PAT 463/563 (Fall 2025)

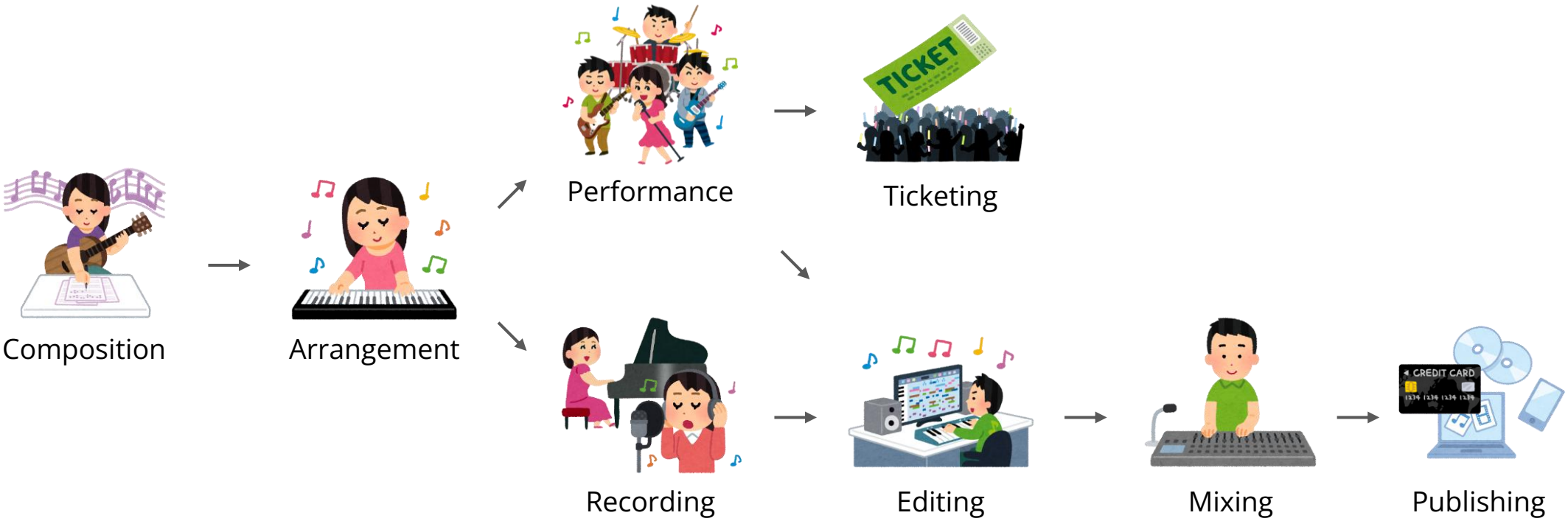
Music & AI

Lecture 17: Music Search & Recommendation

Instructor: Hao-Wen Dong

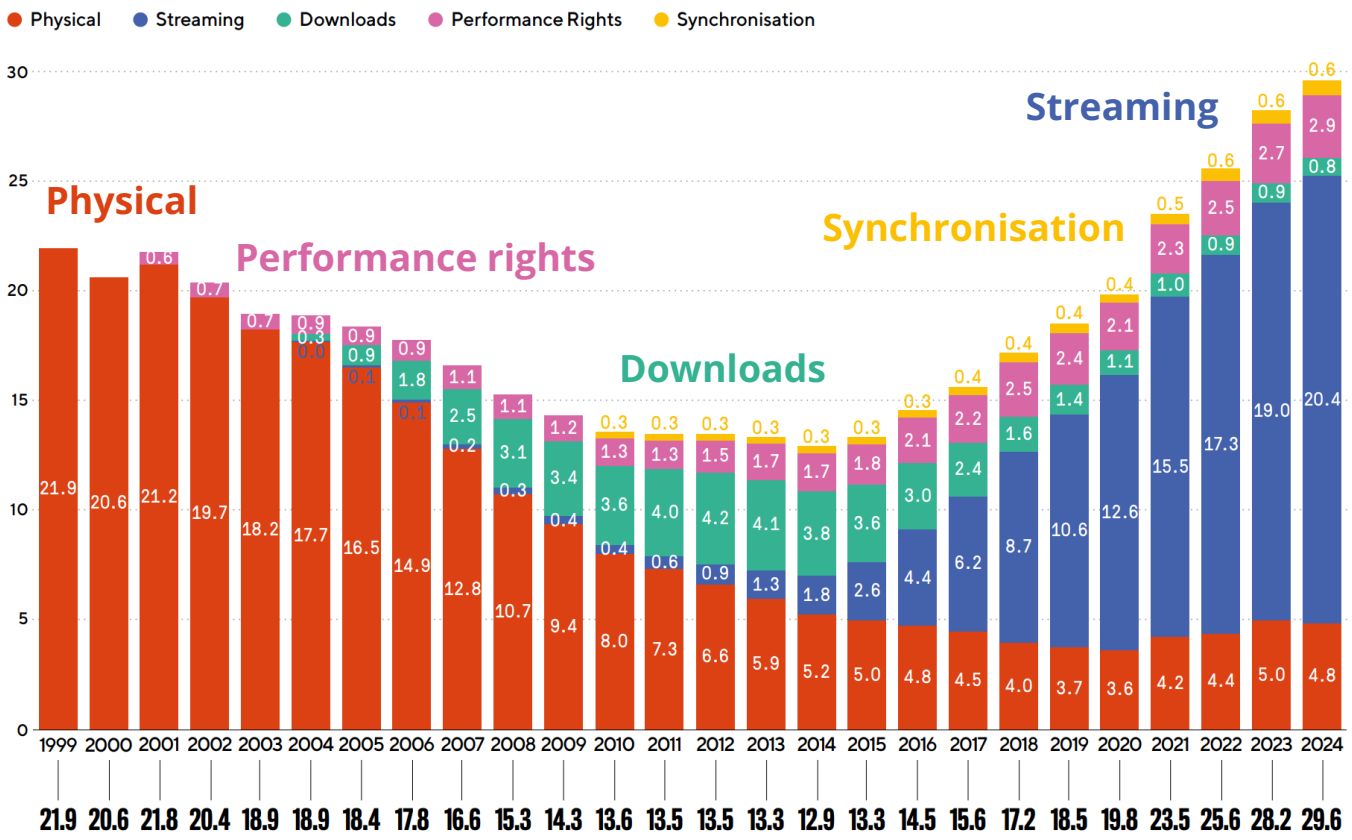
UNIVERSITY OF MICHIGAN

An Overly-Simplified Music Production Workflow

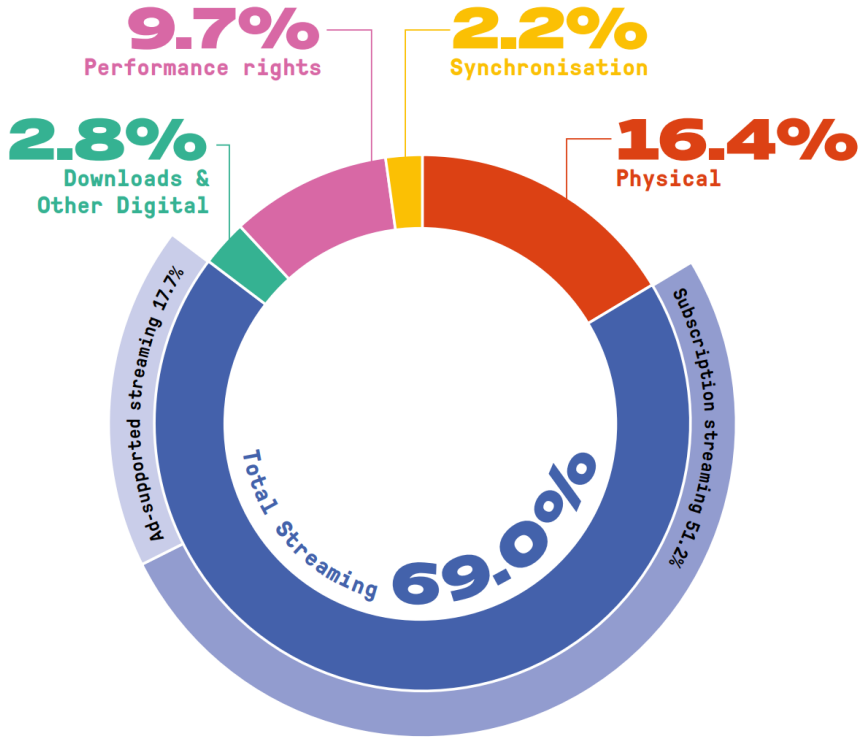


Global Recorded Music Industry Revenues

Global recorded music industry revenues 1999-2024 (US$ billions)

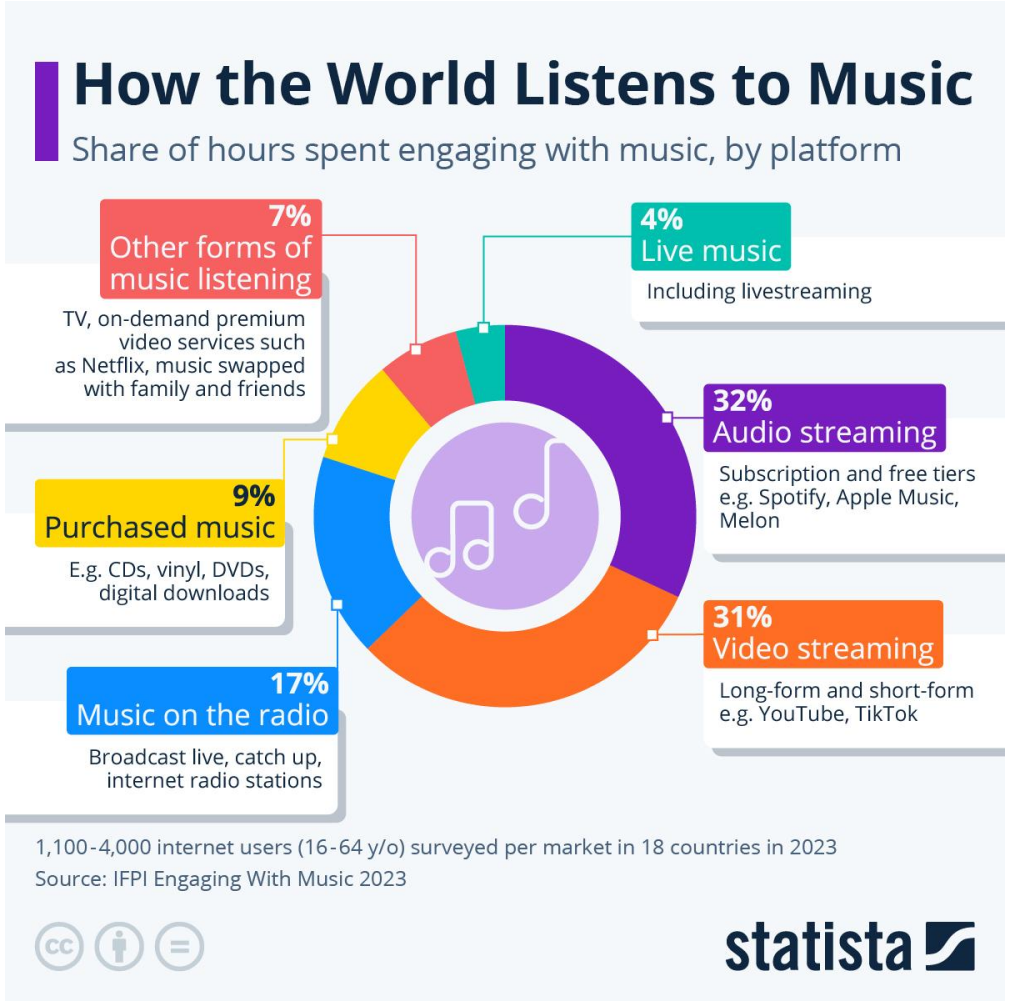


Global recorded music revenue share by format - 2024



Global Music Report 2025, IFPI, 2025

How the World Listens to Music?



