PAT 463/563 (Fall 2026)

Music & AI

Lecture 1: Introduction

Instructor: Hao-Wen (Herman) Dong

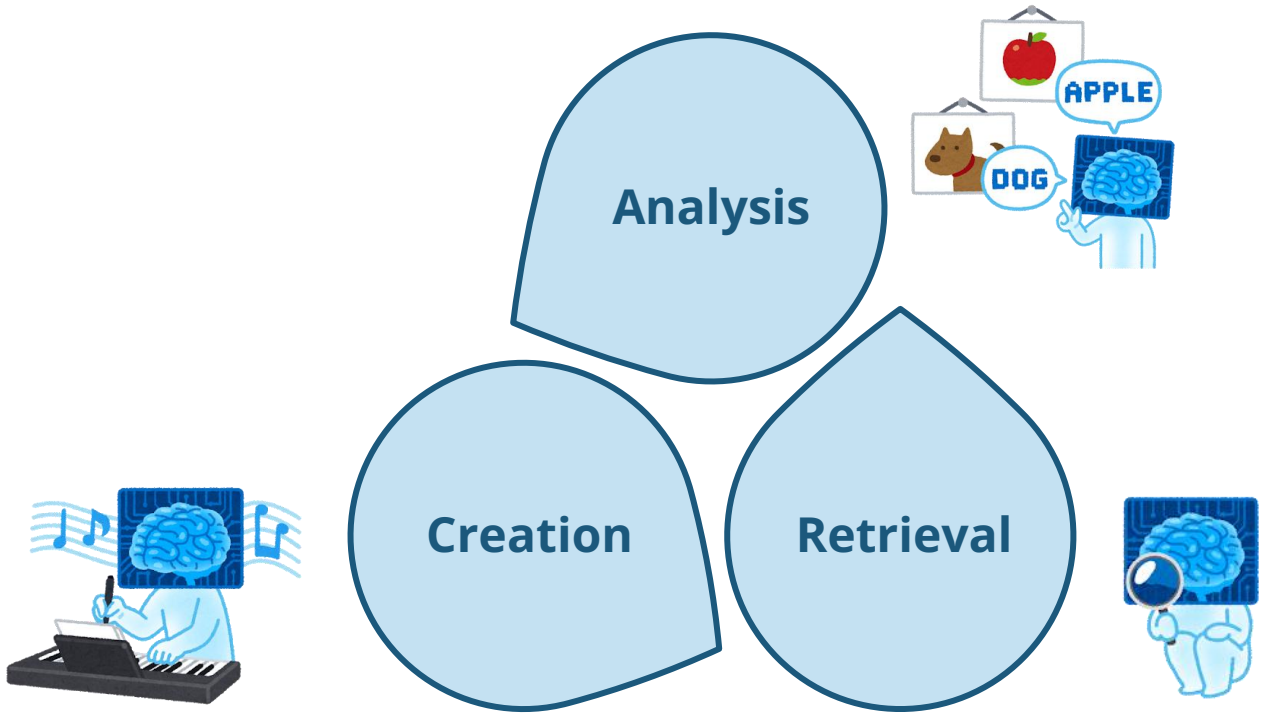
UNIVERSITY OF MICHIGAN

Music & AI

Music & **AI**

What is this course all about?

An introduction to the **emerging field of AI music**. This course introduces students to AI's applications in music from **analysis**, **creation**, **retrieval** to **processing**.



Topics

- **Analysis**
  - Source separation, beat tracking, key/chord recognition, genre/mood classification, transcription, optical music recognition, structural analysis …
- **Creation**
  - Automatic music composition, arrangement, performance synthesis, improvisation …
- **Retrieval & Processing**
  - Search, query by humming, recommendation, production, editing, mixing, mastering …
- **Advanced Topics**
  - Interactive tools, singing voice synthesis, sound design, audiovisual tools …

Learning Objectives

- Gain an **overview of the field of AI music**
- Learn **fundamental AI concepts and principles**
- Learn **representative AI tools** for music analysis, creation, retrieval and processing
- Gain hands-on experience on **using AI music tools**

About Me

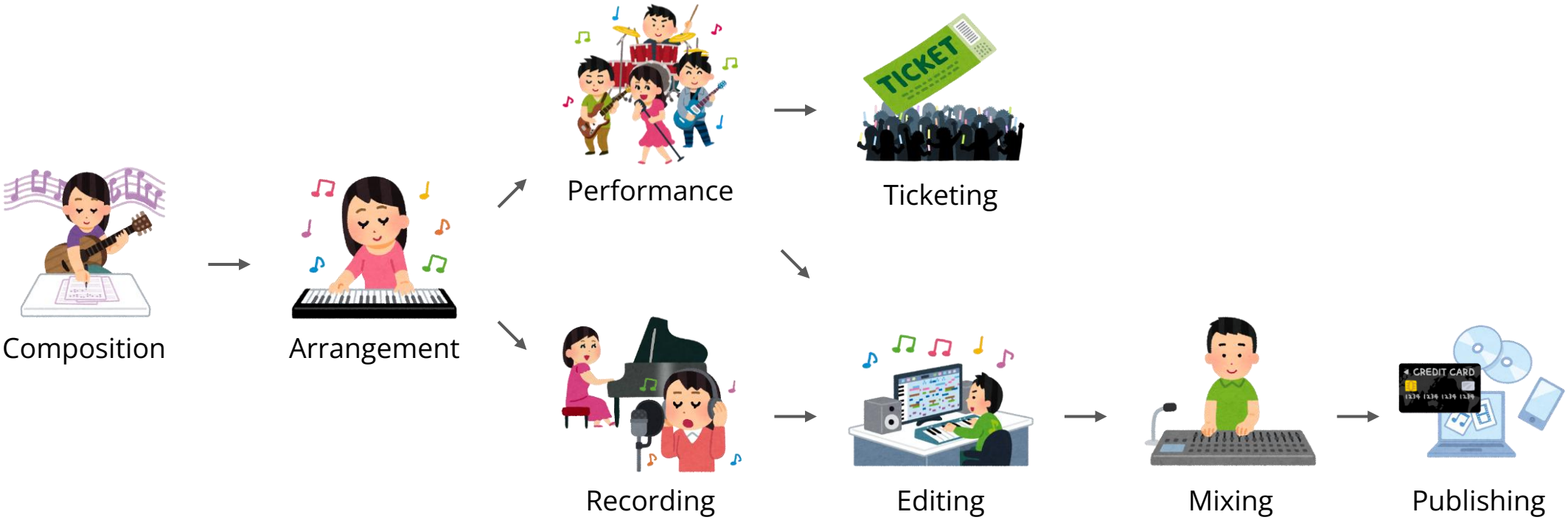
- Hao-Wen (**Herman**) Dong
- Pronouns: he/him
- Email: **hwdong**@umich.edu
- Office: Stearns 131 (15 min walk to the north from Moore)
- Office hours: By appointments (link on the course website)
- Research areas: Generative AI for music, audio, and video

Welcome! Tell Us about Yourself!

- Name
- Pronouns
- Program/year
- What is your **most familiar instrument** (if any)?
- What is your **most familiar programming language** (if any)?
- Have you ever used any **AI Music tools**? Which tools?

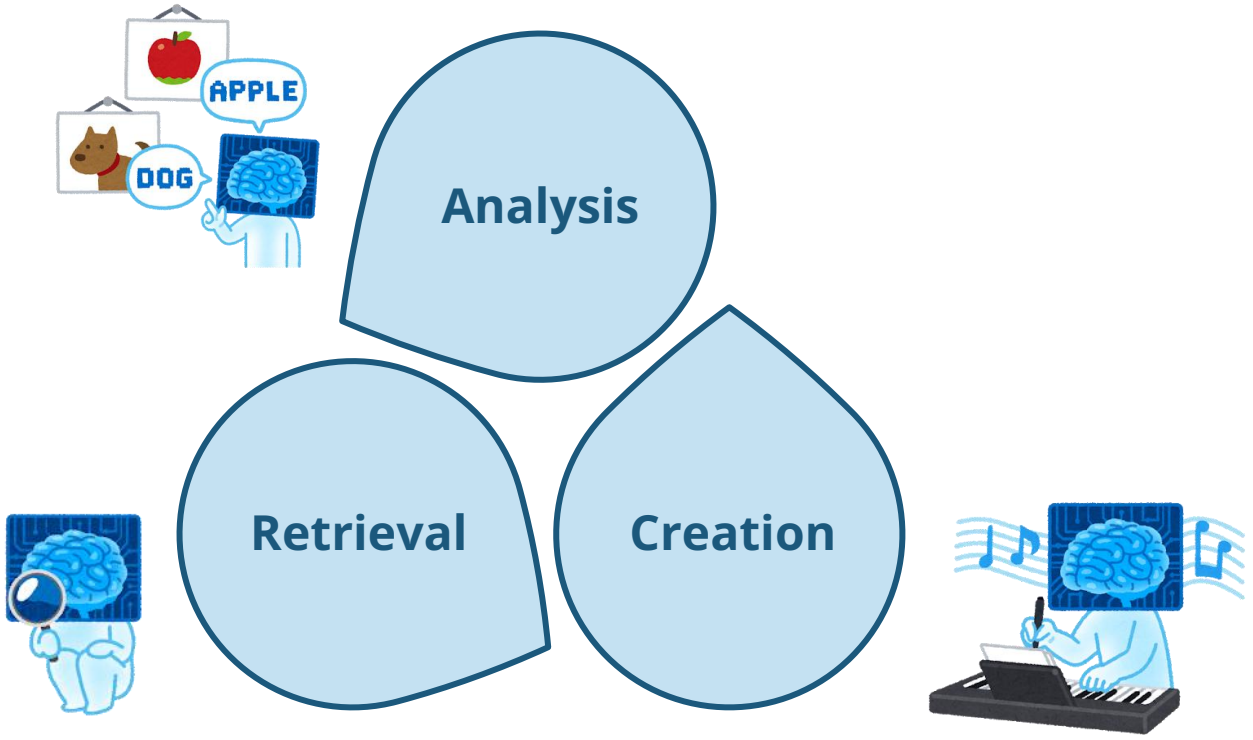
Landscape of AI for Music

An Overly-Simplified Music Production Workflow



Music Information Research (MIR)

- "Intelligent ways to analyze, retrieve and create music" (Yang 2018)



Yang, "Music Information Research," SNHCC, TIGP, lecture notes, April 2018.

Landscape of AI for Music

- **Analysis**
  - Key/chord recognition, genre/mood classification, source separation, beat tracking, automatic music transcription, optical music recognition, structural analysis, computational musicology, …
- **Creation**
  - Automatic music composition, music accompaniment, music arrangement, music style transfer, audio synthesis, singing voice synthesis, lyric generation, robotics improvisation, human-AI music co-creation, …
- **Retrieval & Processing**
  - Search, query by humming, recommendation, production, editing, mixing, mastering, …
- **Advanced Topics**
  - Interactive tools, singing voice synthesis, sound design, audiovisual tools, …

Course Logistics

Communications

- **Course website**: Syllabus, schedule, materials, recordings, etc.
- **Email**: Announcements
- **Google Chat**: Q&A

hermandong.com/teaching/pat463

Prerequisites

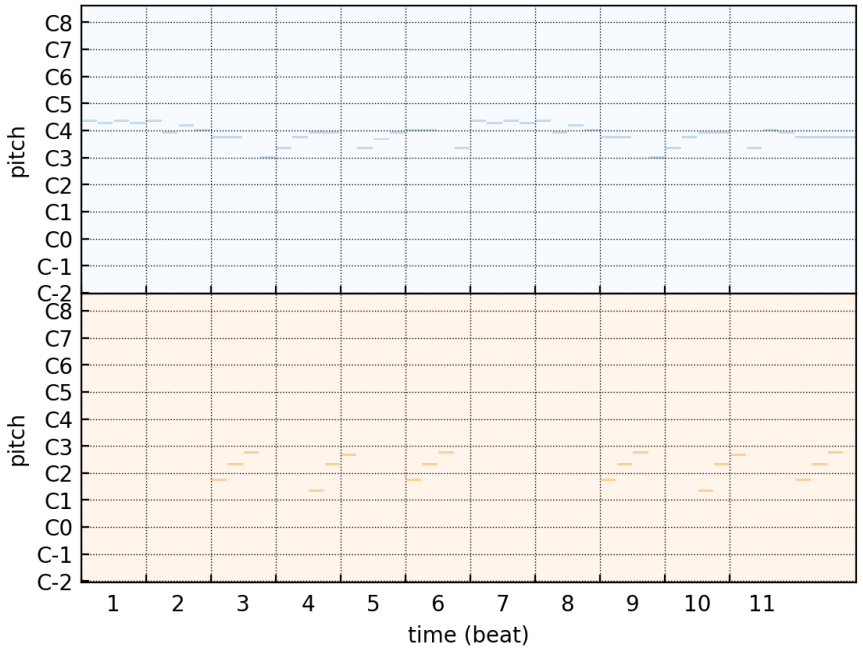
- **Prior coding experience** is recommended!
  - If you've taken any programming course, you should be fine
  - You should be comfortable reading code written by others

(Tentative) Assignments

- **Homework**
  - Real or fake!?
  - Stem separation
  - AI Song Contest 2026
  - AI music tools
- **Programming assignments**
  - Symbolic music processing
  - Spectral analysis
  - Source separation
  - Musical note classification
  - Automatic music instrumentation

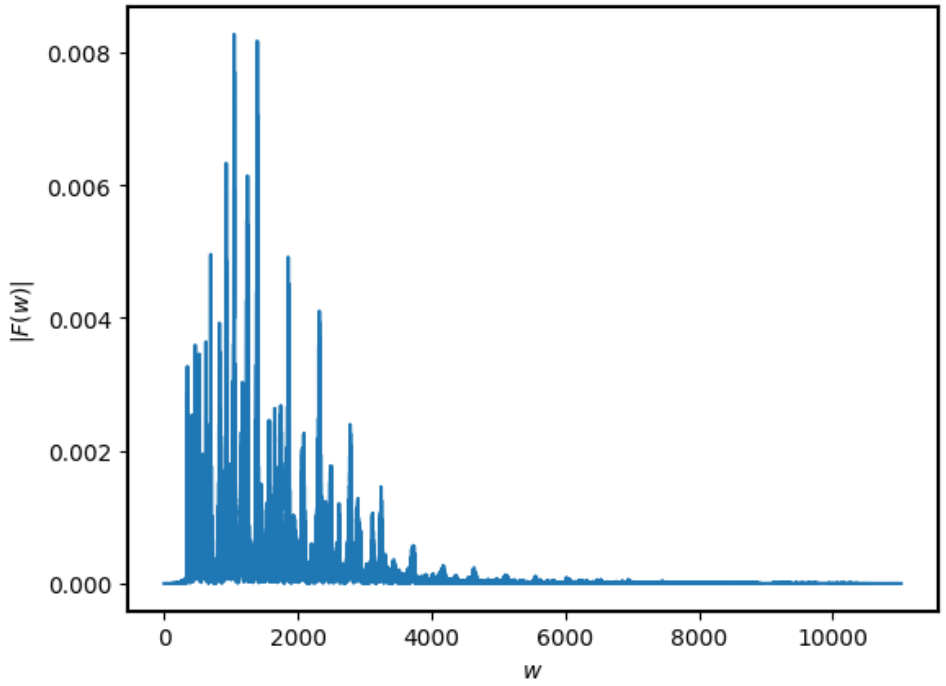
(Tentative) **PA 1**: Symbolic Music Processing

