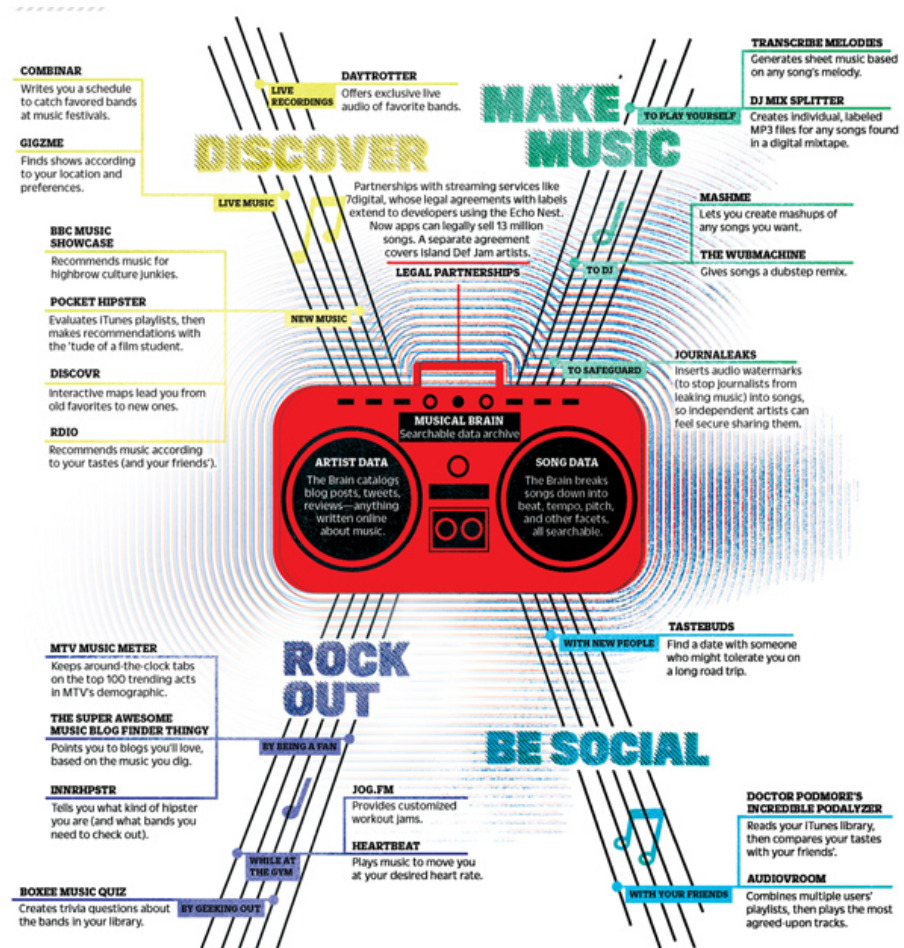
Fundamentals of Music Signal Processing

	 <p>music emotion or genre recognizer</p>	 <p>similarity search or recommendation</p>	 <p>automatic music video generation</p>
Pitch	♪ ♪ ♪	♪ ♪ ♪	♪ ♪ ♪
Rhythm	♪ ♪ ♪	♪ ♪ ♪	♪ ♪ ♪
Timbre	♪ ♪ ♪	♪ ♪ ♪	♪ ♪ ♪

Fundamentals of Music Signal Processing



- **Listens to music**
tempo, instrumentation, key, time signature, energy, harmonic & timbral structures
- **Reads about music**
lyrics, blog posts, reviews, playlists and discussion forums
- **Learns about trends**
online music behavior — who's talking about which artists this week, what songs are being streamed or downloaded
- **Not everything is in audio**

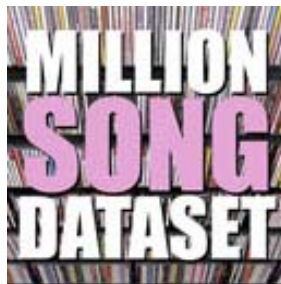


Outline

- Types of music related research/products
- Fundamentals of music signal processing
- **Types of data**

Types of Data

- **Music audio data**
 - not sharable due to copyright issues and business interest
 - however, **audio features** can be shared
 - or, start with **copyright free music**