(Source: Statista)

statista.com/chart/32880/share-of-hours-spent-engaging-with-music/

Music Fingerprinting

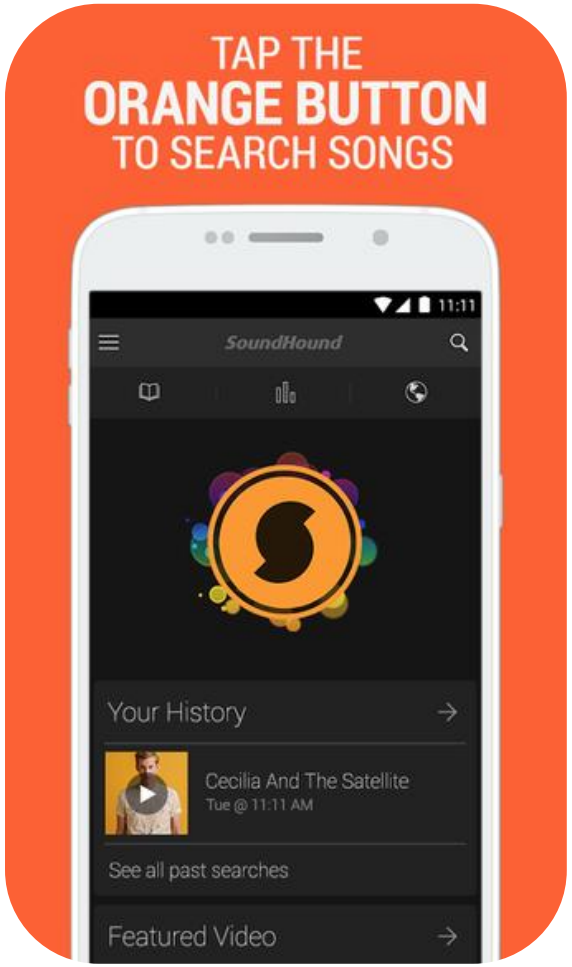
Shazam & Siri

(Source: Shazam User Guide)

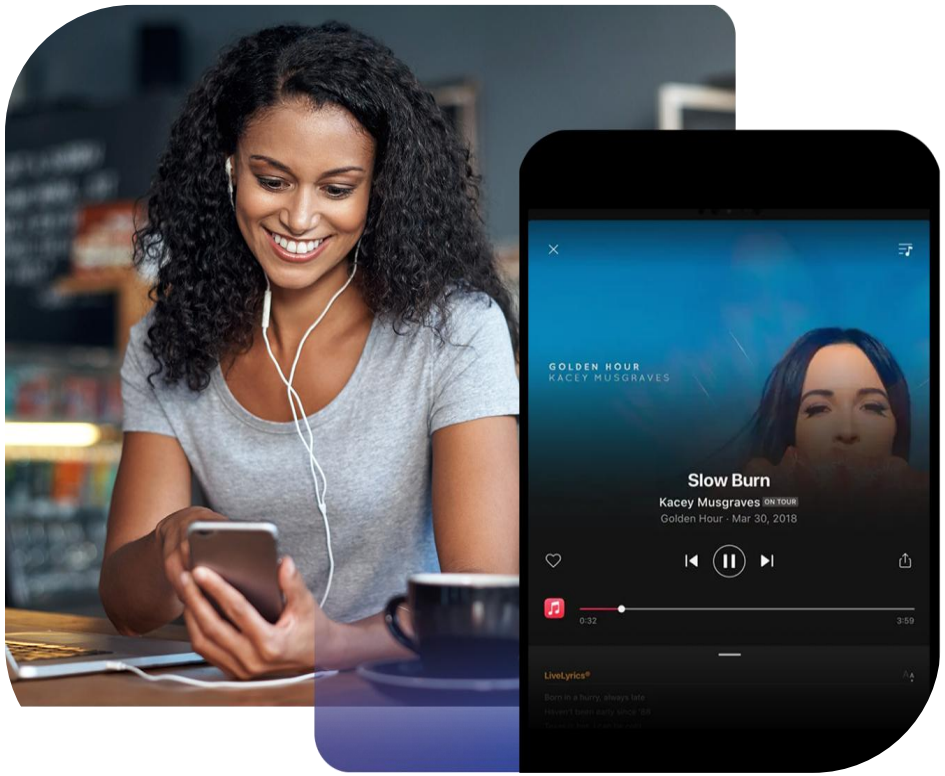
(Source: OSXDaily)

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

SoundHound



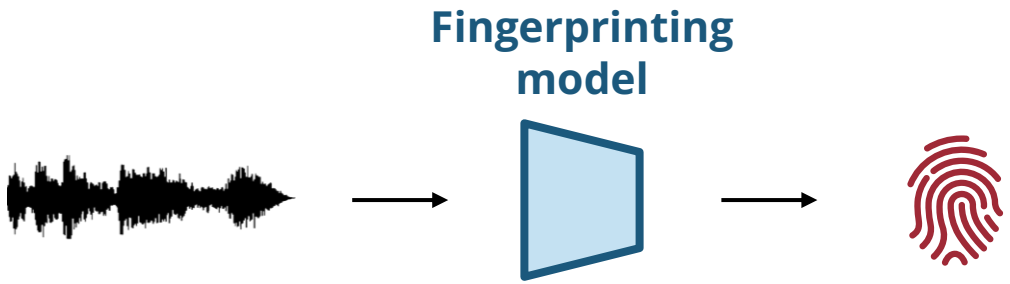
(Source: CNet)



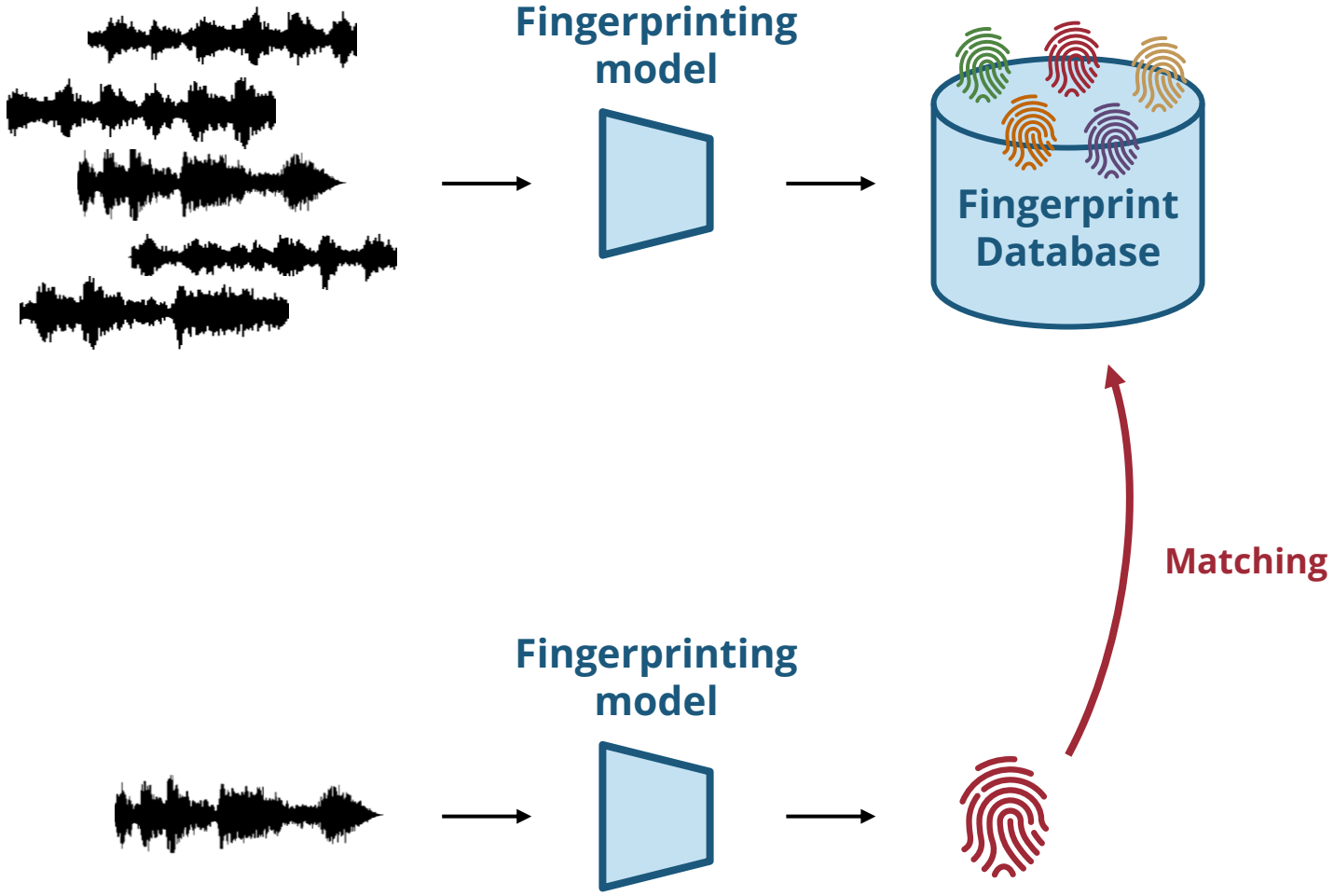
(Source: SoundHound)

download.cnet.com/soundhound-music-discovery-hands-free-player/3000-2141_4-77341748.html
soundhound.com/soundhound/