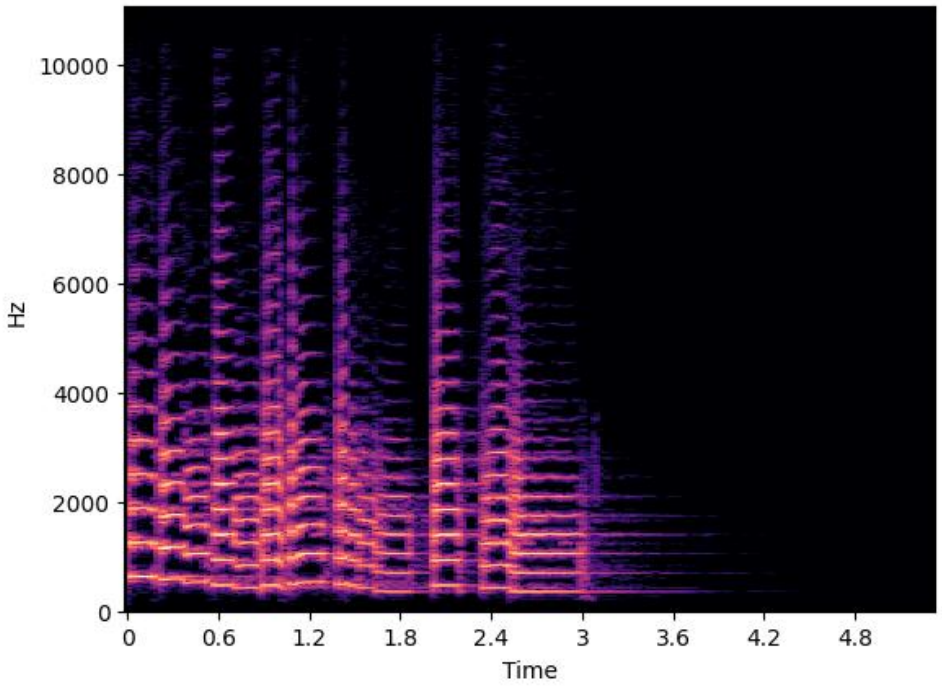
- Use **MusPy** to process symbolic music
  - **Read** MIDI and MusicXML files
  - **Manipulate** notes, instruments, and tempos
  - **Show** as piano roll(s)
  - **Write** as MIDI and MusicXML files
  - **Synthesize** into WAV files
- Learn to code in **Jupyter Notebook**
- Learn about **Google Colab**



(Tentative) **PA 2**: Spectral Analysis

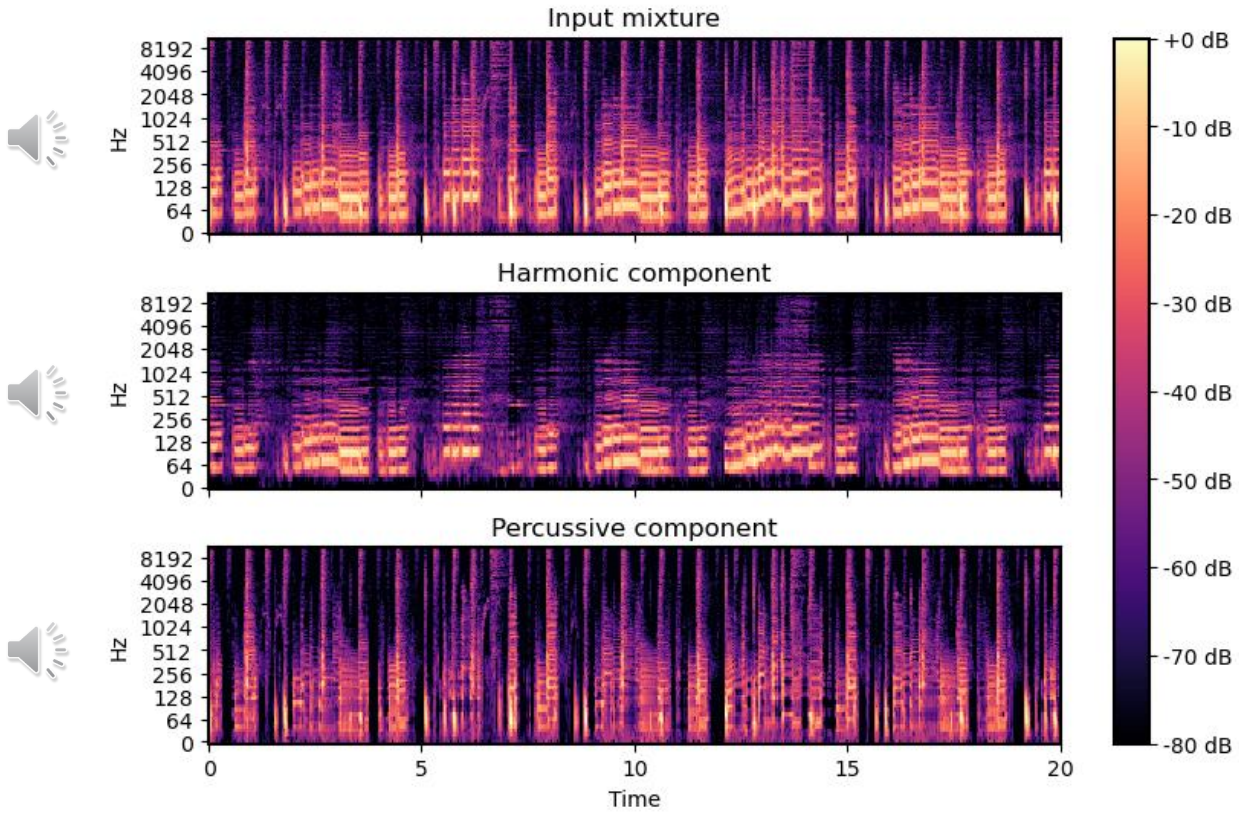
- Use **librosa** to process audio files
  - Fast Fourier transform (**FFT**)
  - Short-time Fourier transform (**STFT**)

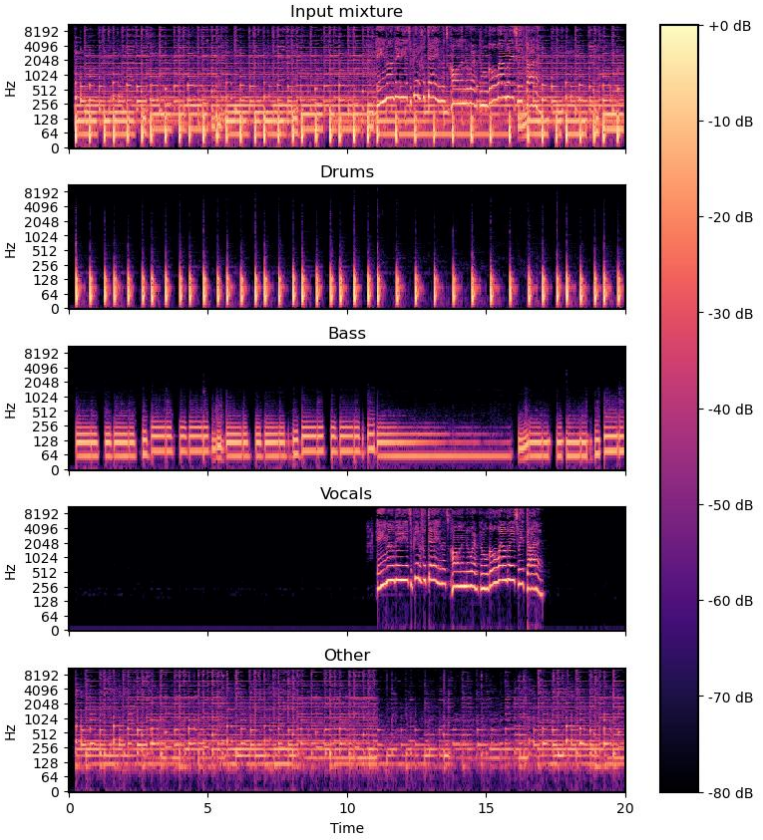




(Tentative) **PA 3**: Source Separation

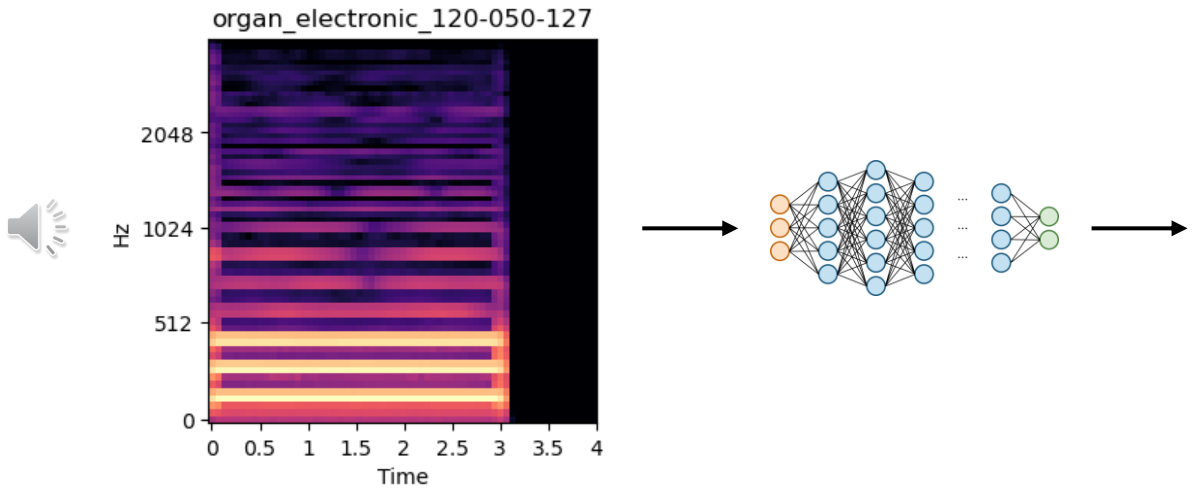
- **Part 1**: Harmonic-Percussive Source Separation (HPSS) using **librosa**
- **Part 2**: Music Source Separation using **Demucs**





(Tentative) **PA 4**: Musical Note Classification using CNNs

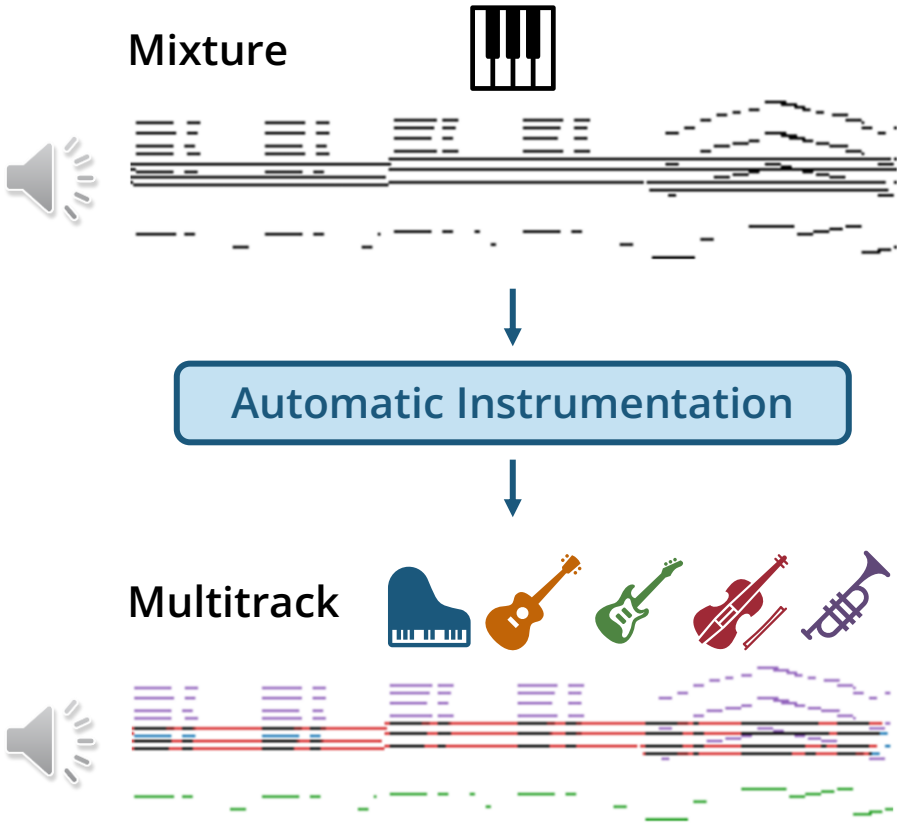
- Train a CNN that can classify audio files into their **instrument families**
  - **Input**: 64x64 mel spectrogram
  - **Output**: 11 instrument classes

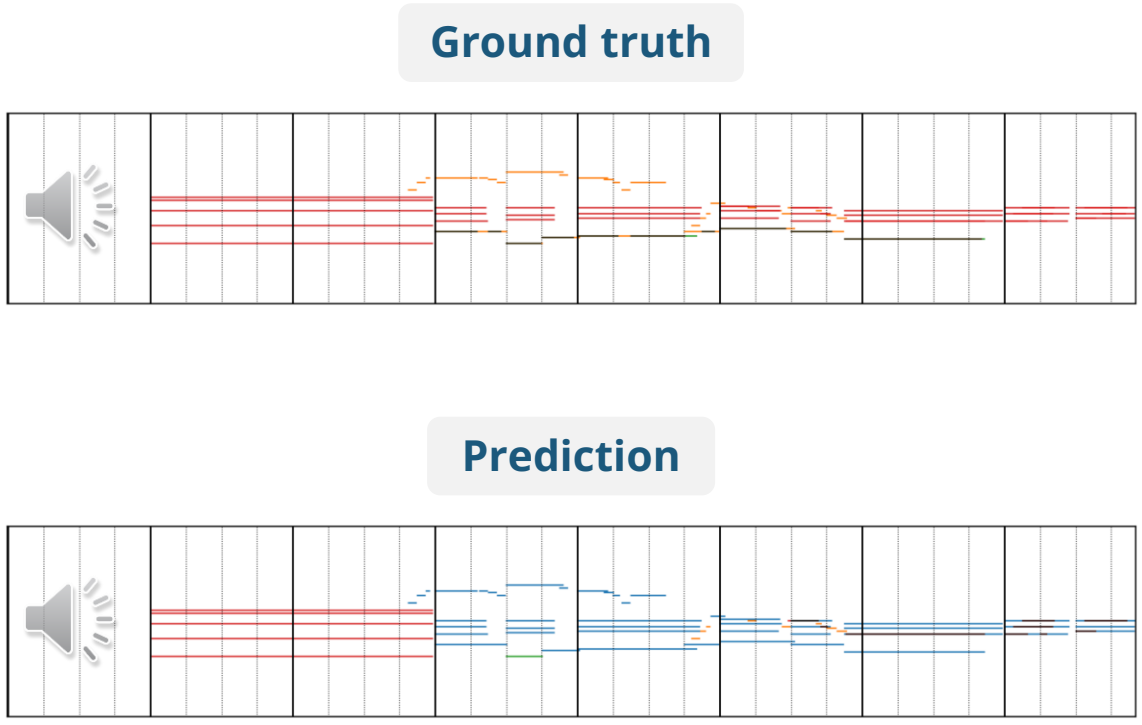


Bass
Brass
Flute
Guitar
Keyboard
Mallet
Organ
Reed
String
Synth
Lead
Vocal

(Tentative) **PA 5**: Automatic Music Instrumentation

- Train a transformer model for automatic instrumentation by **learning to separate parts**





Assignment Policies

- All assignments must be **completed on your own**
  - You are welcome to exchange ideas with your peers, but this should be in the form of concepts and discussion, not in the form of writing and code.
- **Please turn in your work even if it's incomplete for partial credits!**
- Due at **11:59pm ET** on the date specified
- Late submissions: **up to a week**, **1 point deducted per day**

Project

- **Open-ended, individual project**
- Deliverables due at **11:59pm ET** on the date specified
- **No late submissions!** Submit your work early and update it later
- Milestones
  - **Presentation** on **Dec 2** (tentative)
  - **Final report** due on **Dec 11** (tentative)

Project Topics

- Building a new AI music tool
- Exploring creative & artistic use of AI tools
- Analyzing systematically existing AI music tools

Grading

- All **grading** and **regrade requests** will be handled on Gradescope
- **Homework** (**20%**)
- **Programming assignments** (**50%**)
- **Project** (**30%**)

A+	97+	B+	87–89	C+	77–79	D+	67–69	F	<60
A	93–96	B	83–86	C	73–76	D	63–66		
A-	90–92	B-	80–82	C-	70–72	D-	60–62		

Optional Reading

- *Deep and Shallow: Machine Learning in Music and Audio*" by Shlomo Dubnov and Ross Greer
- "*Fundamentals of Music Processing*" by Meinard Müller (notebooks)
- "*Intelligent Music Production*" by Brecht De Man, Ryan Stables, and Joshua D. Reiss
- Course material for "*Deep Learning for Music Analysis and Generation*" by Yi-Hsuan Yang
- Course material for "*Musical Applications of Machine Learning*" by Juhan Nam
- Course material for "*Machine Learning for Music*" by Julian McAuley

Policies: Attendance

- **On-time, in-person attendance is expected**
- Zoom link can be found on the course website
- Course lectures will be recorded and made available on the course website
- Please **attend in-person** for
  - **Project pitch**
  - **Project presentation**

Policies: Generative AI Usage

- Feel free to use GenAI tools in your workflow. However, **you must disclose your usage of GenAI services in your write-ups**.
  - For example, U-M GPT, ChatGPT, Gemini, Claude, GitHub Copilot, DALLE, Midjourney, Stable Diffusion, Firefly, etc.
- **You take full responsibility for AI-generated materials as if you had produced them yourself**: ideas should be attributed and facts should be true.