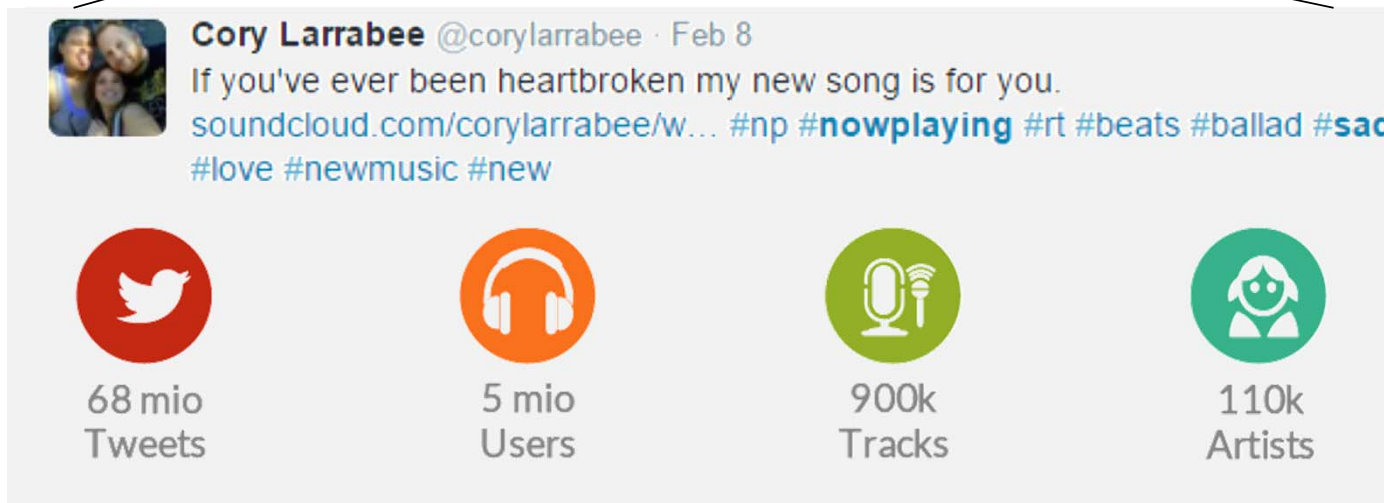
free music
archive

A screenshot of the FMA (Free Music Archive) website. The header includes the FMA logo, a "Sign Up / Log-In" link, and a "Music Search" bar. Below the header, there are navigation tabs: "BY GENRE", "PARTICIPATE!", and "WHAT IS THE FMA?". The "BY GENRE" tab is active, showing a grid of genre categories: Country, Electronic, Experimental, Folk, Hip-Hop, International, Jazz, Old-Time / Historic, Pop, Rock, Soul R&B, and Spoken Word. Below the grid, there is a "Genre Tracks" section with a table of tracks. The table has columns for Artist, Track, Album, Genre, and a "Most Interesting" rating. The tracks listed are: Max Tundra's "Lystine", Lucky Dragons' "Mercy", 2 Tall's "INFO INFINITY 'our' loop 1", 8 Bit Weapon's "Micro Reaction", and Kevin Blechman's "Monster".

ARTIST	TRACK	ALBUM	GENRE	Most Interesting
Max Tundra	"Lystine"	"Live at WFJH on Talk's"	ELECTRONIC, EXPERIMENTAL PO	+ -
Lucky Dragons	"Mercy"	"Dark Falcon"	ELECTRONIC, AUDIO COLLAGE	+ -
2 Tall	"INFO INFINITY 'our' loop 1"	"INFO INFINITY: an explor	AMBIENT ELECTRONIC	+ -
8 Bit Weapon	"Micro Reaction"	"8 Bit Weapon Live at MI	ELECTRONIC	+ -
Kevin Blechman	"Monster"	"Live@PDSin"	ELECTRONIC, EXPERIMENTAL PO	+ -

Types of Data

- **Music listening data**
 - from social platforms via e.g., last.fm API, Spotify API
 - from Twitter: [#nowplaying dataset](#)



Types of Data

- **Big music text data**
 - score, MIDI, lyrics, review, playlist, tags, Wikipedia, etc
 - not everything is in audio
 - some of them are easier to get from non-audio data

Welcome to the **musiXmatch dataset**, the official lyrics collection of the Million Song Dataset.

The WASABI Song Corpus is a large corpus of songs enriched with metadata extracted from music databases on the Web, and resulting from the processing of song lyrics and from audio analysis.

MIR Datasets

<https://mirdata.readthedocs.io/en/latest/>

Supported Datasets and Annotations

★ Dataset Quick Reference ★

This table is provided as a guide for users to select appropriate datasets. The list of annotations omits some metadata for brevity, and we document the dataset's primary annotations only. The number of tracks indicates the number of unique "tracks" in a dataset, but it may not reflect the actual size or diversity of a dataset, as tracks can vary greatly in length (from a few seconds to a few minutes), and may be homogeneous.

"Downloadable" possible values:

✅ : Freely downloadable

🔑 : Available upon request

💻 : Do-it-yourself download

📊 : Features only

❌ : Not available