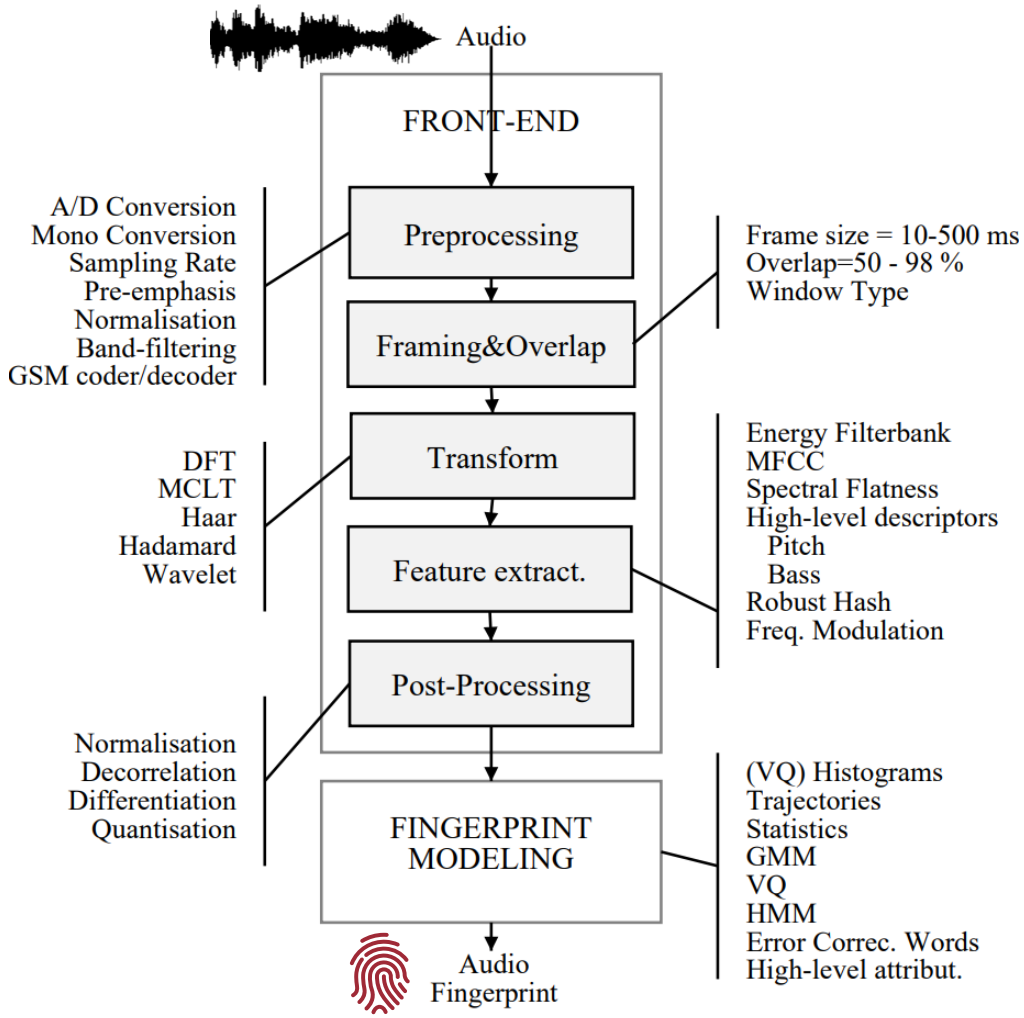
Audio Fingerprinting



Audio Fingerprinting for Audio Identification



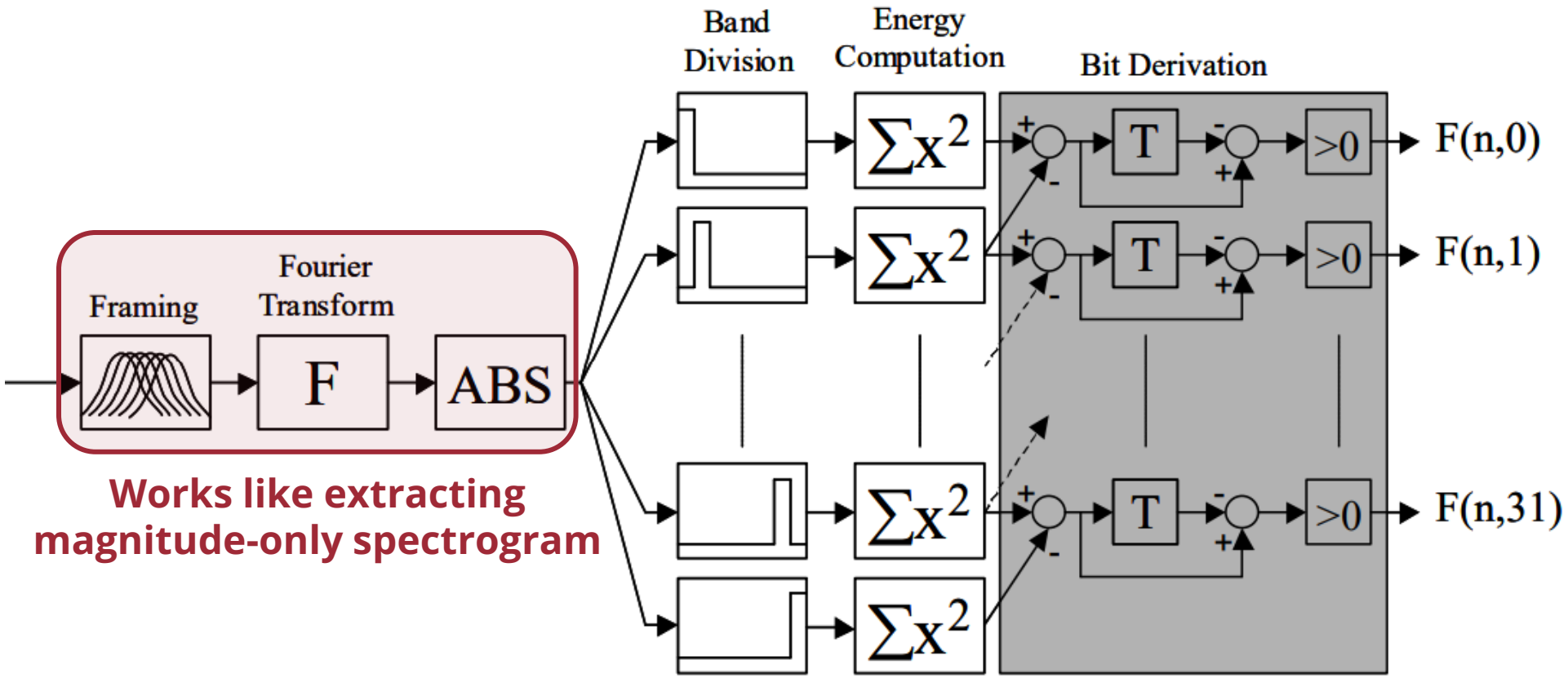
Common Framework of Audio Fingerprinting



(Source: Cano et al., 2002)

Pedro Cano, Eloi Batlle, Ton Kalker, and Jaap Haitsm, "A Review of Algorithms for Audio Fingerprinting," *MMSP*, 2002.

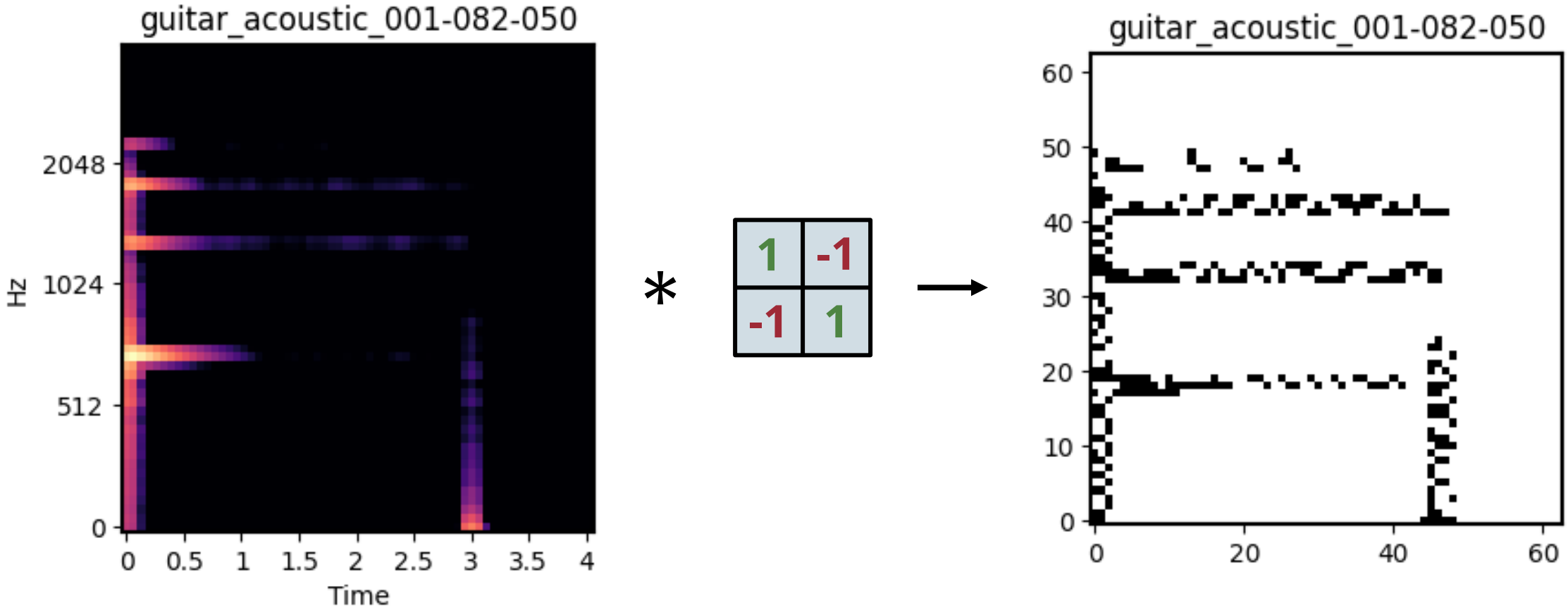
Energy Difference-based Fingerprinting (Haitsma et al., 2002)



(Source: Haitsma et al., 2002)

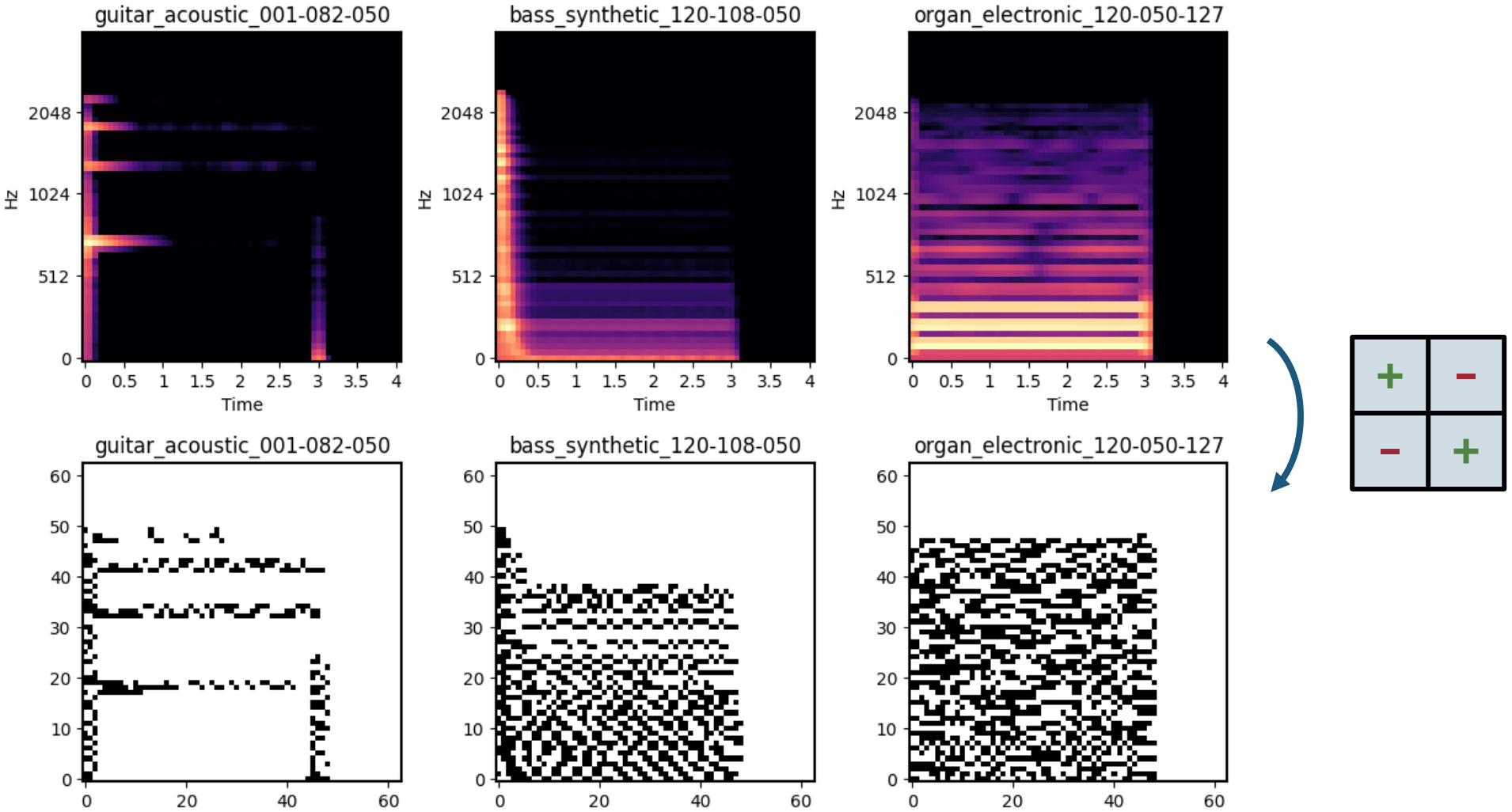
Jaap Haitsma and Ton Kalker, "A Highly Robust Audio Fingerprinting System," *ISMIR*, 2002.