Policies: Academic Integrity

- Plagiarism and cheating violate SMTD's Academic Code of Conduct. **All plagiarism, cheating and other academic misconduct cases will be reported to SMTD's Office of Academic and Student Affairs.**
- **All assignments must be completed on your own**. You are welcome to exchange ideas with your peers, but this should be in the form of concepts and discussion, not in the form of writing and code.
- You must **provide proper citations/references for any external resources** you use in your writing and code.

Performing Arts Technology (PAT)

Minor in Performing Arts Technology (PAT)

Intro (Required)

PAT 100: Music in Technology
PAT 200: Intro to Electronic Music

PAT Theory/Studies (6 credits)

PAT 150: Experiential Music Theory
PAT 205: Intermedia AI Music Practice
PAT 305: Video Game Music
PAT 315: Diversity in Music Technology
PAT 316: NOISE

PAT Practice (3 credits)

PAT 202: Computer Music
PAT 204: Creative Coding
PAT 220: Songwriting Workshop
PAT 280: Sound Reinforcement
PAT 412: Digital Music Ensemble
PAT 413: Electronic Chamber Music

Electives (6 credits)

PAT 421: Advanced Psychoacoustics
PAT 422: Technical Ear Training & Critical Listening
PAT 424: Dialog of the Senses
PAT 431/432: Contemporary Practice in Studio Production I/II
PAT 441: Sound for Film and Games
PAT 443: Immersive Media
PAT 451/452: Interactive Media Design I/II
PAT 454: Digital Fabrication for Acoustics
PAT 461: Performance Systems
PAT 462: Digital Sound Synthesis
PAT 463: Music & AI
PAT 464: Generative AI for Music & Audio Creation
PAT 472: Business of Music

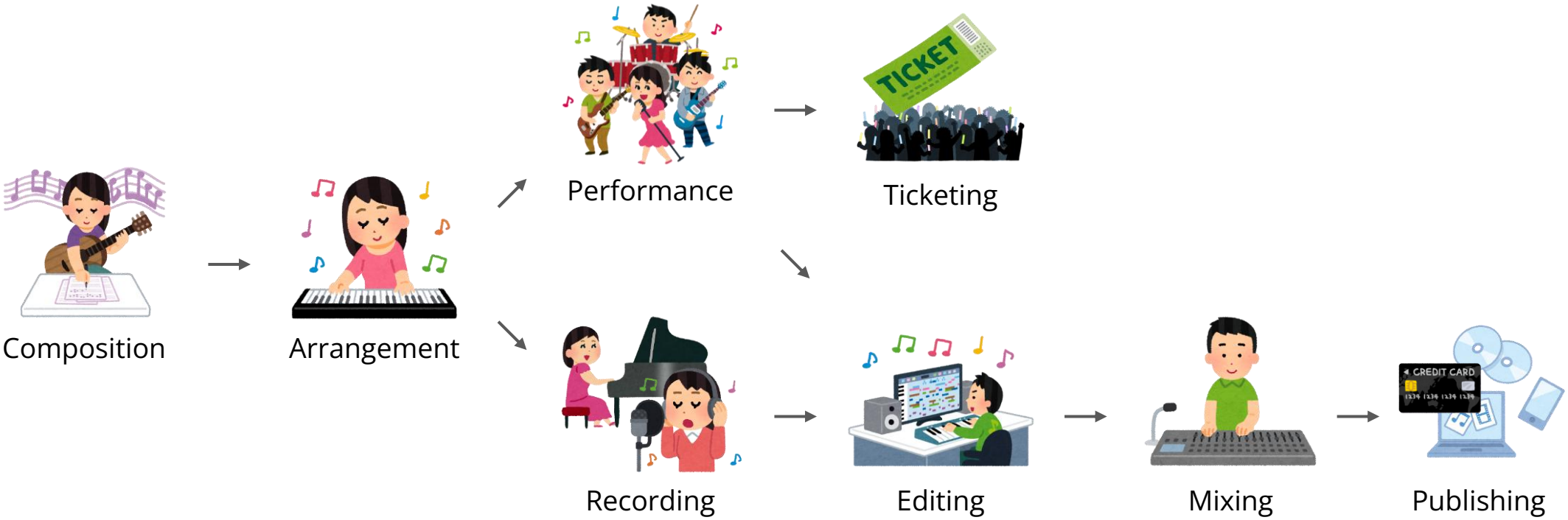
PAT PERFORMING ARTS TECHNOLOGY
UNIVERSITY OF MICHIGAN

🤔 Questions on the Syllabus?

hermandong.com/teaching/pat463

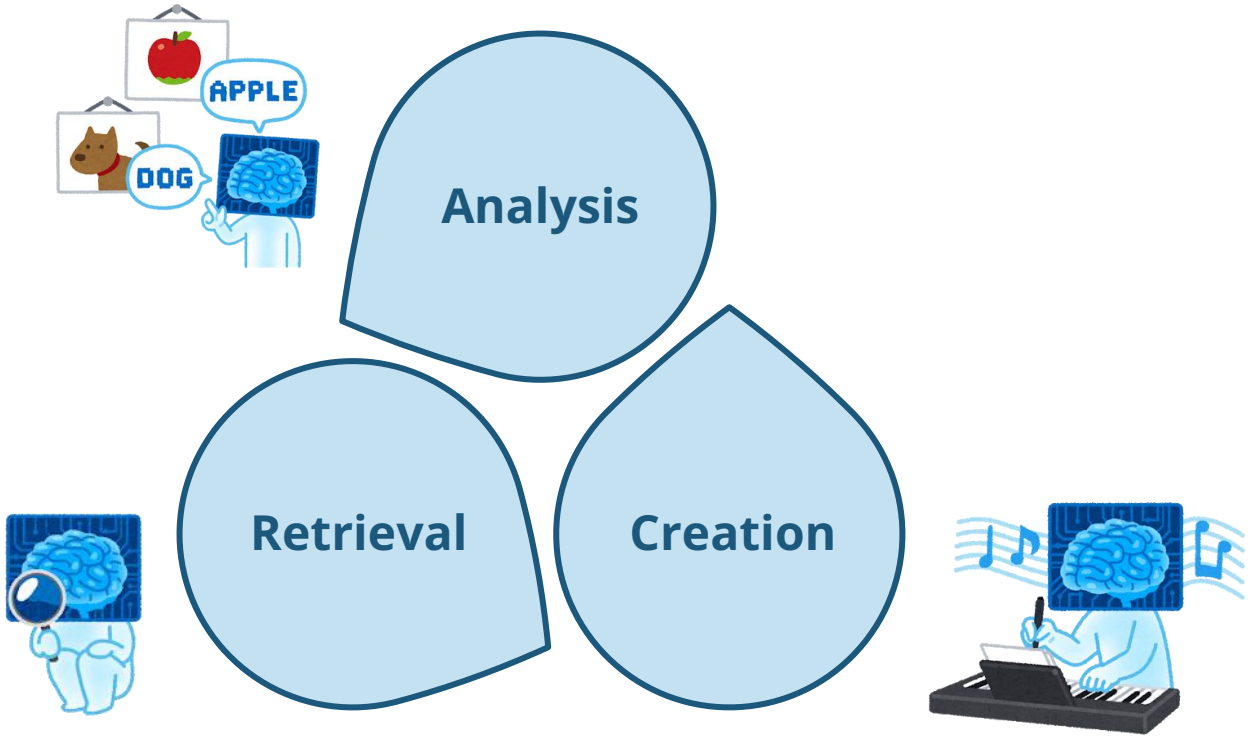
Landscape of AI for Music

An Overly-Simplified Music Production Workflow



Music Information Research (MIR)

- "Intelligent ways to analyze, retrieve and create music" (Yang 2018)



Yang, "Music Information Research," SNHCC, TIGP, lecture notes, April 2018.

Landscape of AI for Music

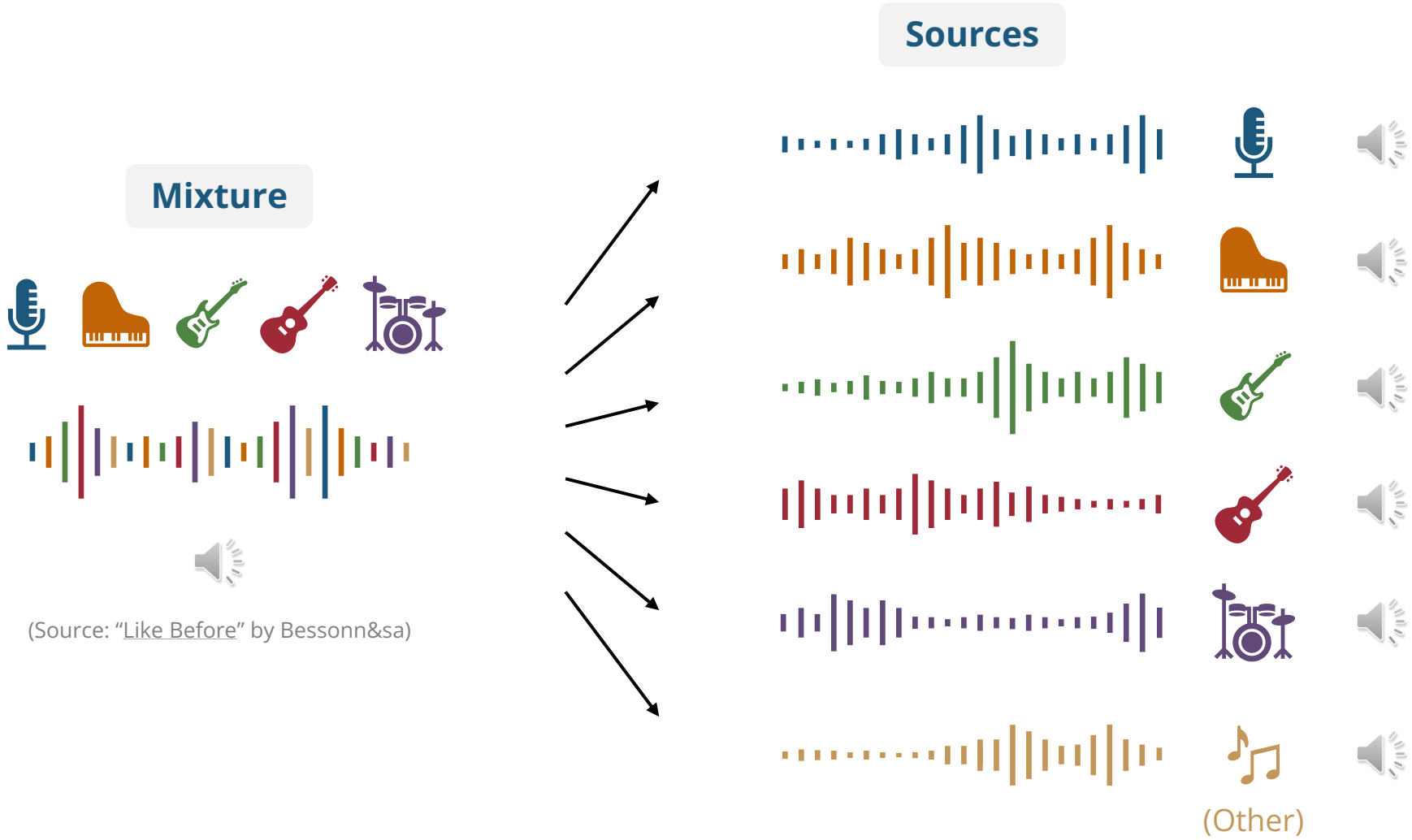
- **Analysis**
  - Key/chord recognition, genre/mood classification, source separation, beat tracking, automatic music transcription, optical music recognition, structural analysis, computational musicology, …
- **Creation**
  - Automatic music composition, music accompaniment, music arrangement, music style transfer, audio synthesis, singing voice synthesis, lyric generation, robotics improvisation, human-AI music co-creation, …
- **Retrieval & Processing**
  - Search, query by humming, recommendation, production, editing, mixing, mastering, …
- **Advanced Topics**
  - Interactive tools, singing voice synthesis, sound design, audiovisual tools, …

Analysis

Retrieval

Creation

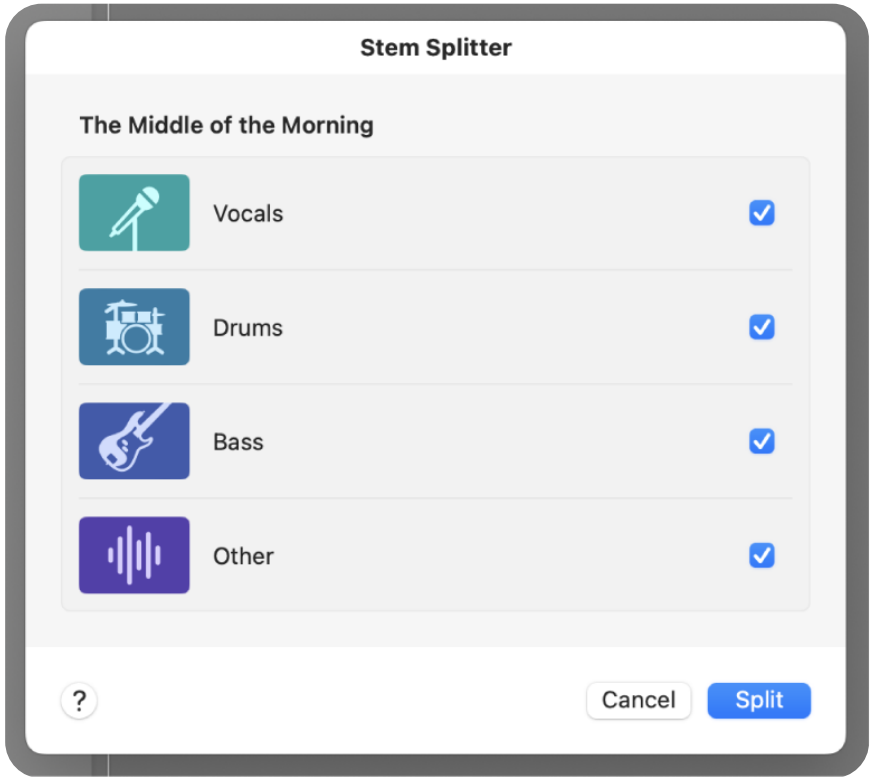
Source Separation



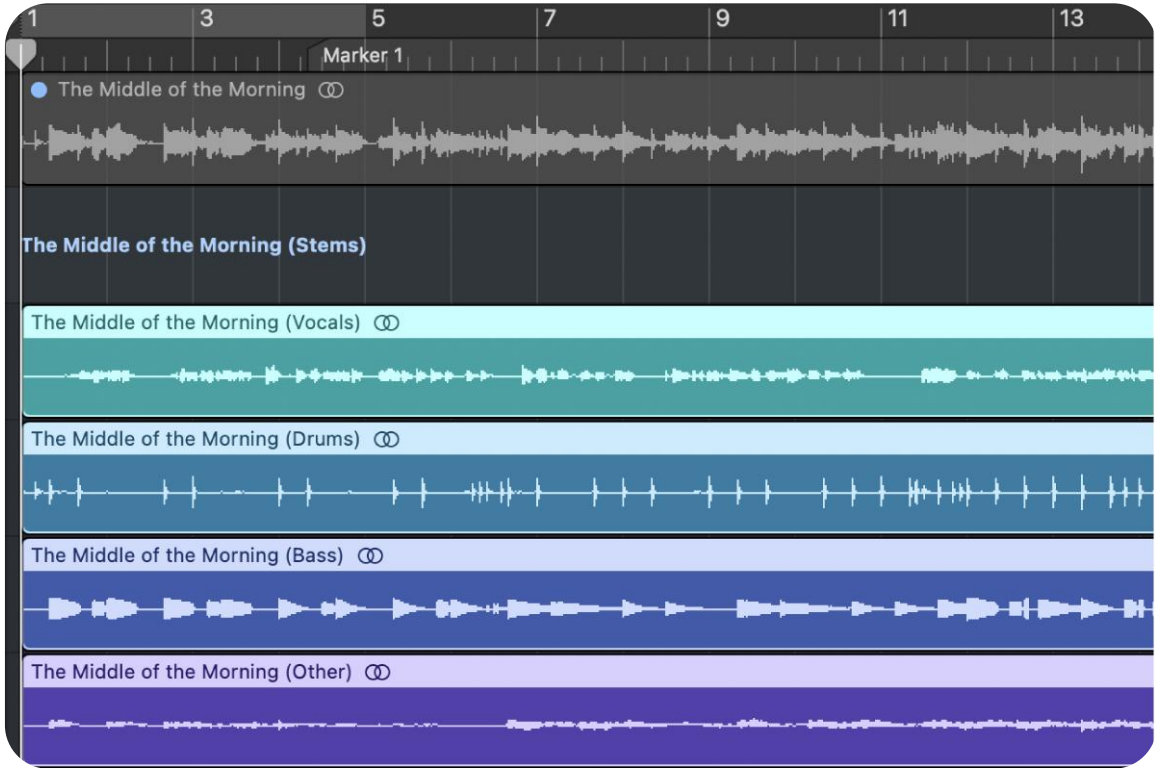
(Source: "Like Before" by Bessonn&sa)

Bessonn&sa, CC BY-NC-SA, via Jamendo
Illustration inspired by Ethan Manilow, Prem Seetharman, and Justin Salamon's Tutorial on "Open Source Tools & Data for Music Source Separation" at ISMIR 2020.