Energy Difference-based Fingerprinting (Haitsma et al., 2002)

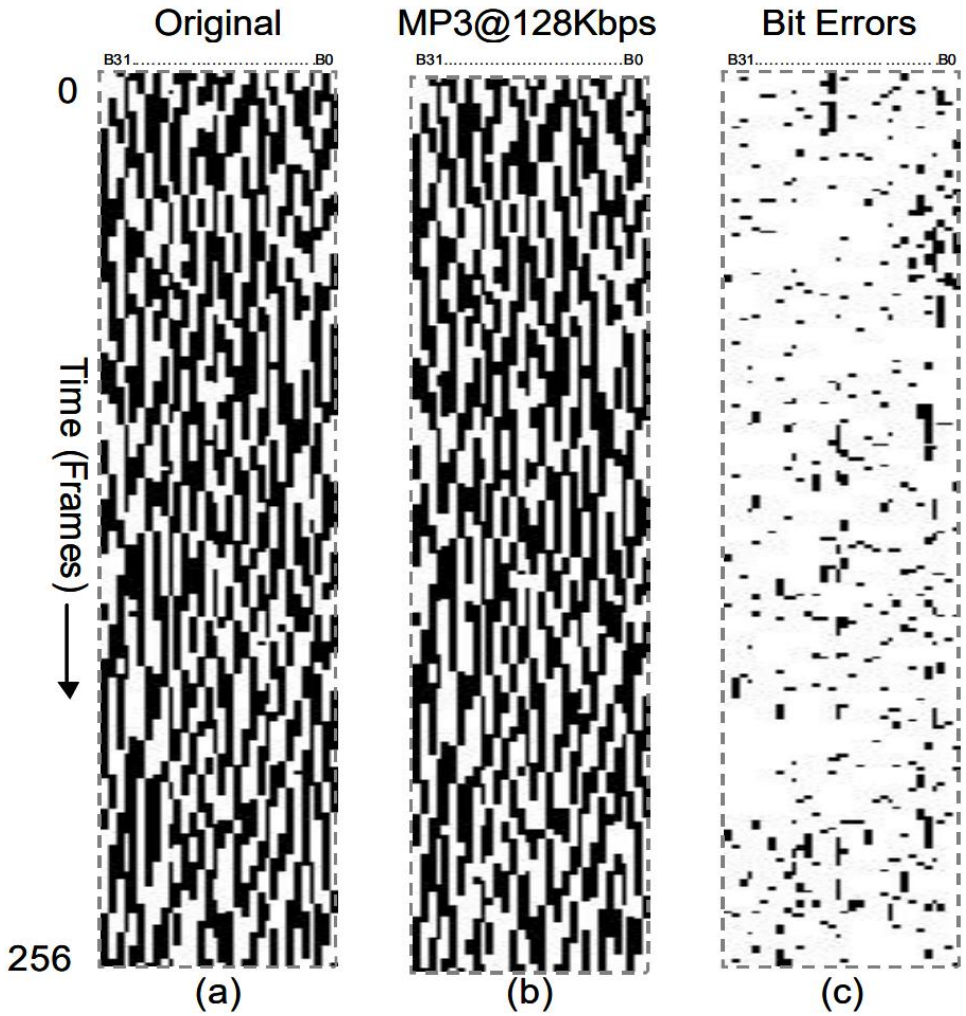


Jaap Haitsma and Ton Kalker, "A Highly Robust Audio Fingerprinting System," *ISMIR*, 2002.

Energy Difference-based Fingerprinting (Haitsma et al., 2002)



Energy Difference-based Fingerprinting (Haitsma et al., 2002)

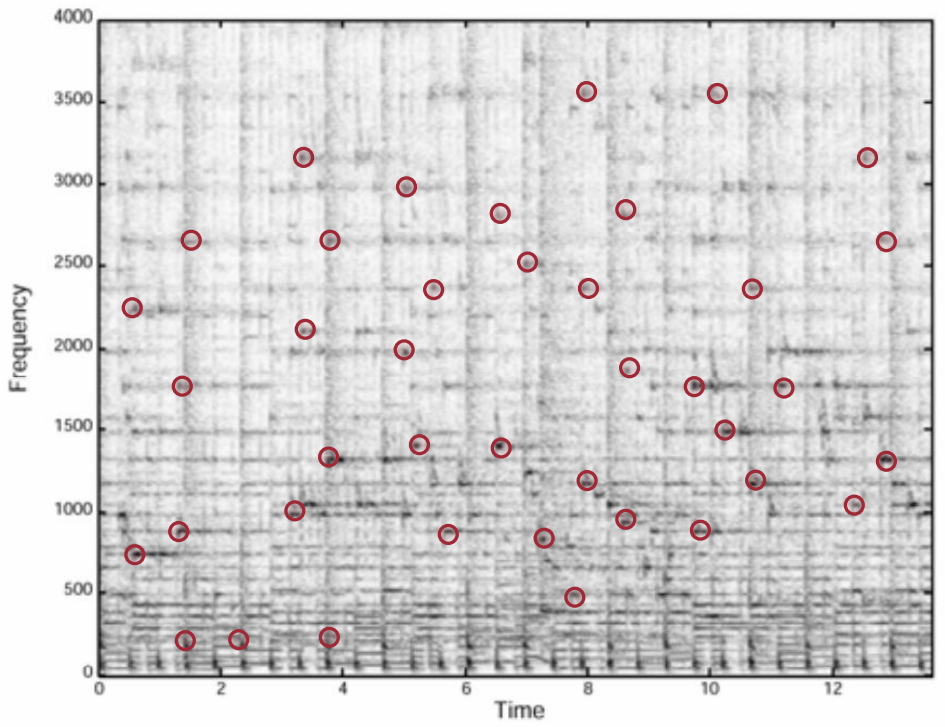


(Source: Haitsma et al., 2002)

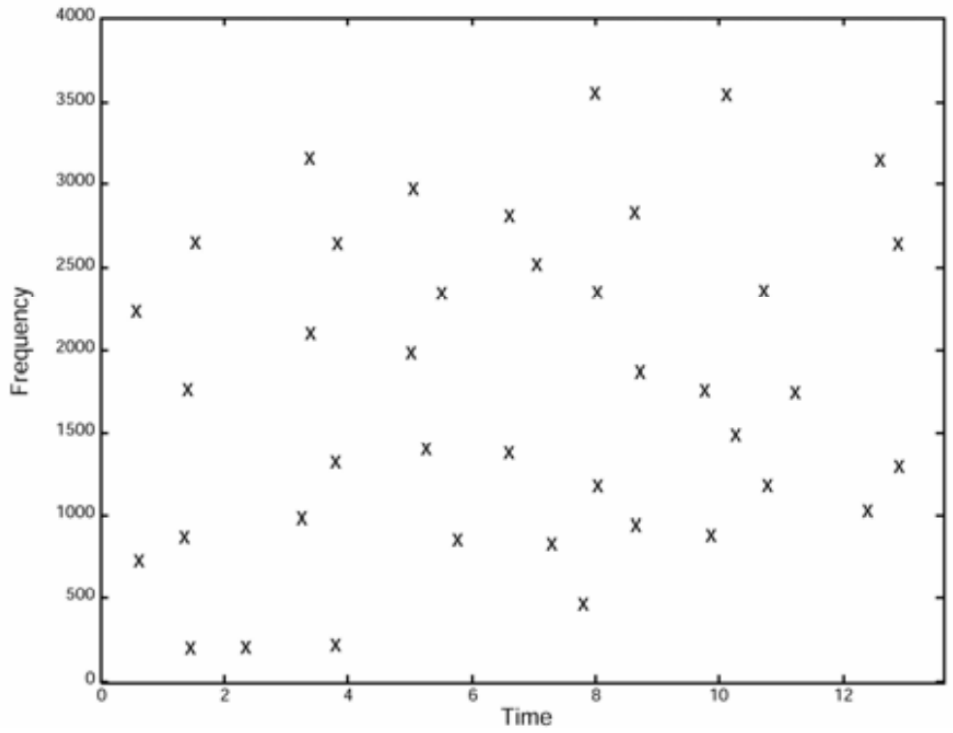
Jaap Haitsma and Ton Kalker, "A Highly Robust Audio Fingerprinting System," *ISMIR*, 2002.

Peak-based Audio Fingerprinting (Wang et al., 2003)

Spectrogram



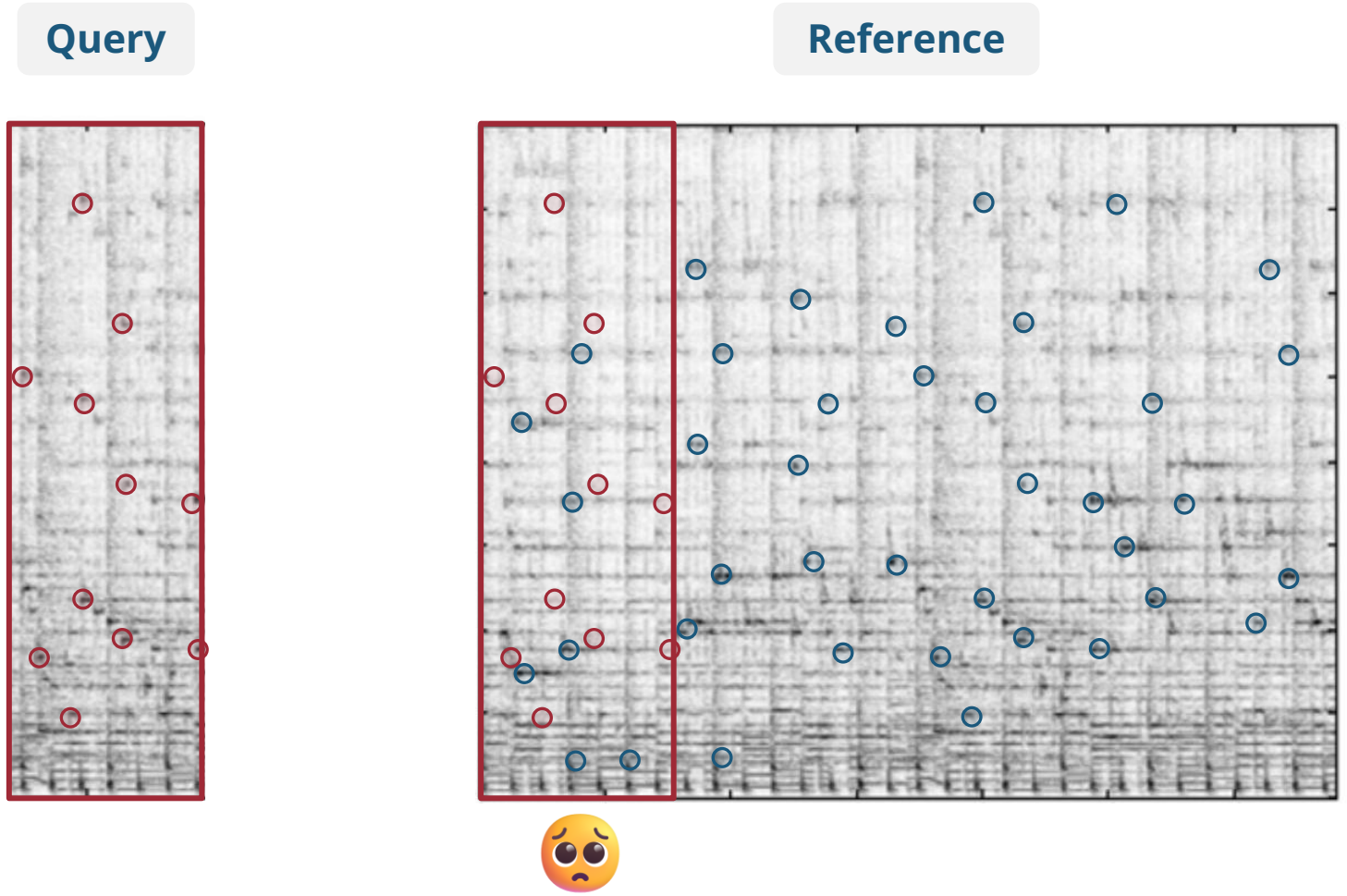
Constellation Map



(Source: Wang et al., 2003)

Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

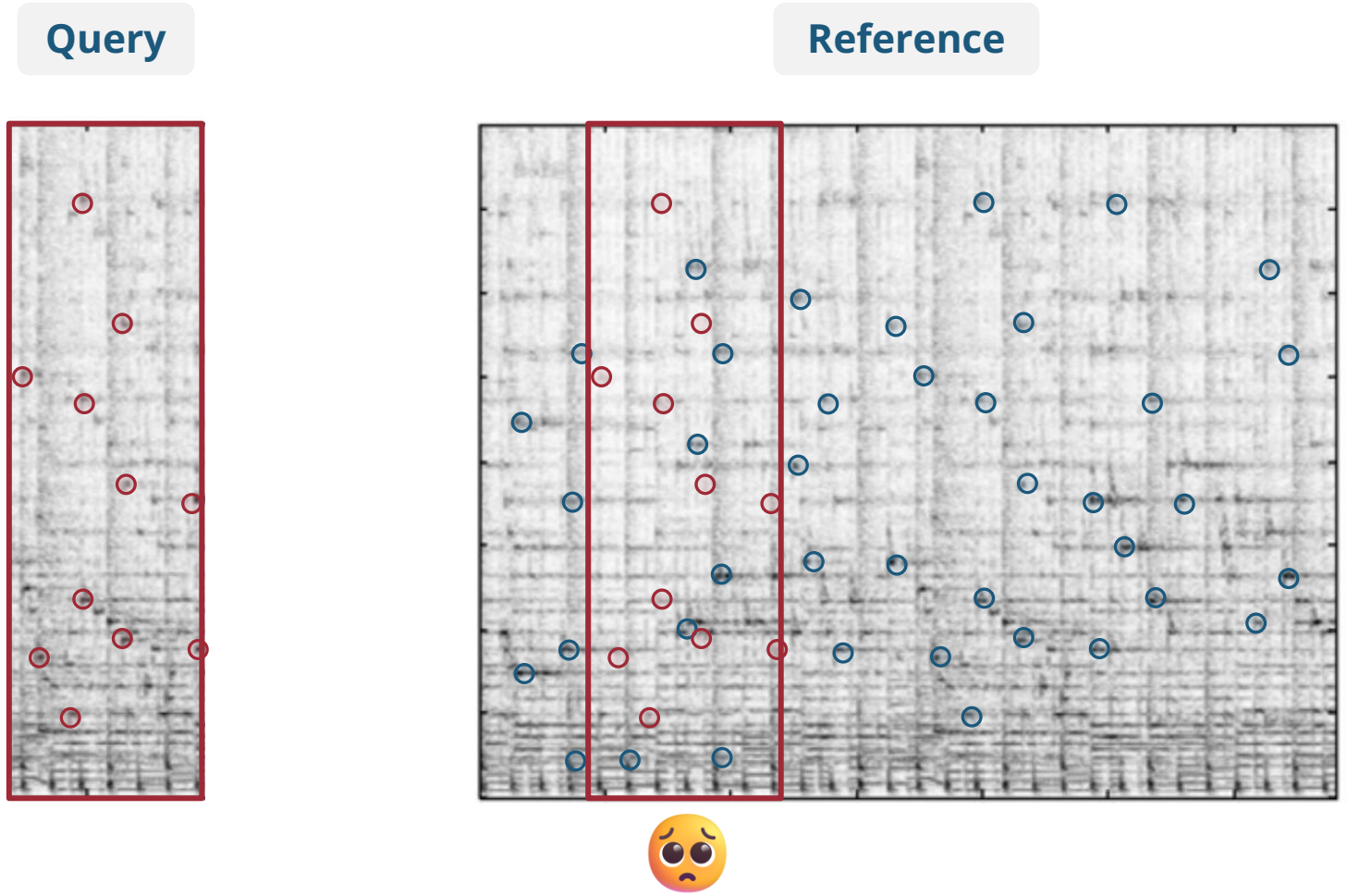
Peak-based Audio Fingerprinting (Wang et al., 2003)



(Source: Wang et al., 2003)

Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

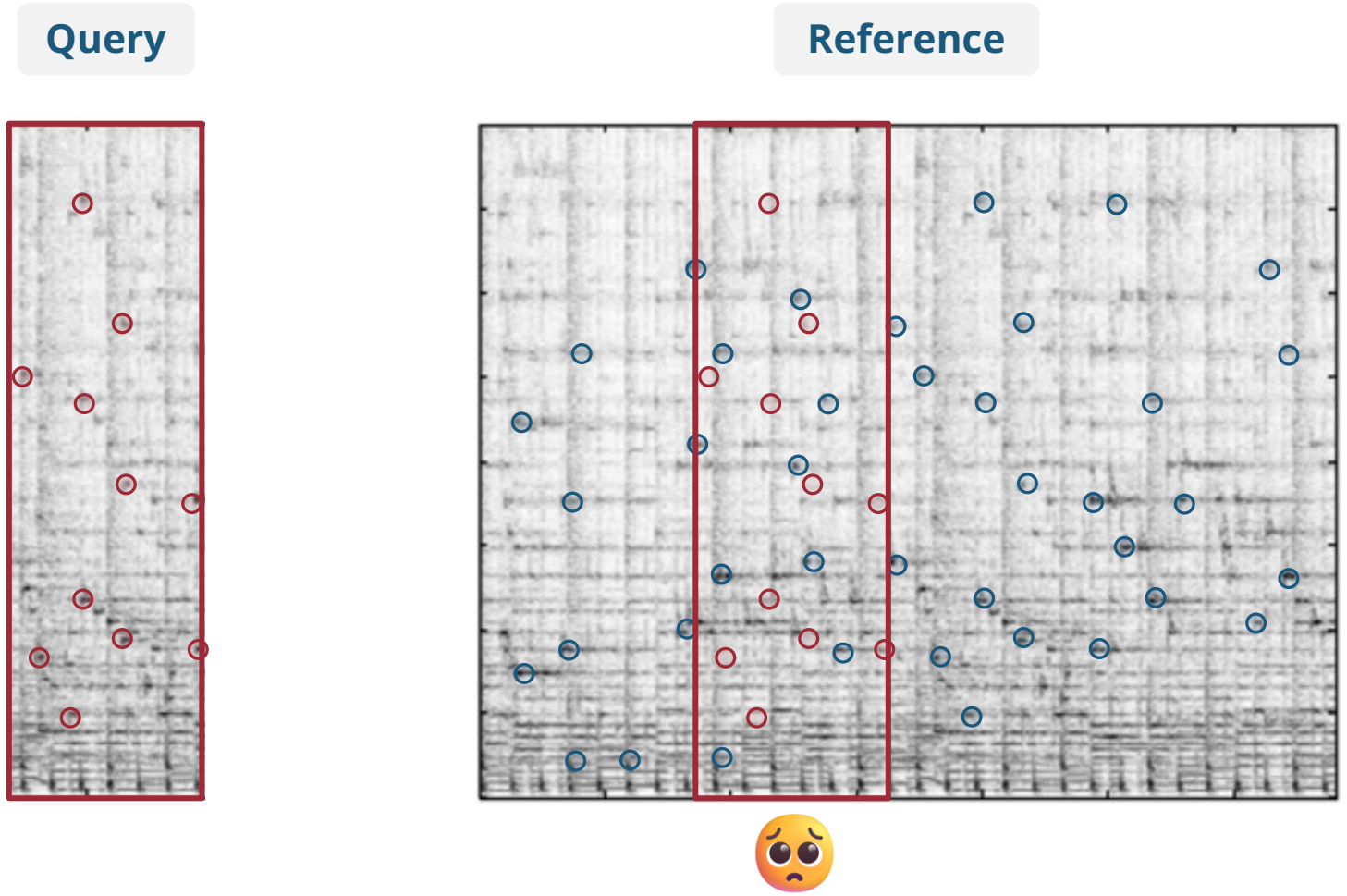
Peak-based Audio Fingerprinting (Wang et al., 2003)



(Source: Wang et al., 2003)

Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

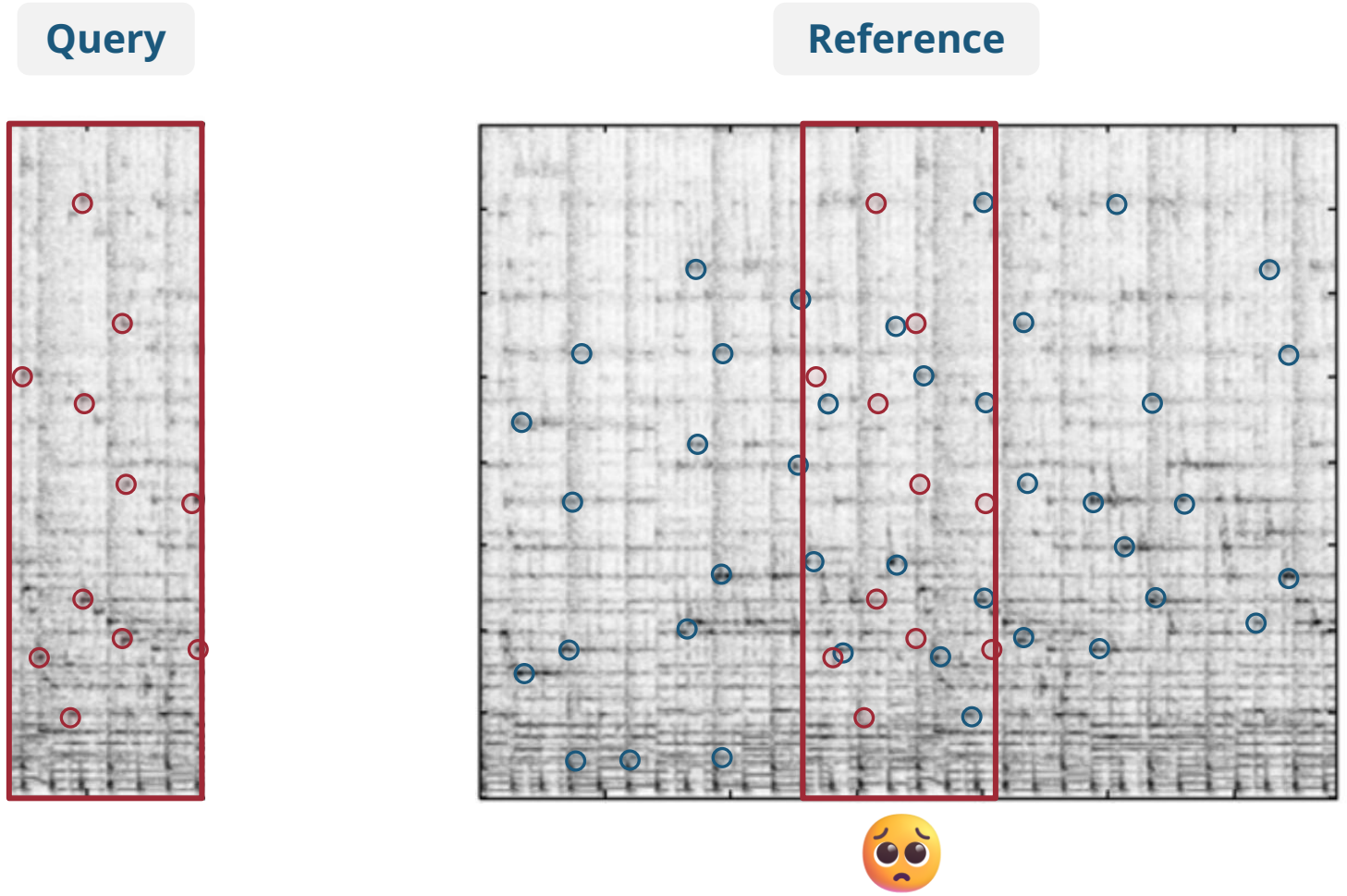
Peak-based Audio Fingerprinting (Wang et al., 2003)



(Source: Wang et al., 2003)

Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

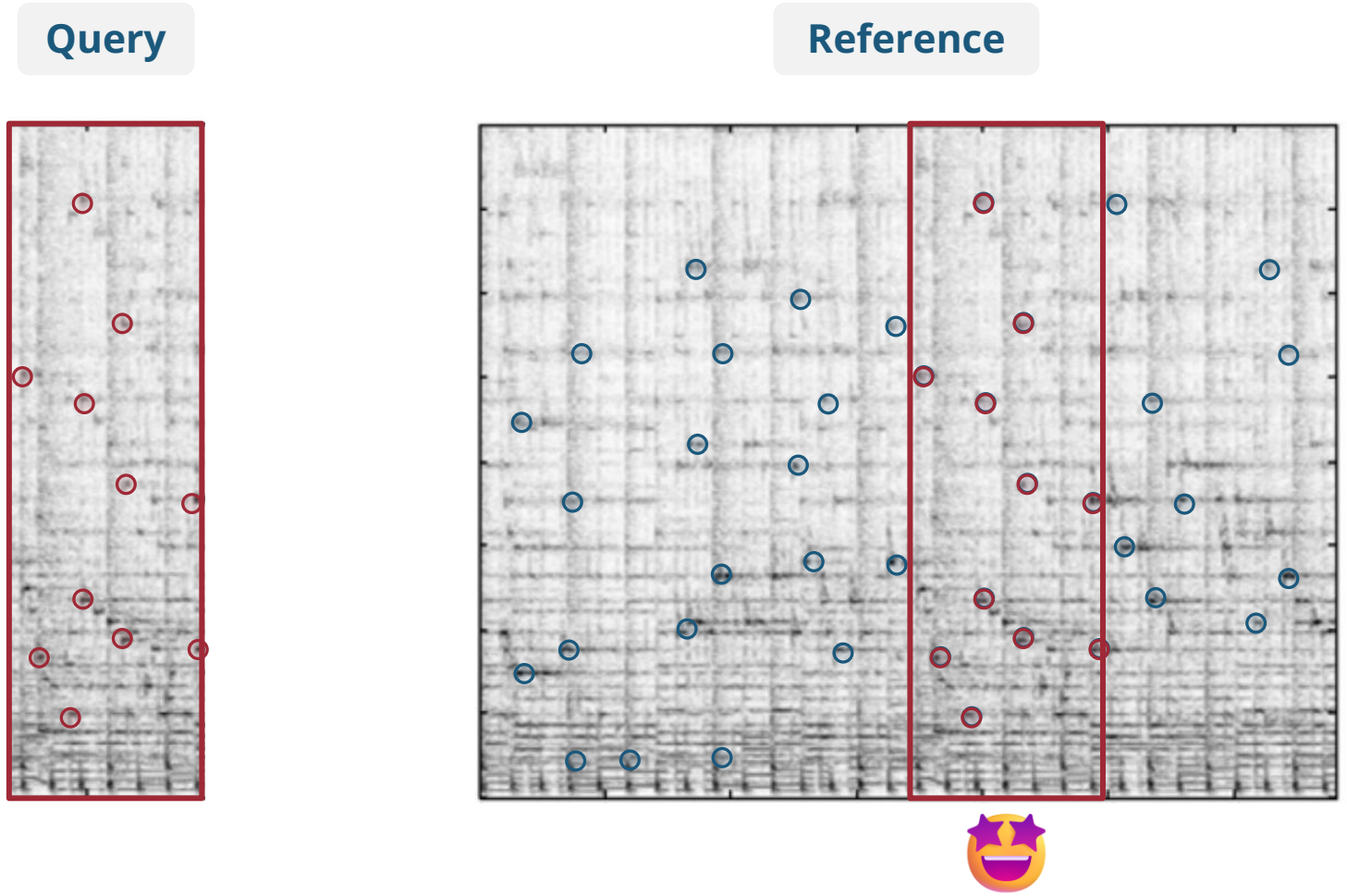
Peak-based Audio Fingerprinting (Wang et al., 2003)



(Source: Wang et al., 2003)

Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

Peak-based Audio Fingerprinting (Wang et al., 2003)

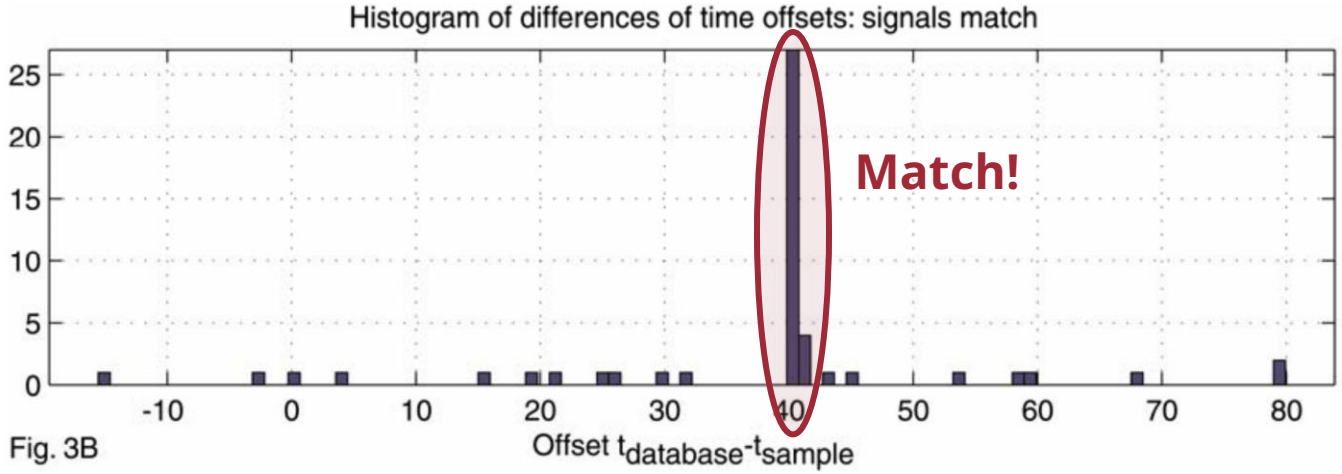


(Source: Wang et al., 2003)

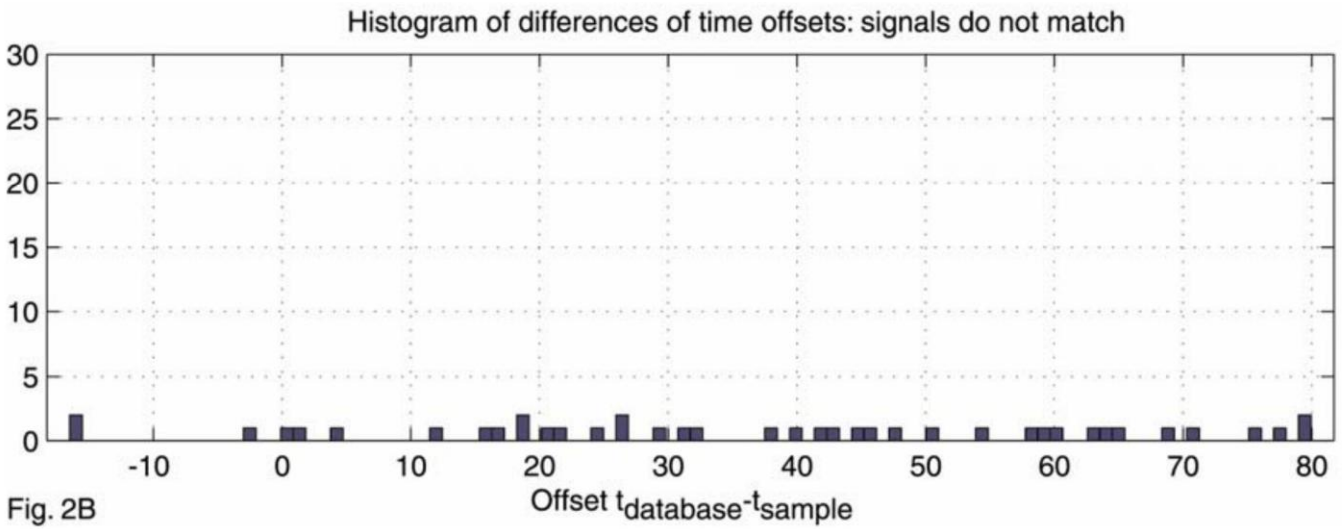
Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

Peak-based Audio Fingerprinting (Wang et al., 2003)

Match found



No match

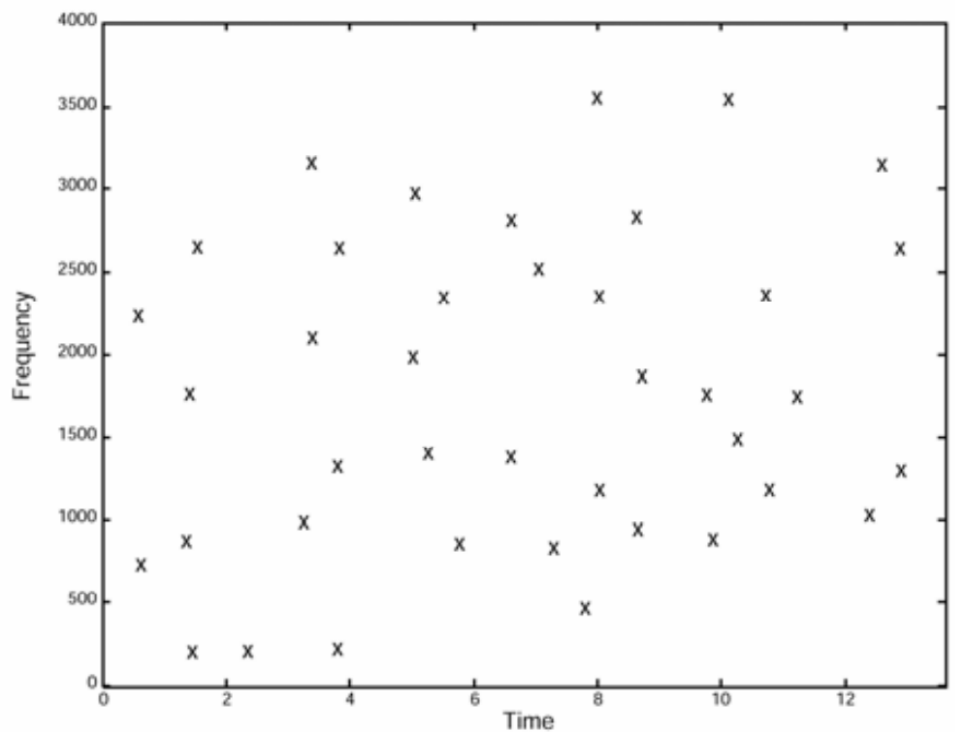


(Source: Wang et al., 2003)