Stem Splitter in Logic Pro



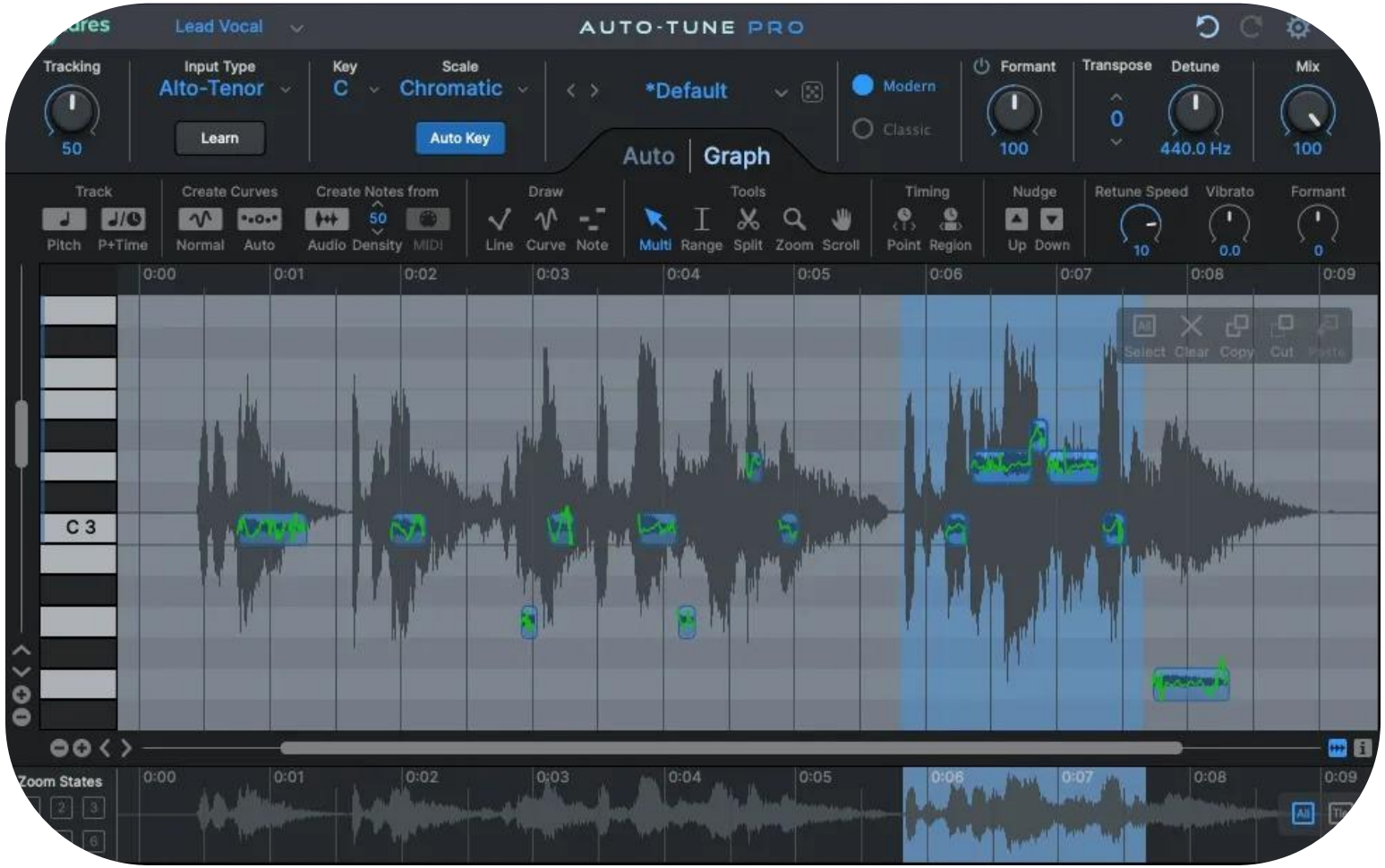
(Source: Logic Pro User Guide)



(Source: Logic Pro User Guide)

support.apple.com/guide/logicpro/extract-vocal-instrumental-stems-stem-lgcp61bae908/mac

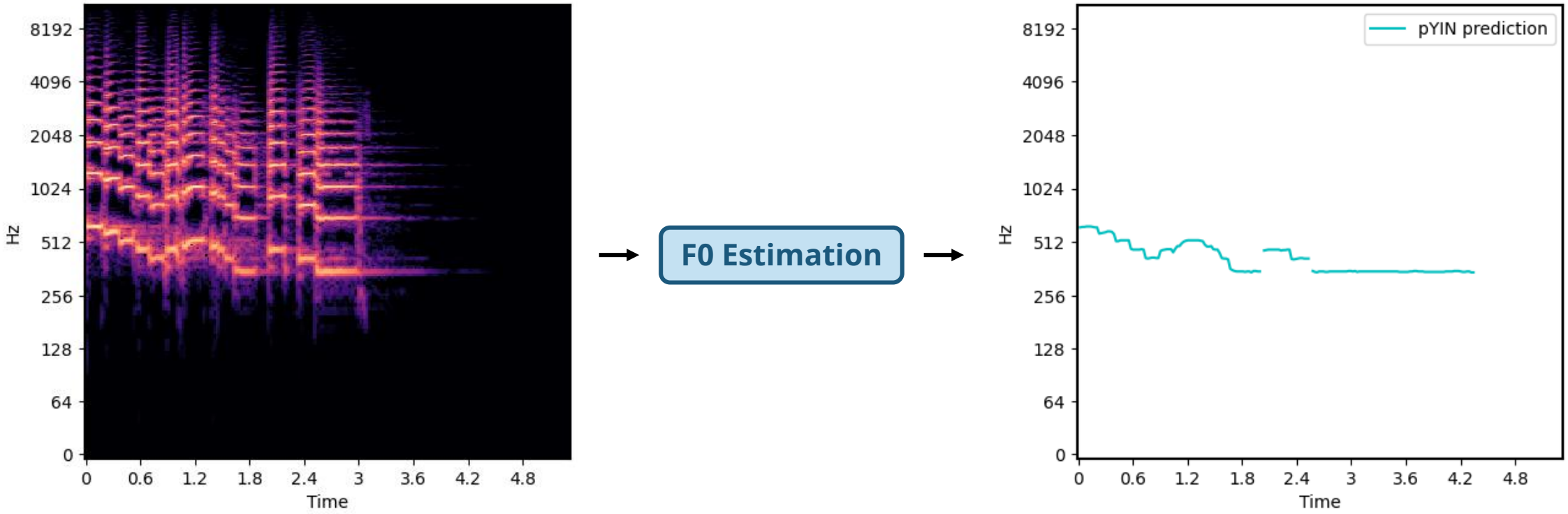
Auto-tune Pro



(Source: Antares Audio Technologies)

antarestech.com/documentation/auto-tune-pro/chapter-2

Fundamental Frequency (F0) Estimation



Chord Detection

youtube.com/shorts/_N3b_GARMfA

Structure Analysis

Music segmentation

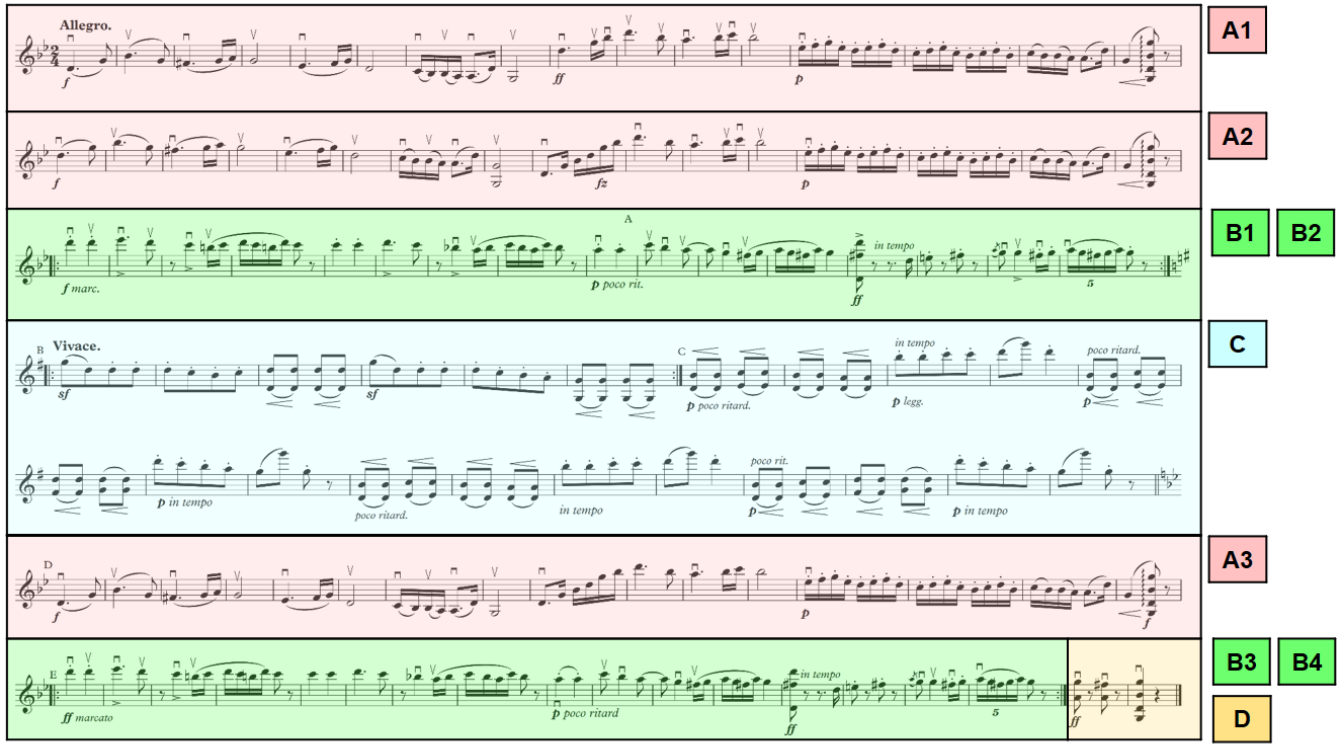


Figure 4.5 following [Müller, FMP, Springer 2015]

(Source: Müller & Zalkow, 2019)

Hierarchical music segmentation

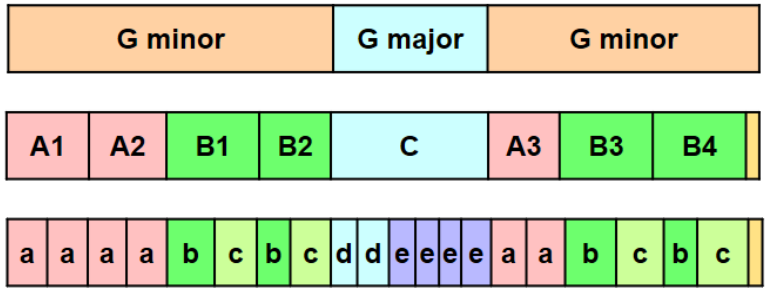


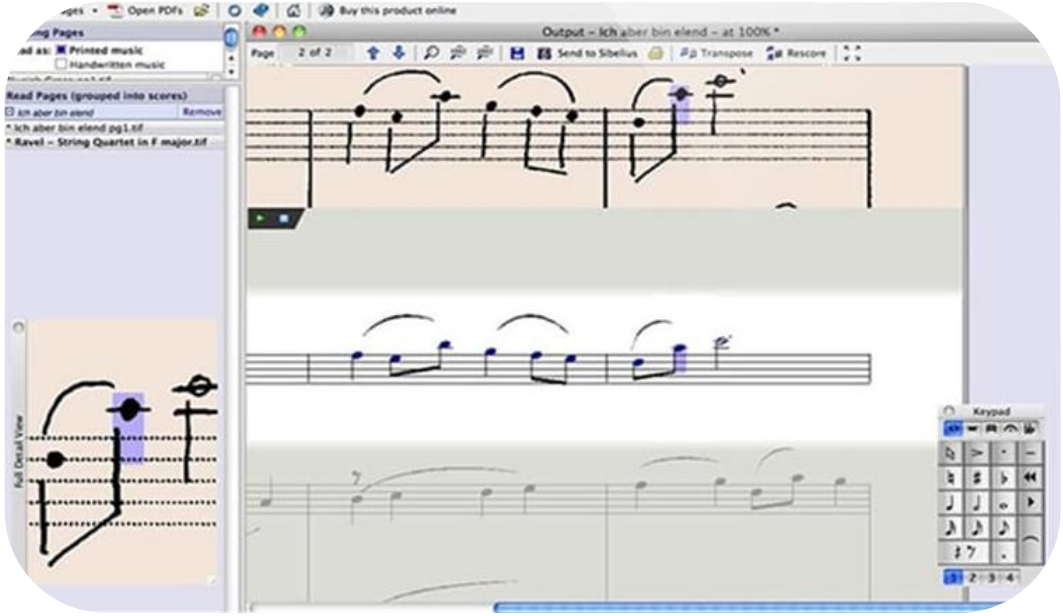
Figure 4.28 from [Müller, FMP, Springer 2015]

(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.

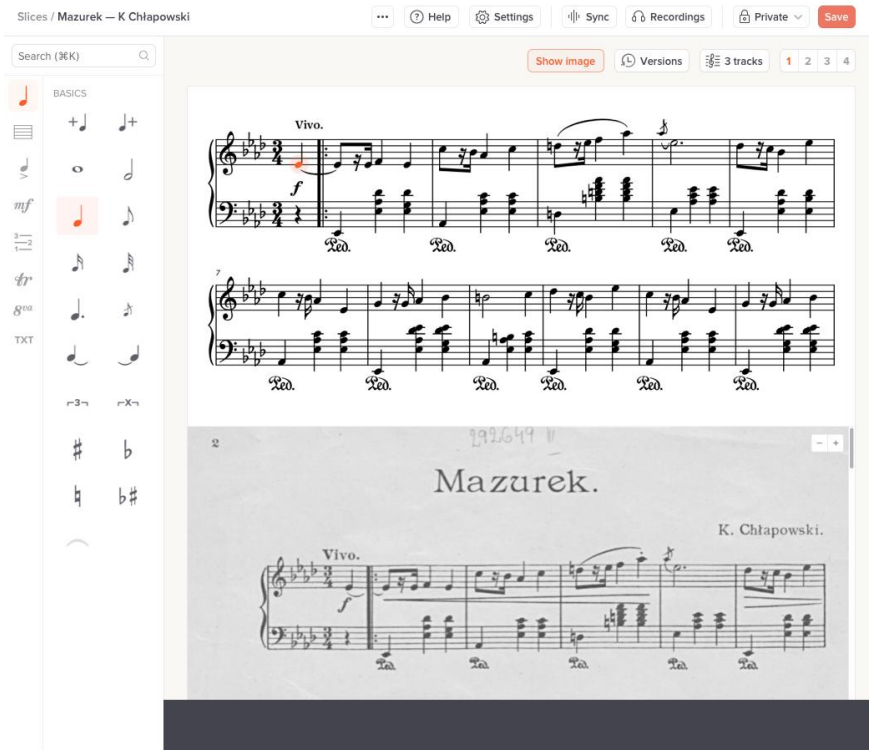
Optical Music Recognition (OMR)

PhotoScore & NotateMe in Sibelius



(Source: Avid)

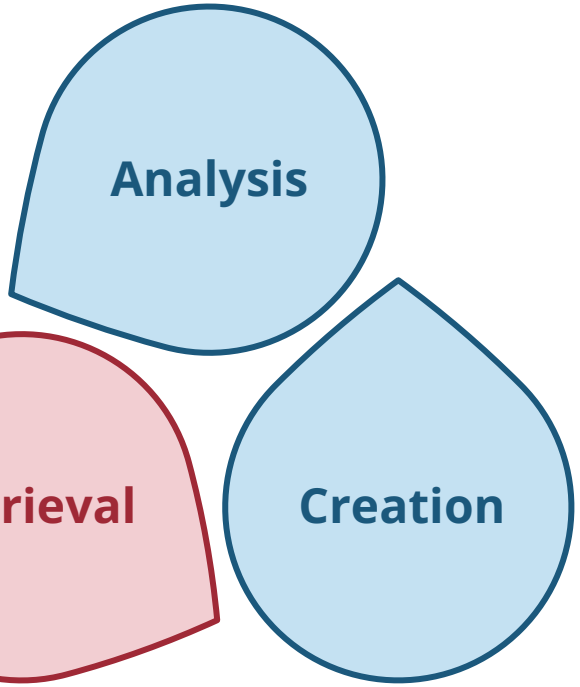
Soundslice



(Source: Soundslice)

avid.com/products/photoscore-and-notateme-ultimate
soundslice.com/sheet-music-scanner/


APPLE
DOG


Analysis
Retrieval
Creation

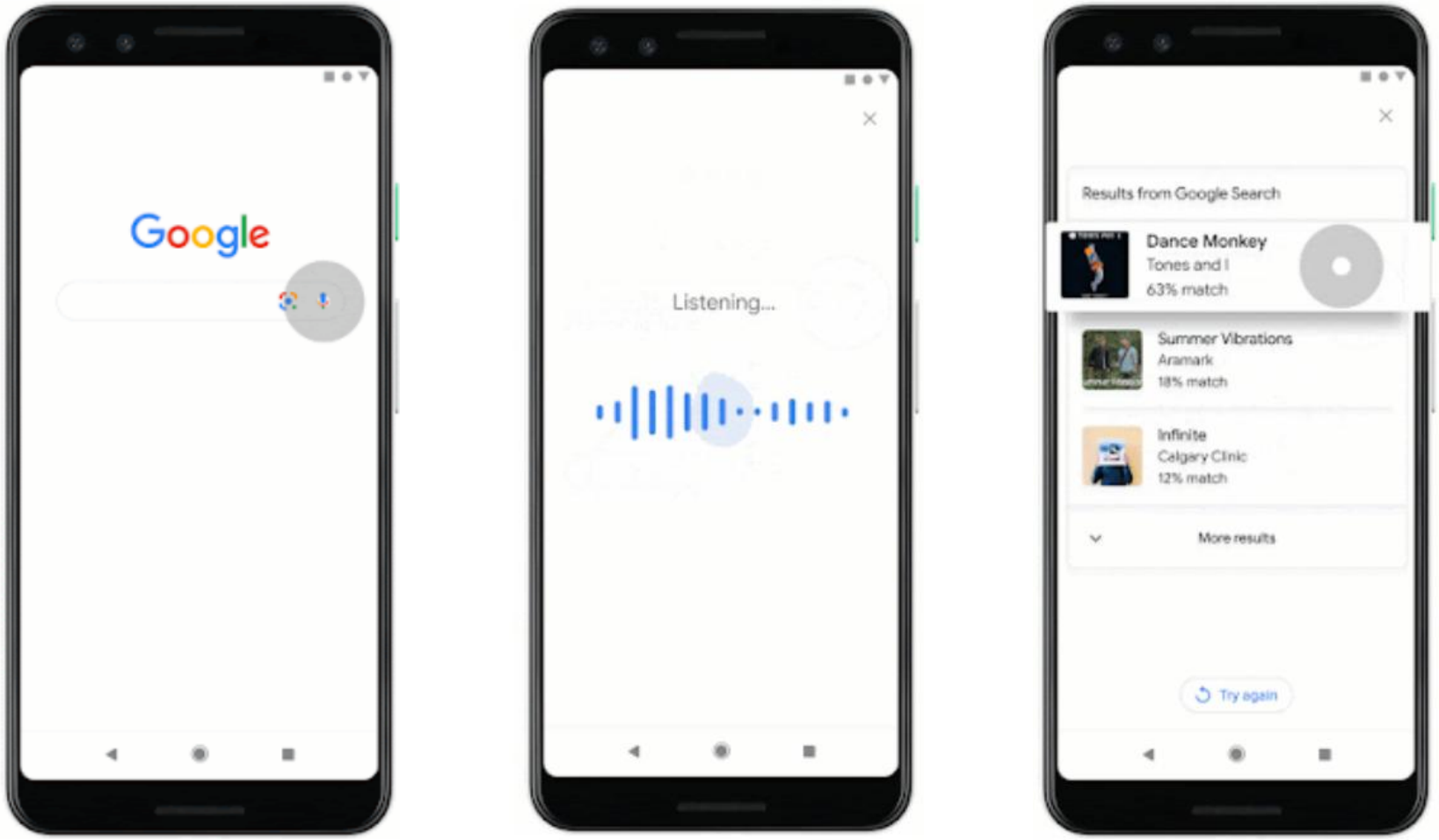
Shazam & Siri

(Source: Shazam User Guide)

(Source: OSXDaily)

osxdaily.com/2017/09/08/identify-song-playing-mac-siri/
support.apple.com/en-is/guide/shazam/dev9748744b6/web

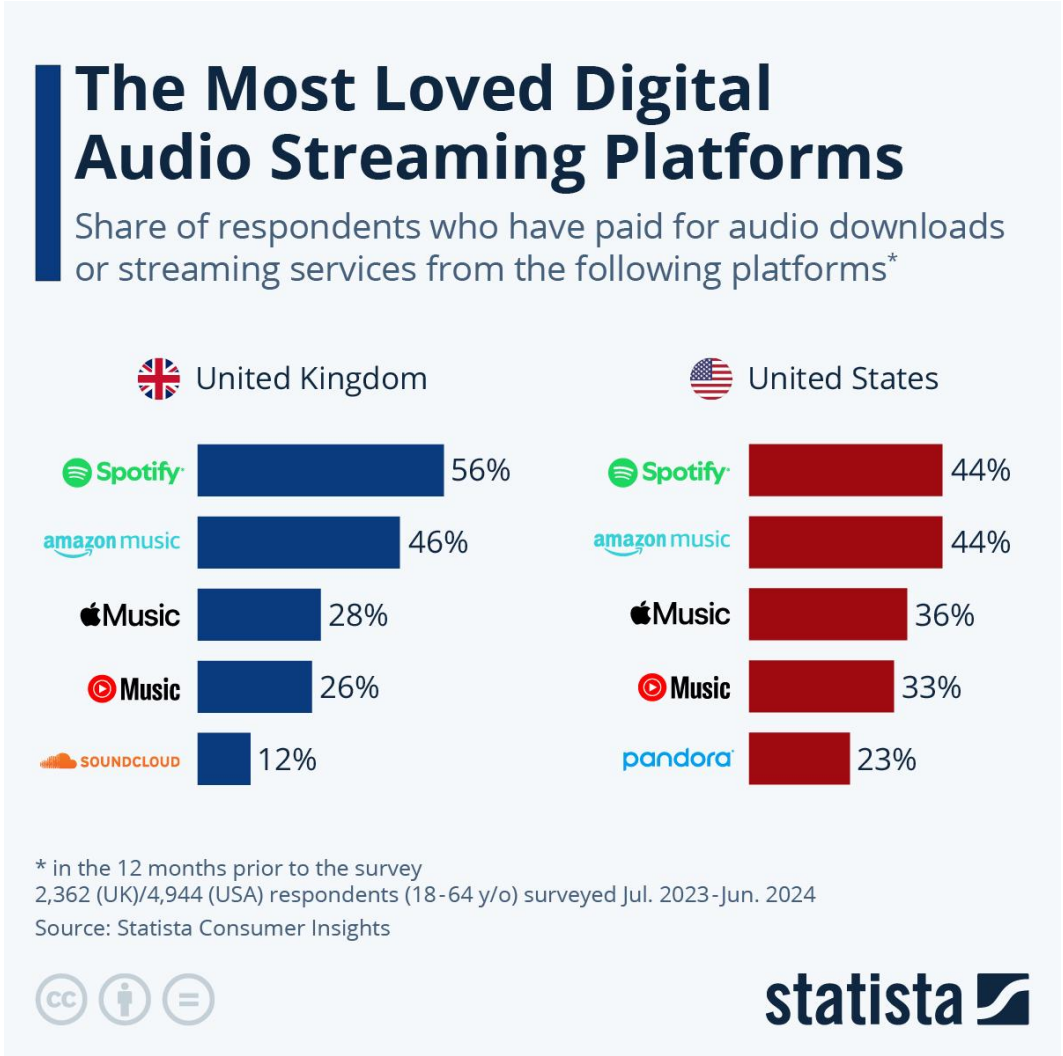
Hum to Search (Google)



(Source: Google Research Blog)

research.google/blog/the-machine-learning-behind-hum-to-search/

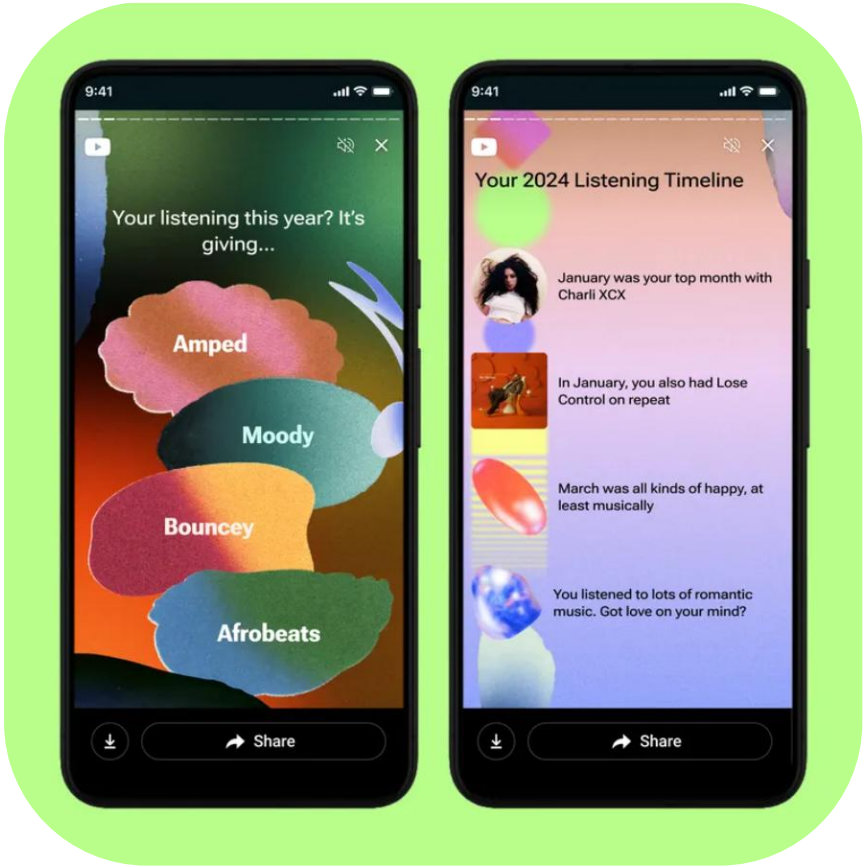
Music Recommendation



Anna Fleck, "The Most Loved Digital Audio Streaming Platforms," *Statista*, September 30, 2024.

Listening Behavior Analysis

YouTube's Music Recap



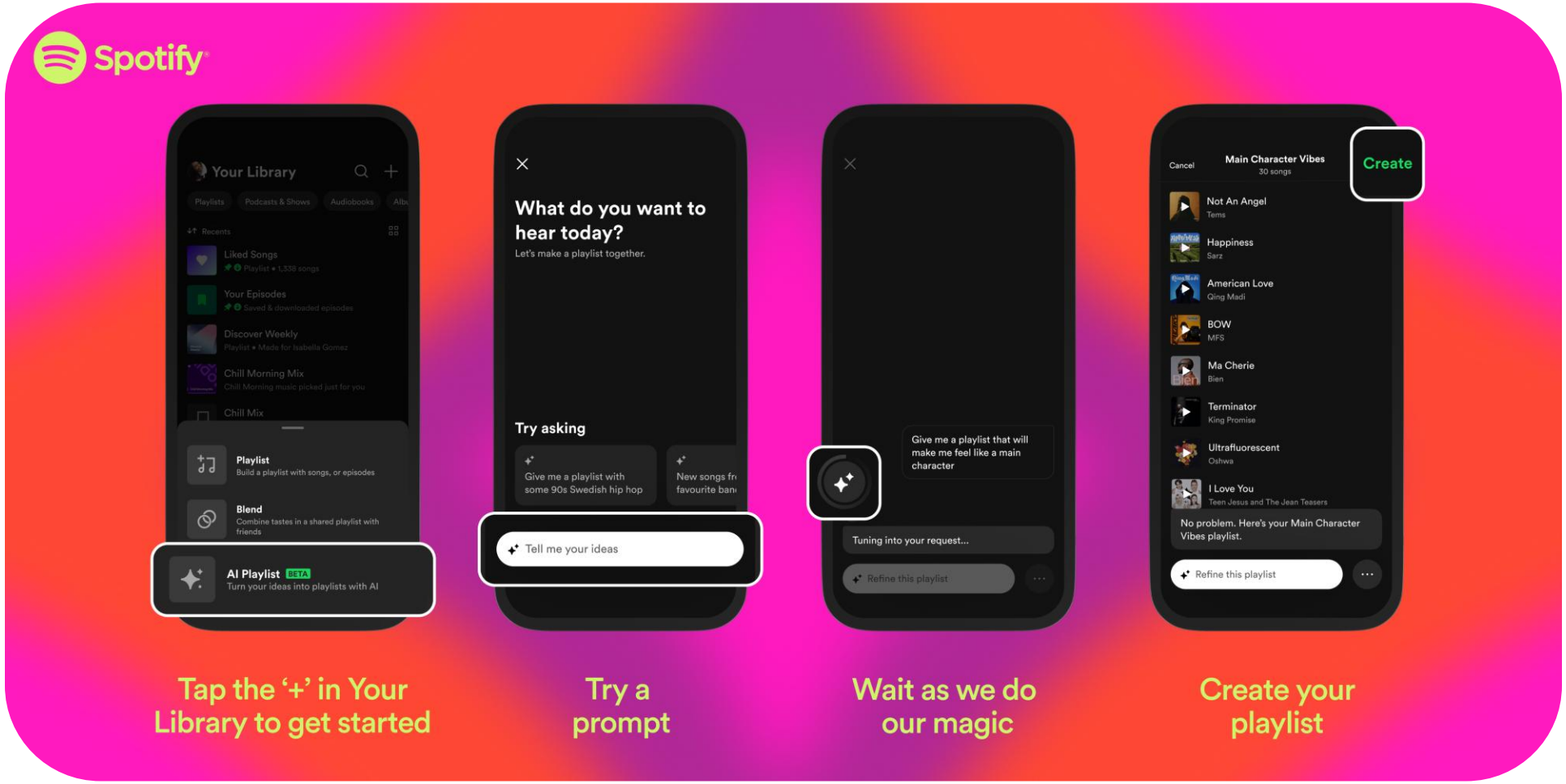
(Source: YouTube)

Spotify's Listening Personality

(Source: Spotify)

blog.youtube/news-and-events/2024-music-recap-youtube/
engineering.atspotify.com/2023/01/whats-a-listening-personality/