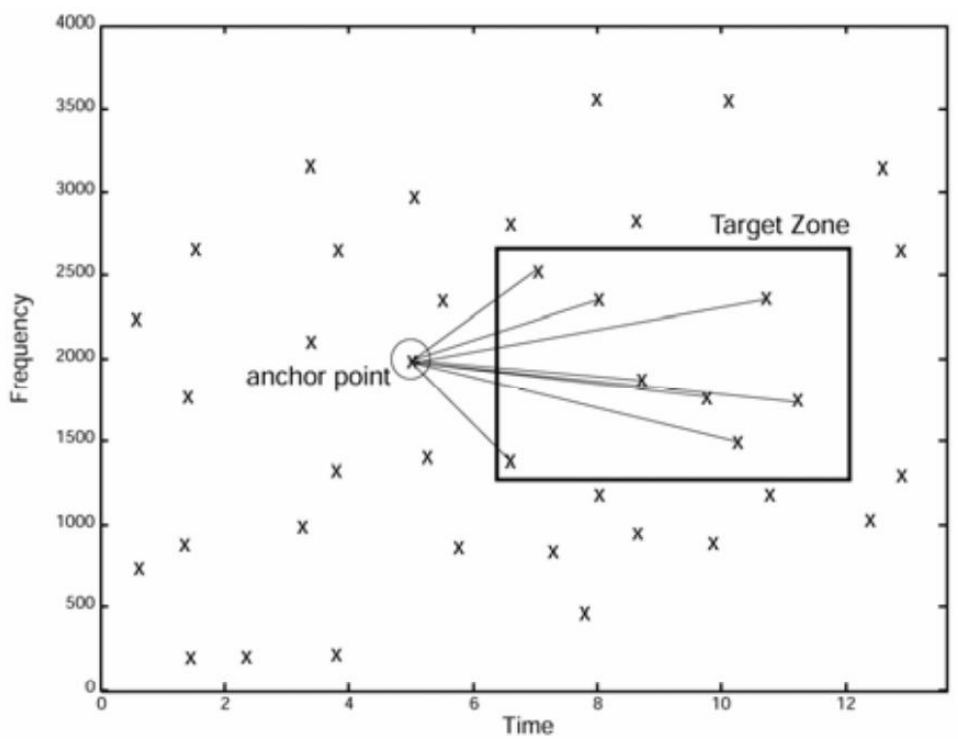
Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

Peak-based Audio Fingerprinting (Wang et al., 2003)

Constellation Map



Combinatorial Hashing

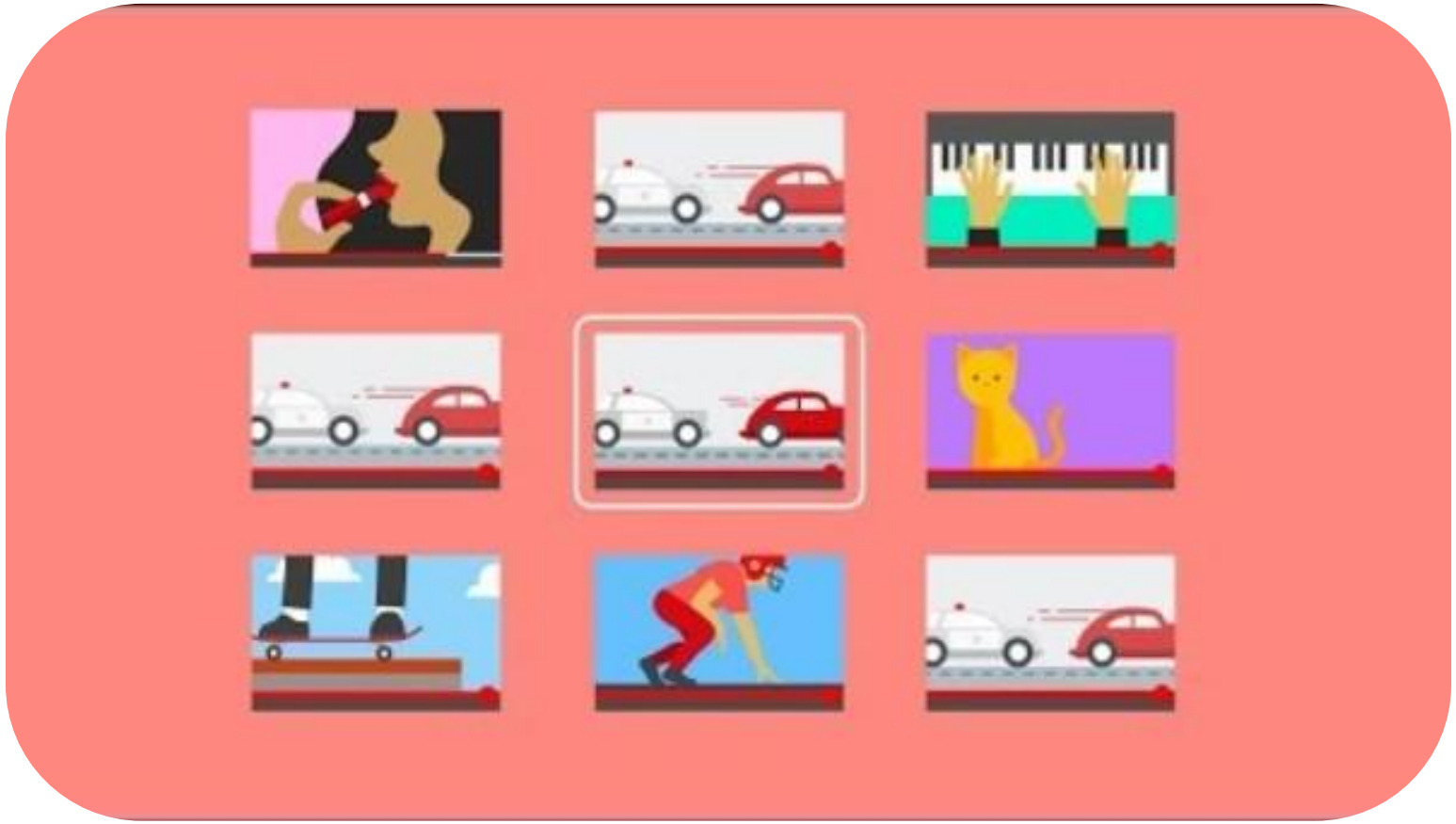


(Source: Wang et al., 2003)

Apply hashing algorithms to speed up the matching process!

Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

YouTube’s Content ID



youtu.be/9g2U12SsRns

Resources on Fingerprinting

- Pedro Cano, Eloi Batlle, Ton Kalker, and Jaap Haitsma, "A Review of Audio Fingerprinting," Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology, 2005.

Version Identification

Version Identification: Beyond Fingerprinting & Covers

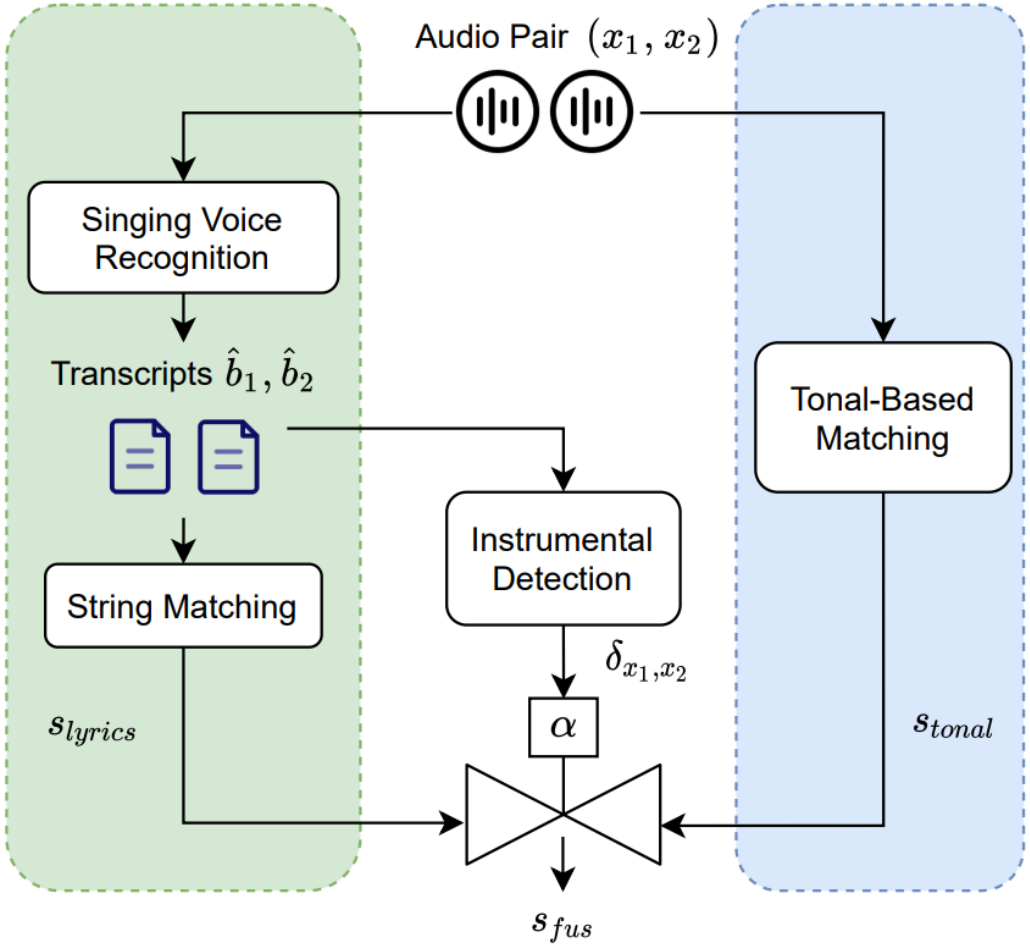
Musical Characteristic	Version Type																	
	Duplicate	Remaster	Radio Edit	Translation	Performance	Demo	Parody	Within-Genre	Karaoke	Live	Standard	Mashup	Acoustic	Medley	Remix	Cross-Genre	Arrangement	Quotation
Melody	0	0	0	0	0	1	0	1	2	1	1	0	1	1	1	2	2	2
Harmony	0	0	0	0	0	1	0	1	0	0	2	0	1	1	2	2	2	3
Tempo	0	0	0	0	2	1	1	1	0	2	1	3	2	2	3	2	2	3
Timing	0	0	0	0	2	1	1	1	0	2	1	3	2	2	2	3	3	3
Structure	0	0	1	0	1	1	1	1	1	2	2	3	2	3	3	2	3	3
Lyrics	0	0	1	3	0	1	3	0	3	1	0	0	0	1	1	1	1	2
Key	0	0	0	0	1	1	1	1	1	1	1	2	2	2	2	3	3	3
Timbre	0	0	0	0	1	1	1	1	2	2	3	2	3	2	3	3	3	3
Noise	0	1	1	1	3	3	2	3	2	3	3	2	3	3	2	3	3	3

Degree of Potential Difference	0	1	2	3
	Likely the same	May be variations	May be major differences	May be unrelated

(Source: Yesiler et al., 2021)

Furkan Yesiler, Guillaume Doras, Rachel M. Bittner, Christopher J. Tralie, and Joan Serrà, "Audio-based Musical Version Identification: Elements and Challenges," *IEEE Signal Processing Magazine*, 2021.