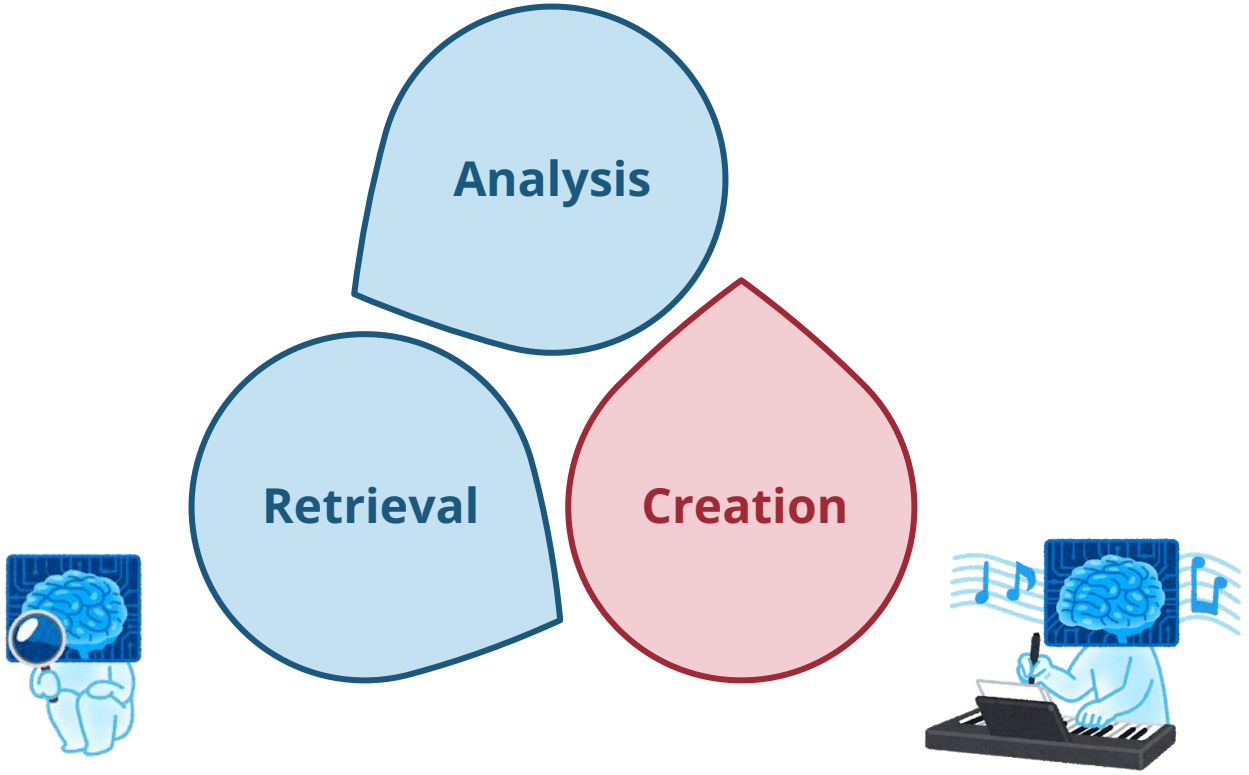
AI Playlist (Spotify)



(Source: Spotify)

newsroom.spotify.com/2024-04-07/spotify-premium-users-can-now-turn-any-idea-into-a-personalized-playlist-with-ai-playlist-in-beta/


APPLE
DOG


Analysis
Retrieval
Creation

Piano Genie (Donahue et al., 2018)

piano-genie.glitch.me/

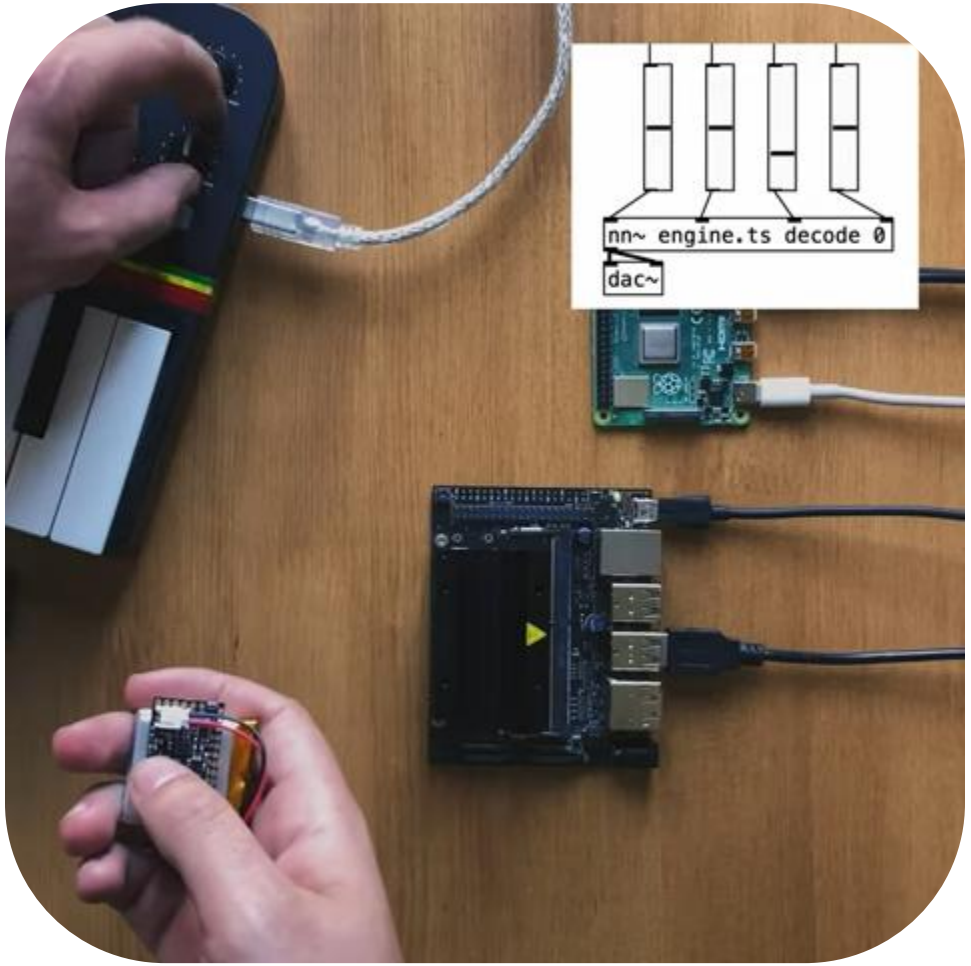
youtu.be/YRb0XAnUpIk & magenta.tensorflow.org/pianogenie

Chris Donahue, Ian Simon, and Sander Dieleman, "Piano Genie," IUI, 2019.

`unloop` (Garcia et al., 2023)

youtu.be/yzBI8Vcjd2s & github.com/hugofloresgarcia/unloop

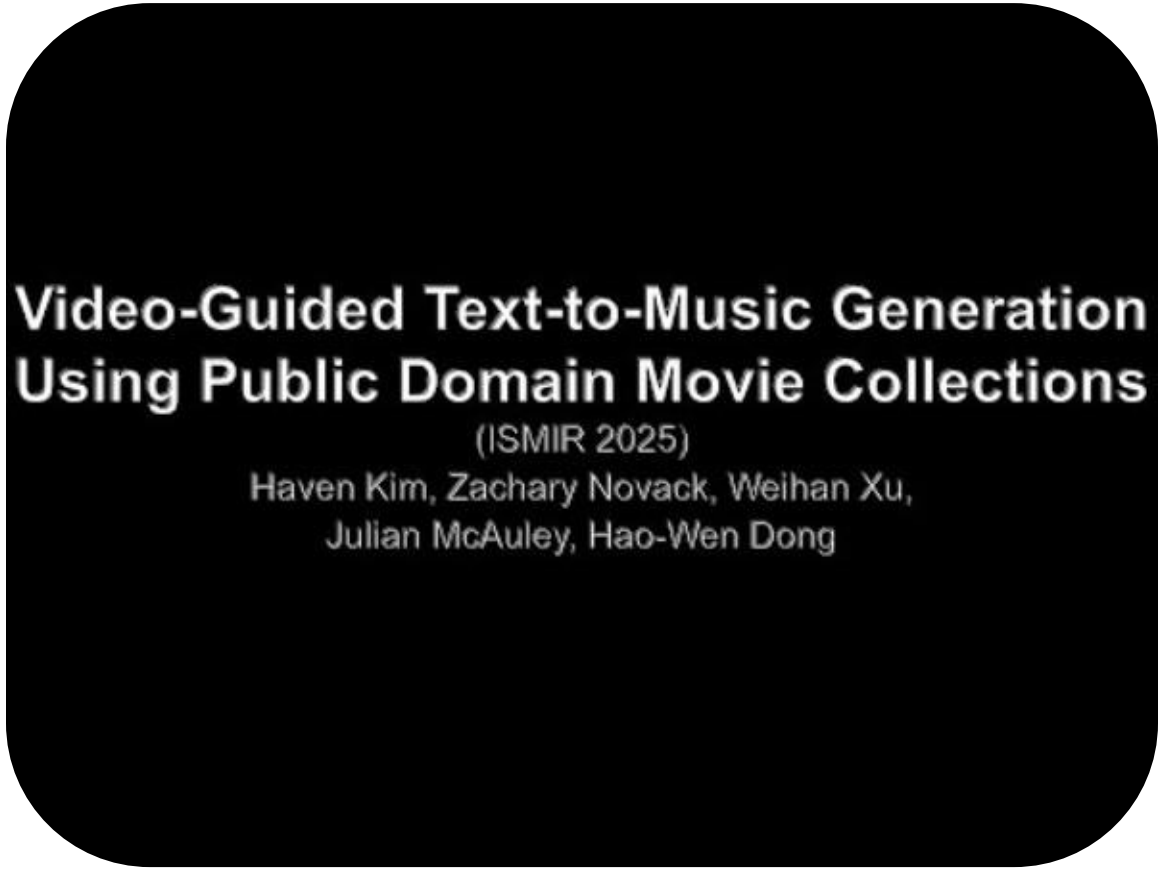
RAVE: Real-time Audio Synthesis (Caillon & Esling, 2022)



youtu.be/jAIRf4nGgYI

Antoine Caillon and Philippe Esling, "RAVE: A variational autoencoder for fast and high-quality neural audio synthesis," *arXiv preprint arXiv:2111.05011*, 2021.

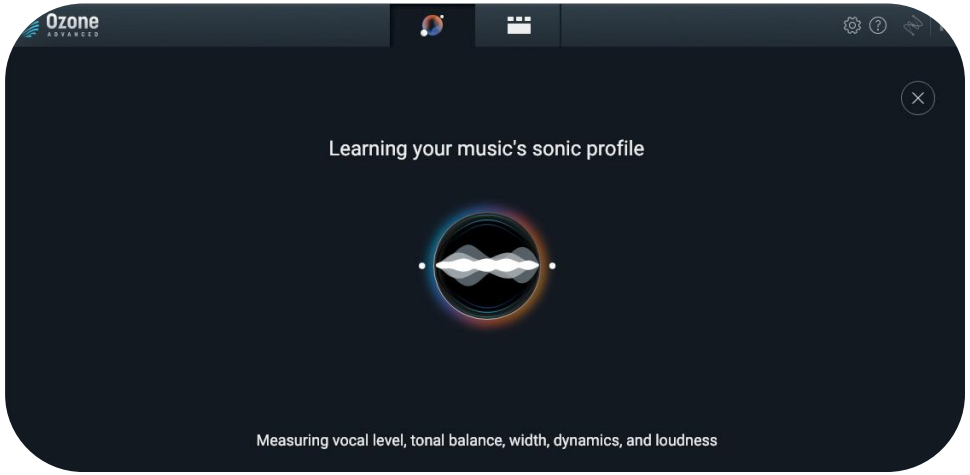
Video-Guided Text-to-Music Generation (Kim et al., 2025)



youtu.be/S0BMicbdzmg

Haven Kim, Zachary Novack, Weihan Xu, Julian McAuley, and Hao-Wen Dong , "Video-Guided Text-to-Music Generation Using Public Domain Movie Collections," *ISMIR*, 2025.

iZotope Ozone's Master Assistant



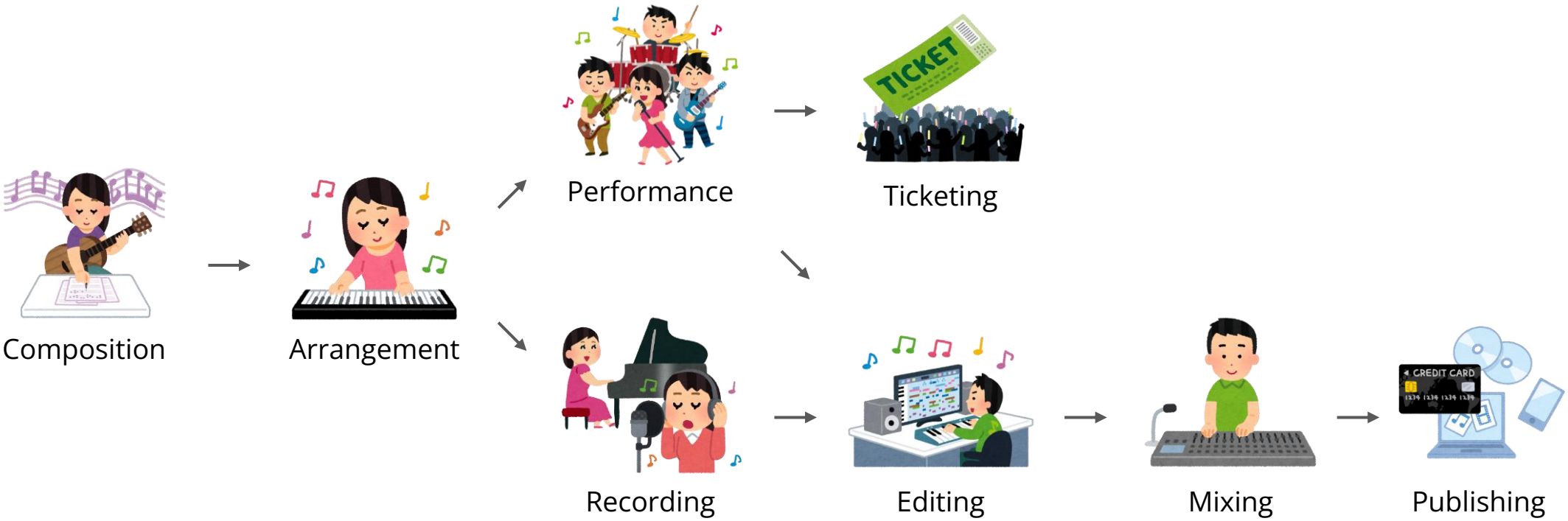
(Source: iZotope)

(Source: iZotope)

izotope.com/en/learn/10-steps-to-a-quick-master-in-ozone.html

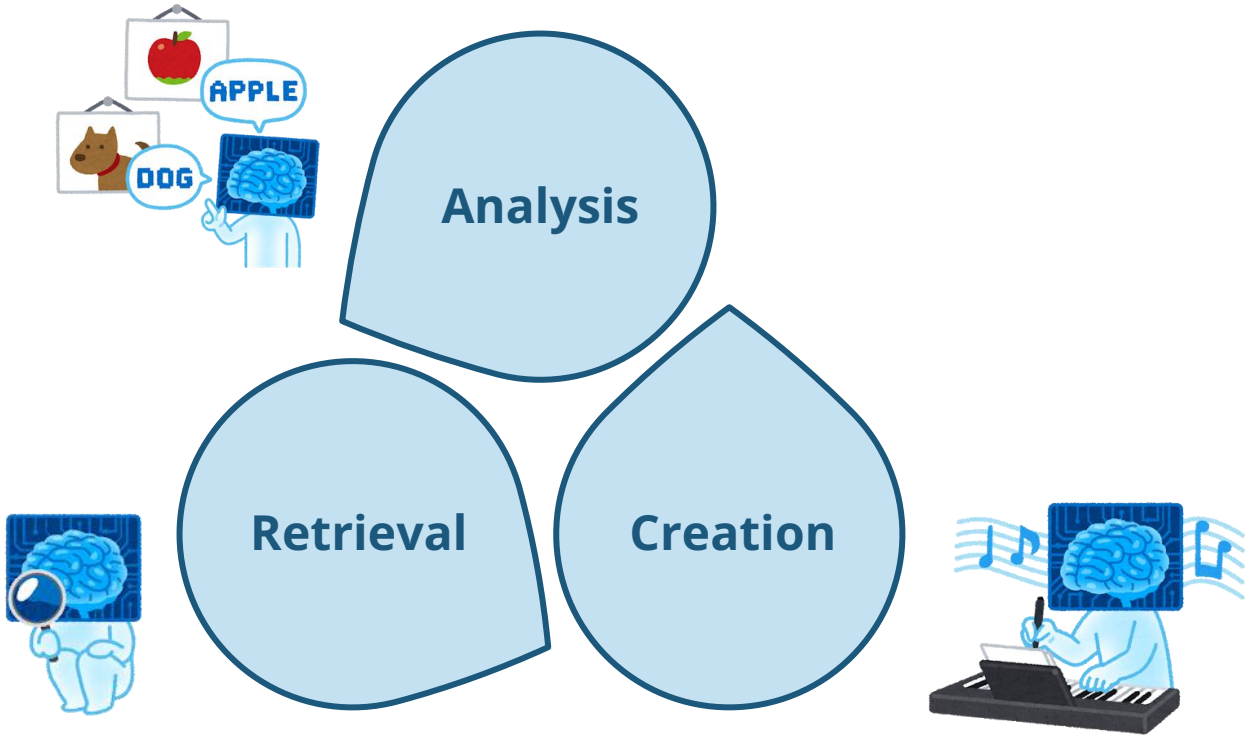
Recap

An Overly-Simplified Music Production Workflow



Music Information Research (MIR)

- "Intelligent ways to analyze, retrieve and create music" (Yang 2018)

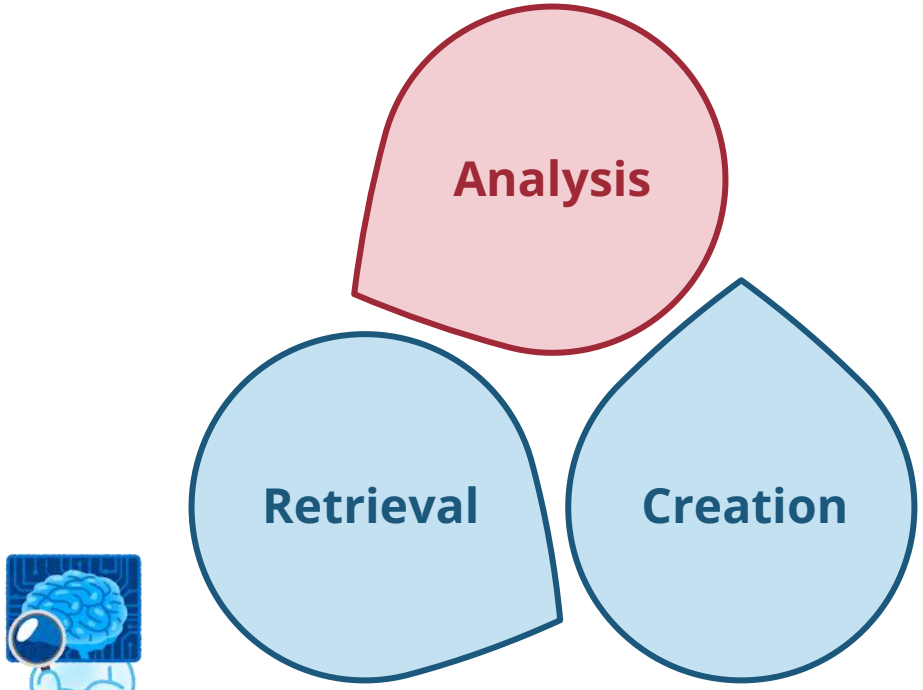


Yang, "Music Information Research," SNHCC, TIGP, lecture notes, April 2018.

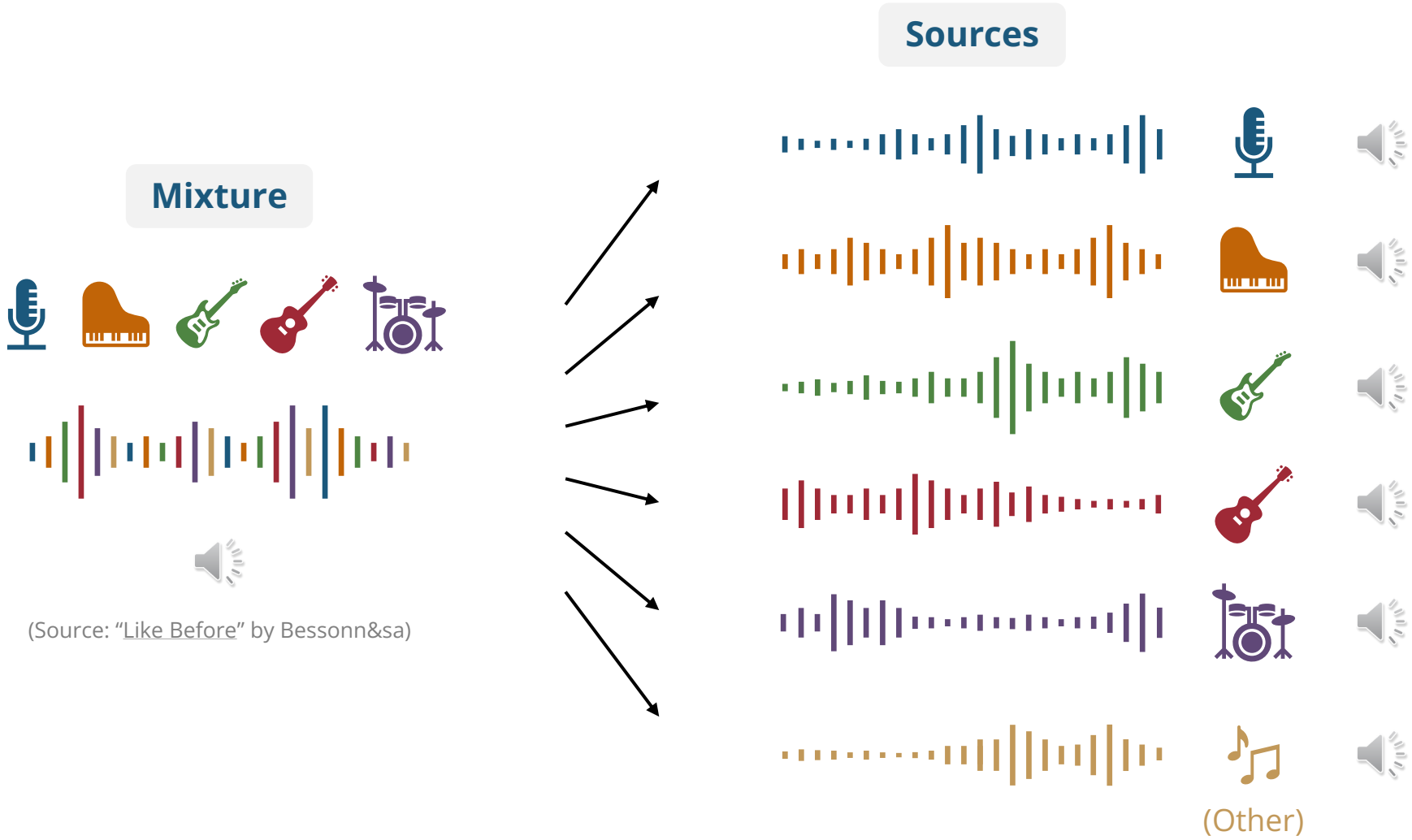
Landscape of AI for Music

- **Analysis**
  - Key/chord recognition, genre/mood classification, source separation, beat tracking, automatic music transcription, optical music recognition, structural analysis, computational musicology, …
- **Creation**
  - Automatic music composition, music accompaniment, music arrangement, music style transfer, audio synthesis, singing voice synthesis, lyric generation, robotics improvisation, human-AI music co-creation, …
- **Retrieval & Processing**
  - Search, query by humming, recommendation, production, editing, mixing, mastering, …
- **Advanced Topics**
  - Interactive tools, singing voice synthesis, sound design, audiovisual tools, …


APPLE
DOG


Analysis
Retrieval
Creation

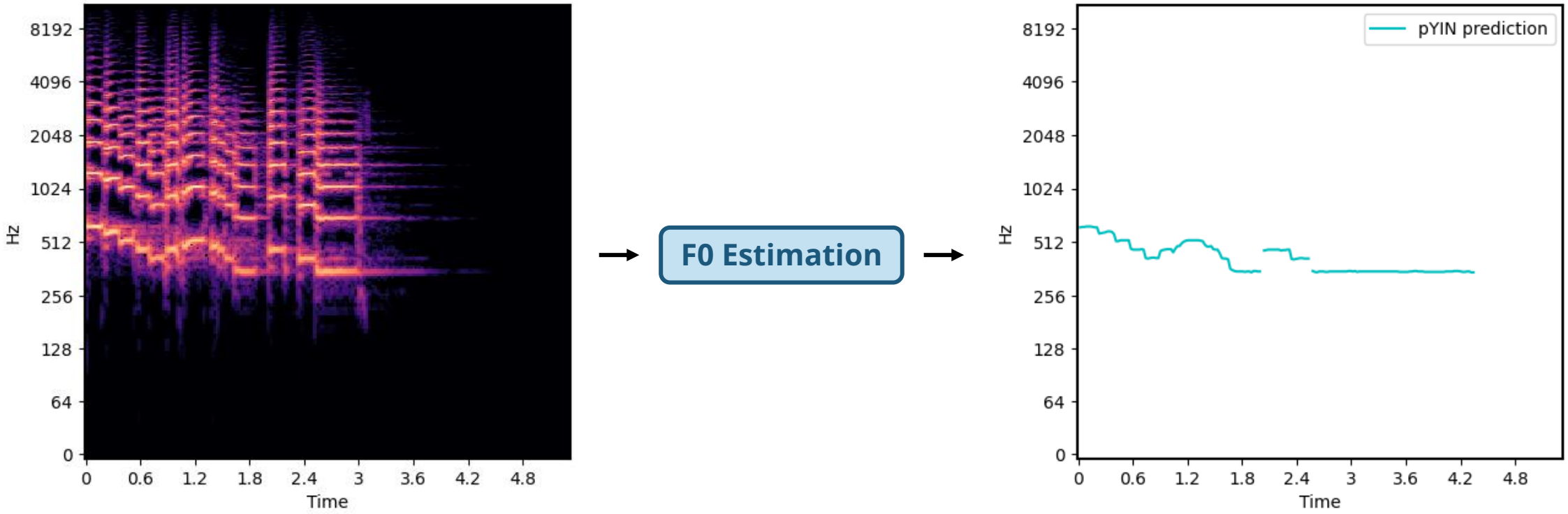
Source Separation



(Source: "Like Before" by Bessonn&sa)

Bessonn&sa, CC BY-NC-SA, via Jamendo
Illustration inspired by Ethan Manilow, Prem Seetharman, and Justin Salamon's Tutorial on "Open Source Tools & Data for Music Source Separation" at ISMIR 2020.

Fundamental Frequency (F0) Estimation



Structure Analysis

Music segmentation

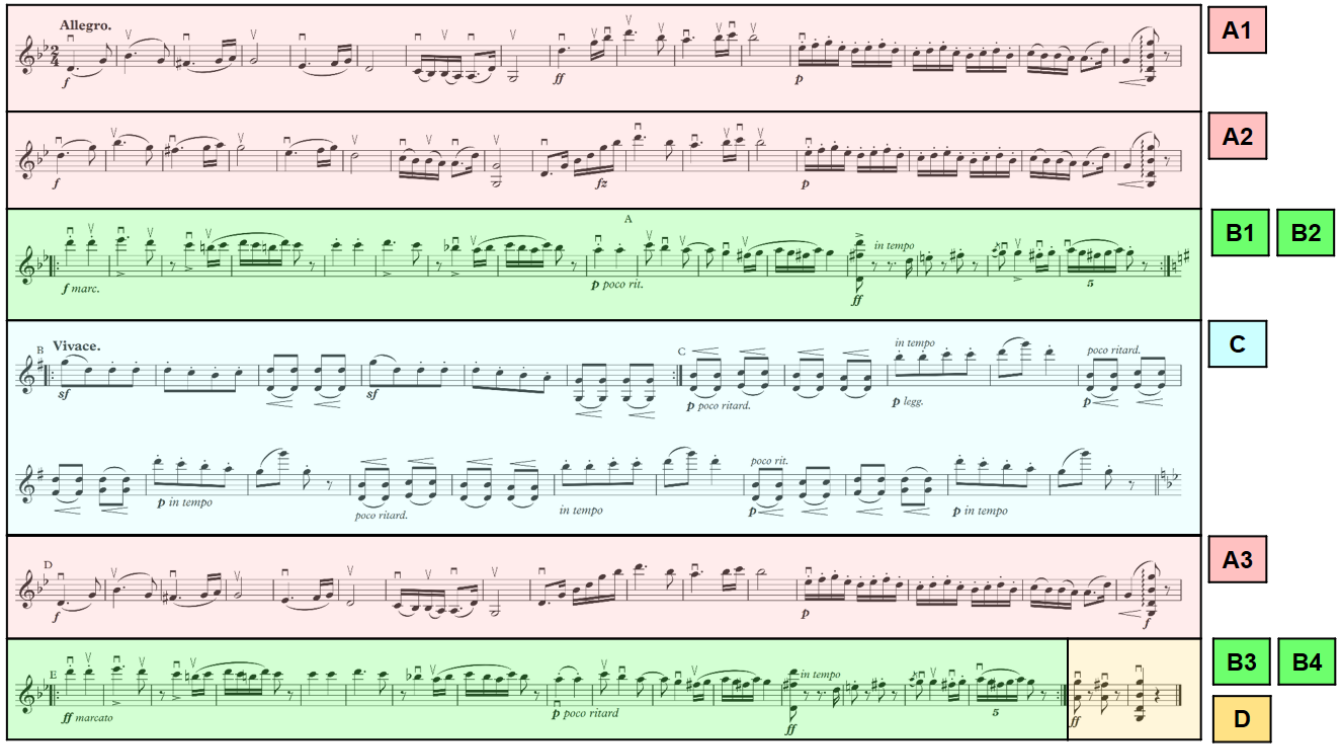


Figure 4.5 following [Müller, FMP, Springer 2015]

(Source: Müller & Zalkow, 2019)

Hierarchical music segmentation

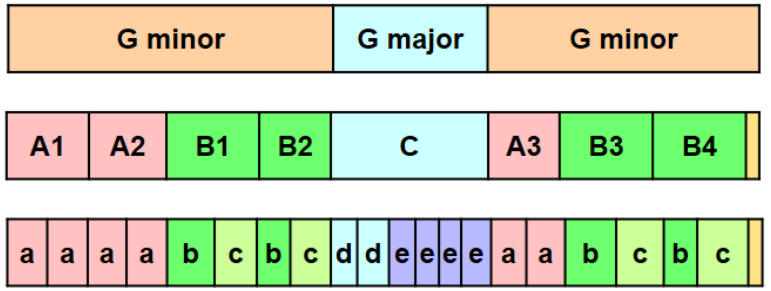


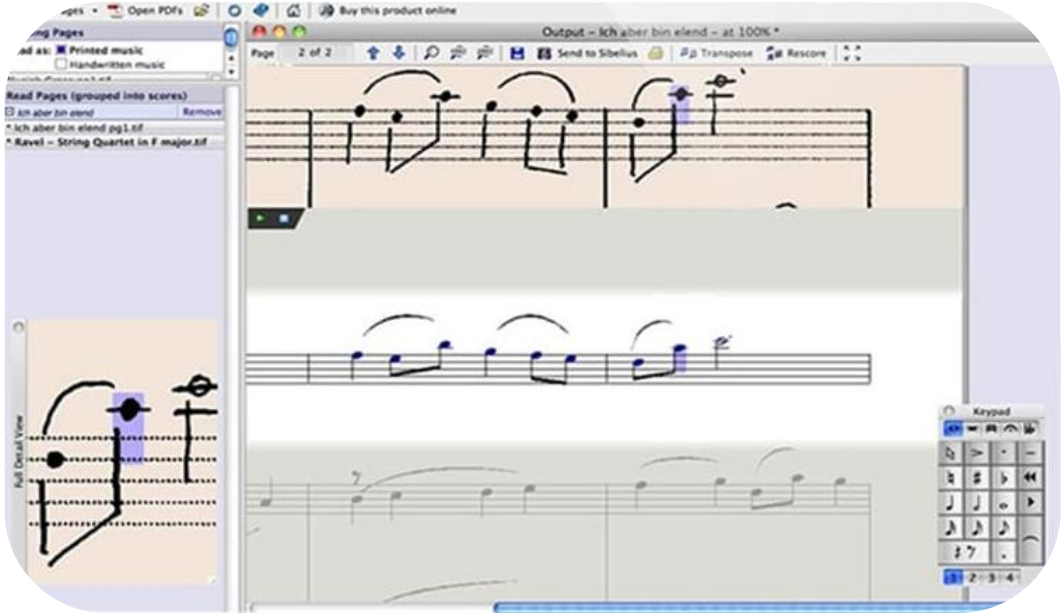
Figure 4.28 from [Müller, FMP, Springer 2015]

(Source: Müller & Zalkow, 2019)

Meinard Müller and Frank Zalkow, "FMP Notebooks: Educational Material for Teaching and Learning Fundamentals of Music Processing," *ISMIR*, 2019.

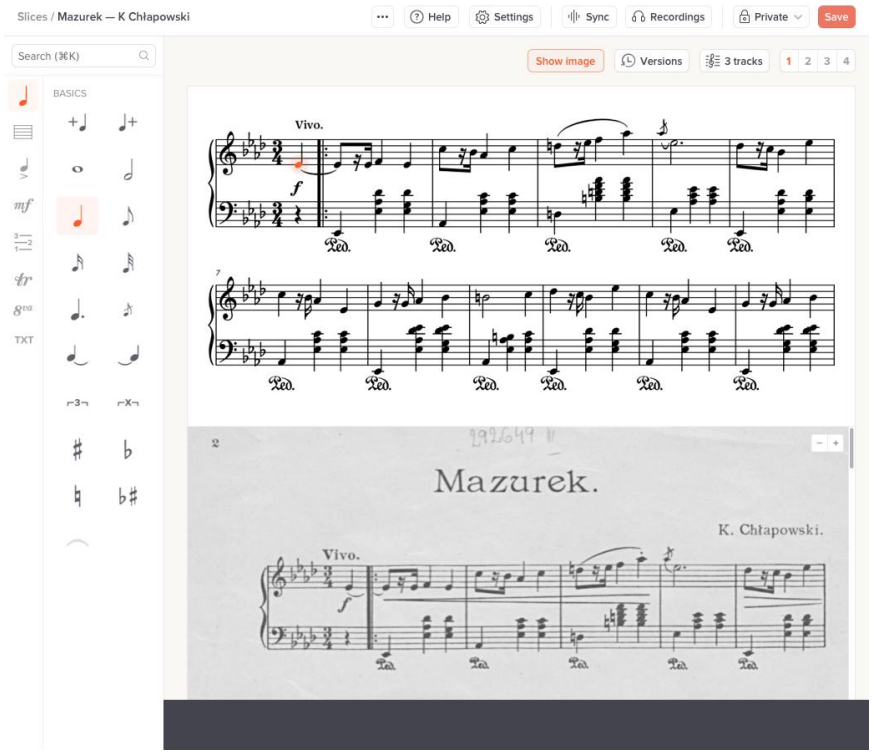
Optical Music Recognition (OMR)

PhotoScore & NotateMe in Sibelius



(Source: Avid)

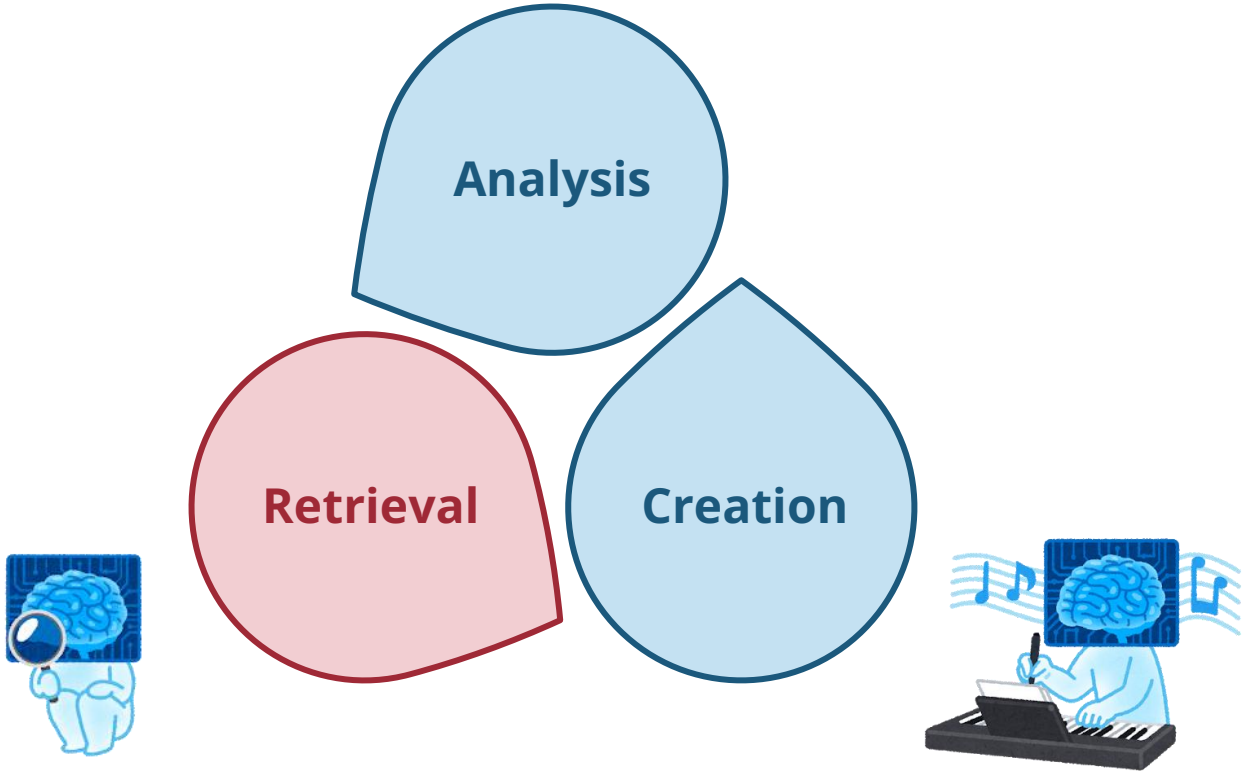
Soundslice



(Source: Soundslice)

avid.com/products/photoscore-and-notateme-ultimate
soundslice.com/sheet-music-scanner/


APPLE
DOG


Analysis
Retrieval
Creation

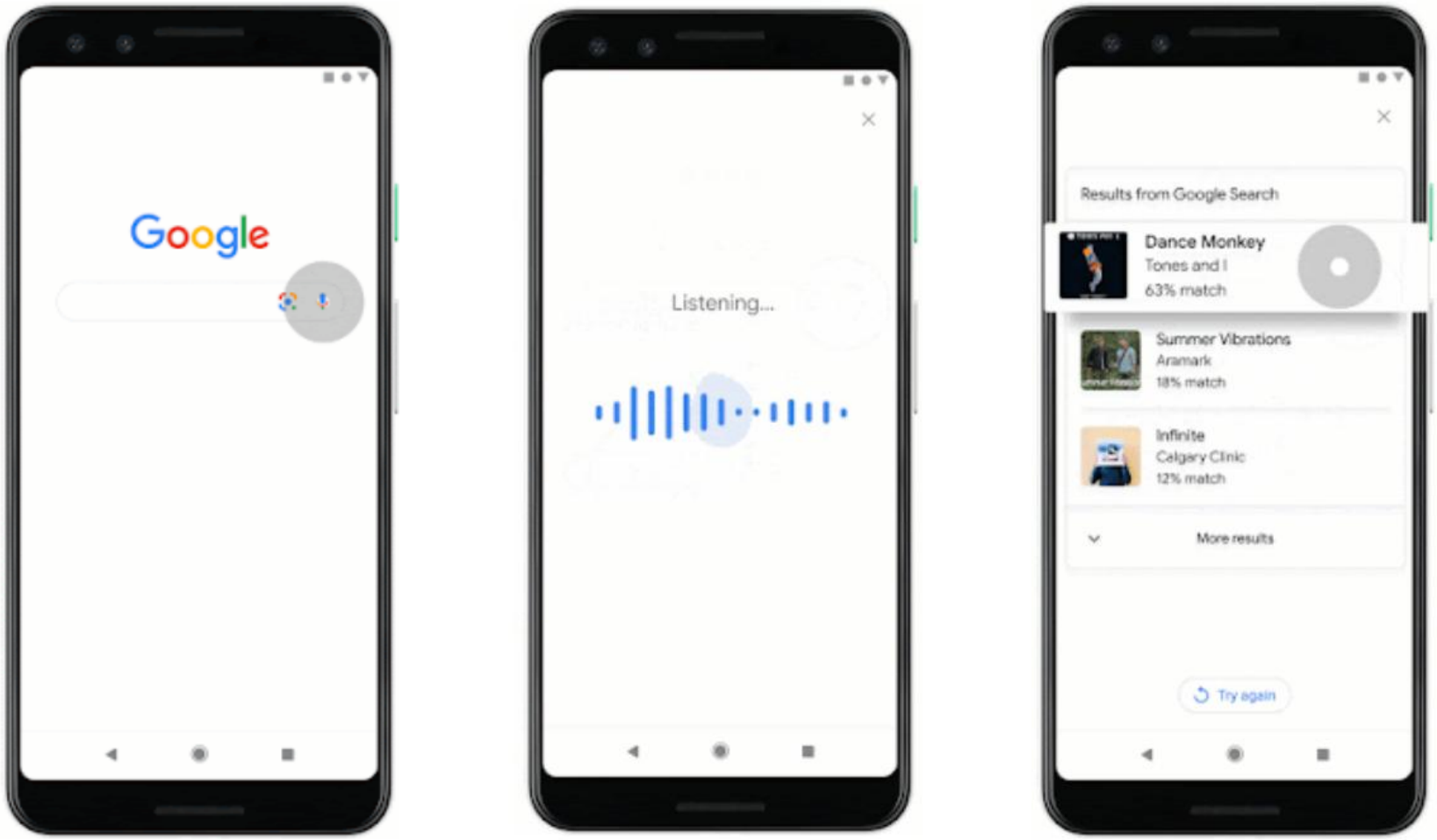
Shazam & Siri

(Source: Shazam User Guide)

(Source: OSXDaily)

osxdaily.com/2017/09/08/identify-song-playing-mac-siri/
support.apple.com/en-is/guide/shazam/dev9748744b6/web

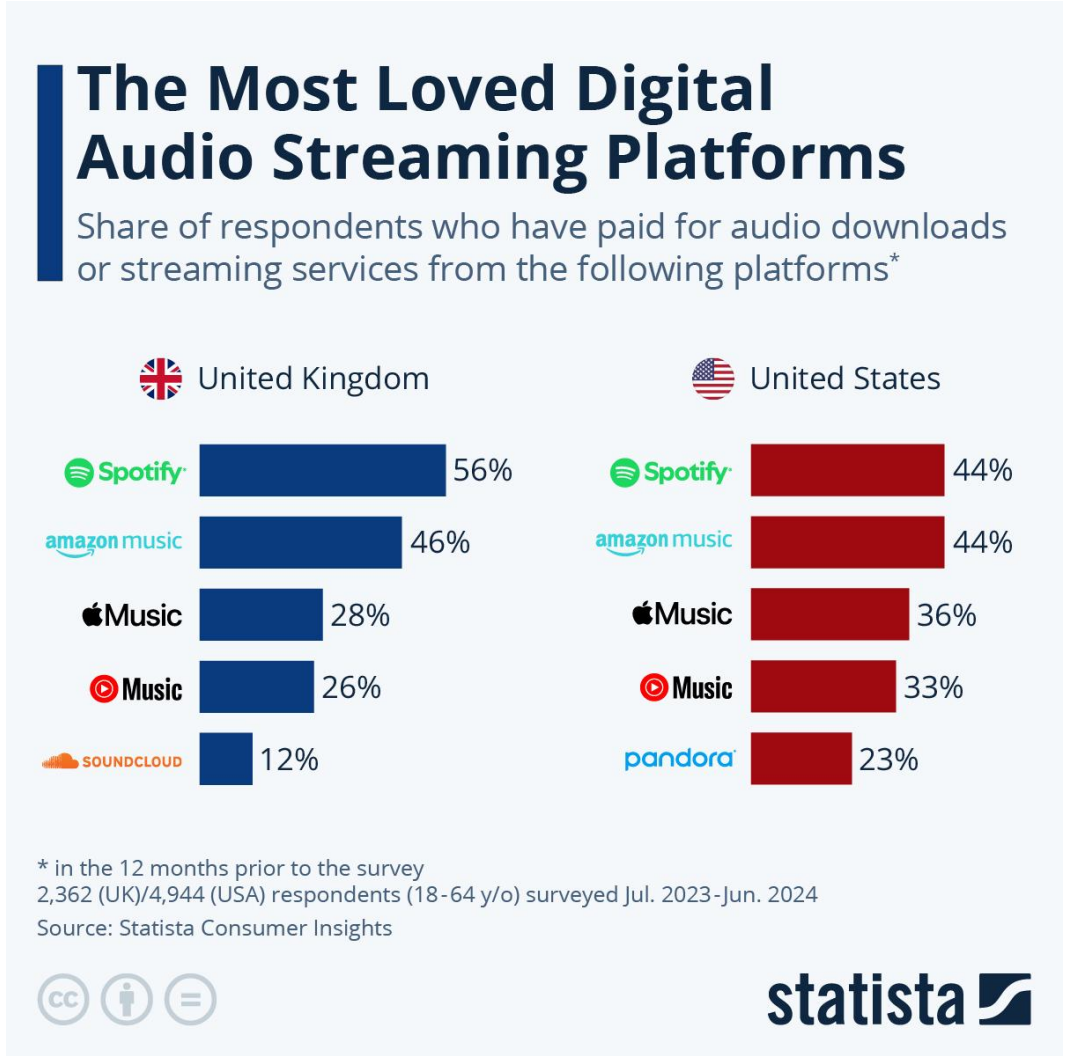
Hum to Search (Google)



(Source: Google Research Blog)

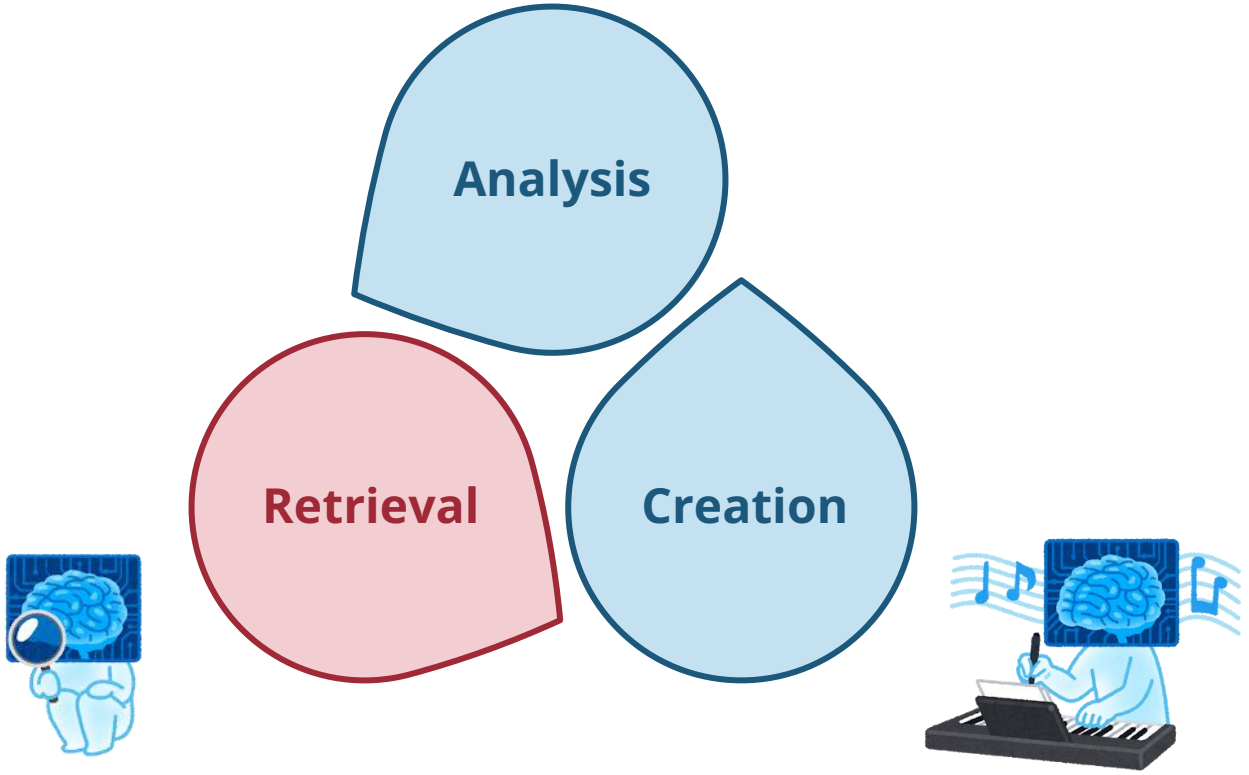
research.google/blog/the-machine-learning-behind-hum-to-search/

Music Recommendation



Anna Fleck, "The Most Loved Digital Audio Streaming Platforms," *Statista*, September 30, 2024.


APPLE
DOG


Analysis
Retrieval
Creation

Next Lecture

AI & ML Fundamentals