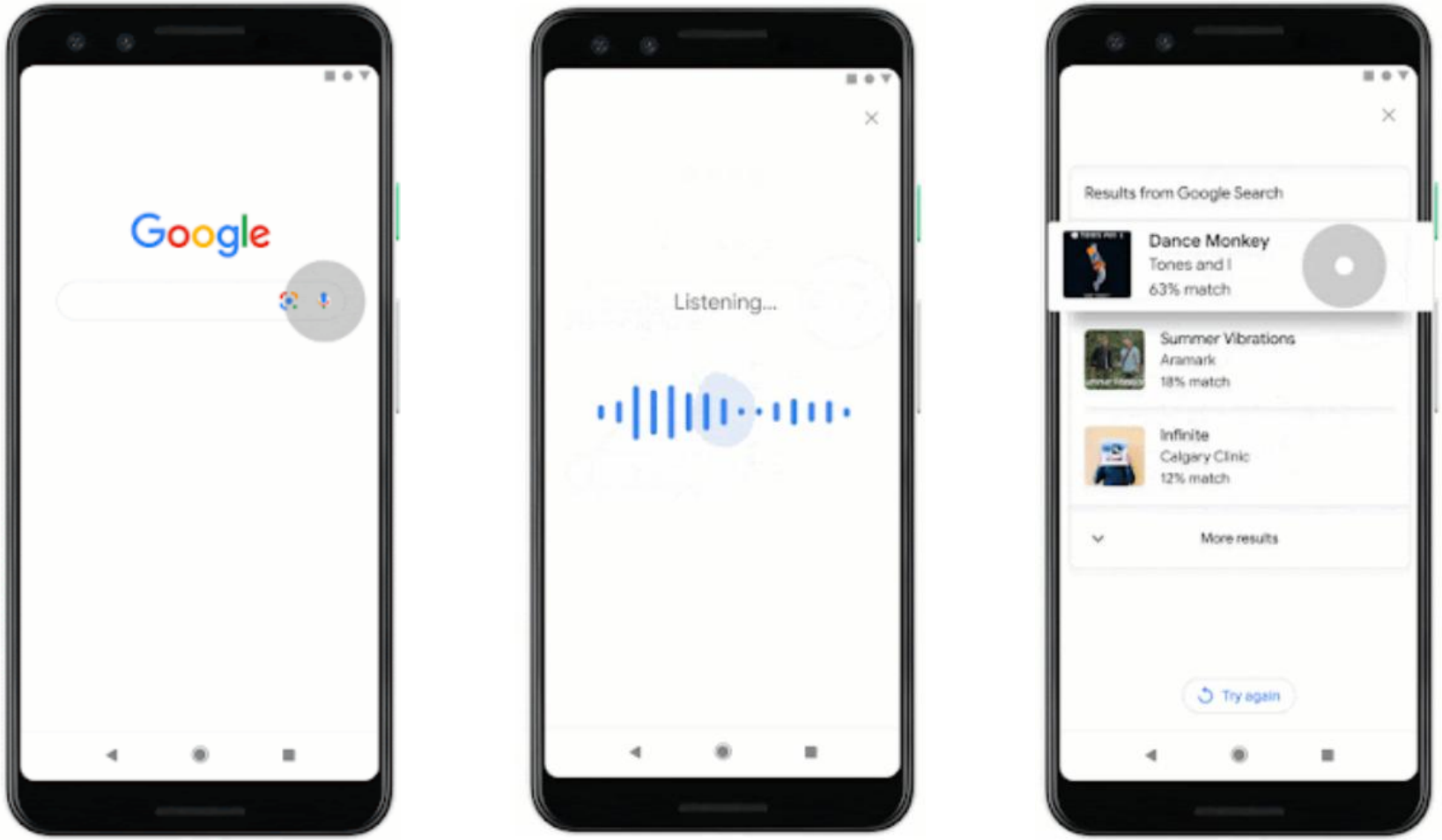
Cover Song Identification



(Source: Vaglio et al., 2021)

Andrea Vaglio, Romain Hennequin, Manuel Moussallam, and Gaël Richard, "The Words Remain the Same: Cover Detection with Lyrics Transcription," *ISMIR*, 2021.

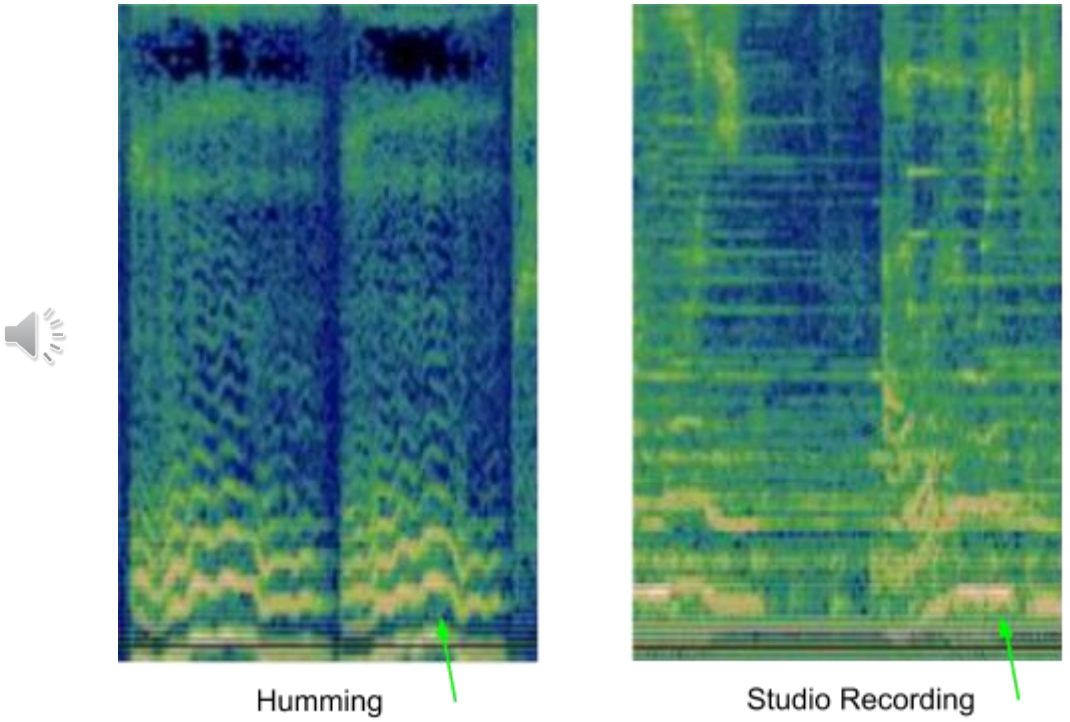
Google's Hum to Search (2020)



(Source: Google Research Blog)

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

Google's Hum to Search (2020)

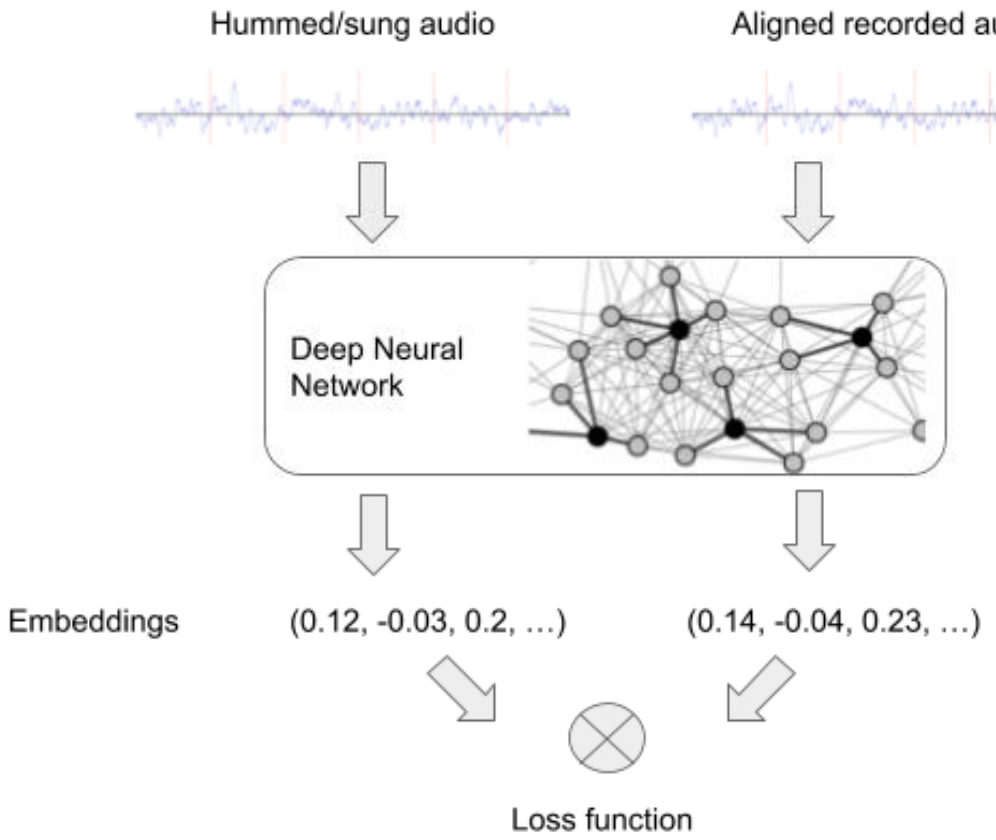


(Source: Google Research Blog)

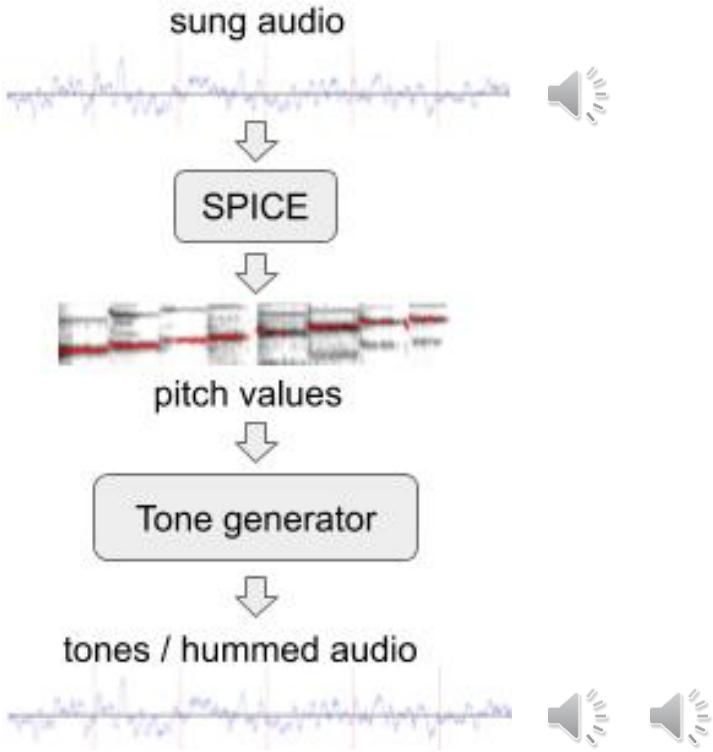
research.google/blog/the-machine-learning-behind-hum-to-search/

Google's Hum to Search (2020)

Audio encoder



Data augmentation



(Source: Google Research Blog)

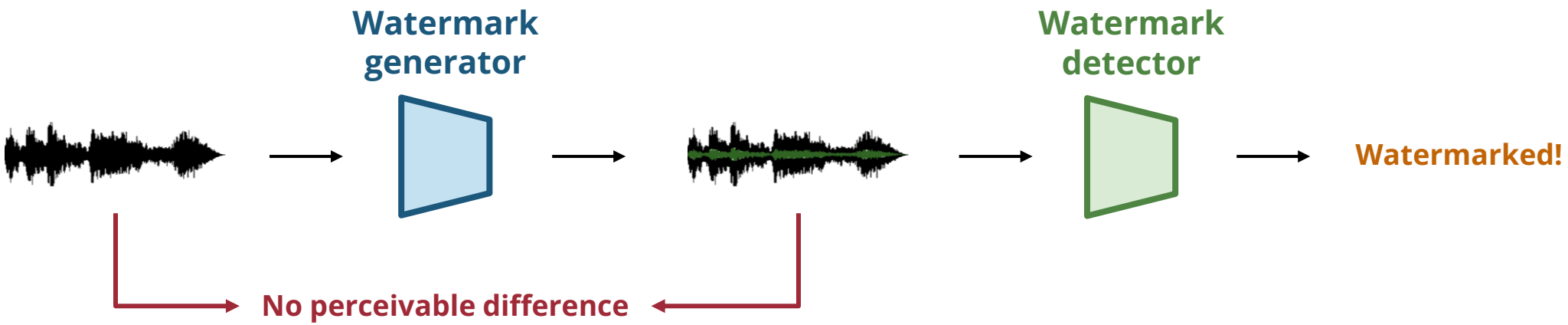
research.google/blog/the-machine-learning-behind-hum-to-search/

Resources on Version Identification

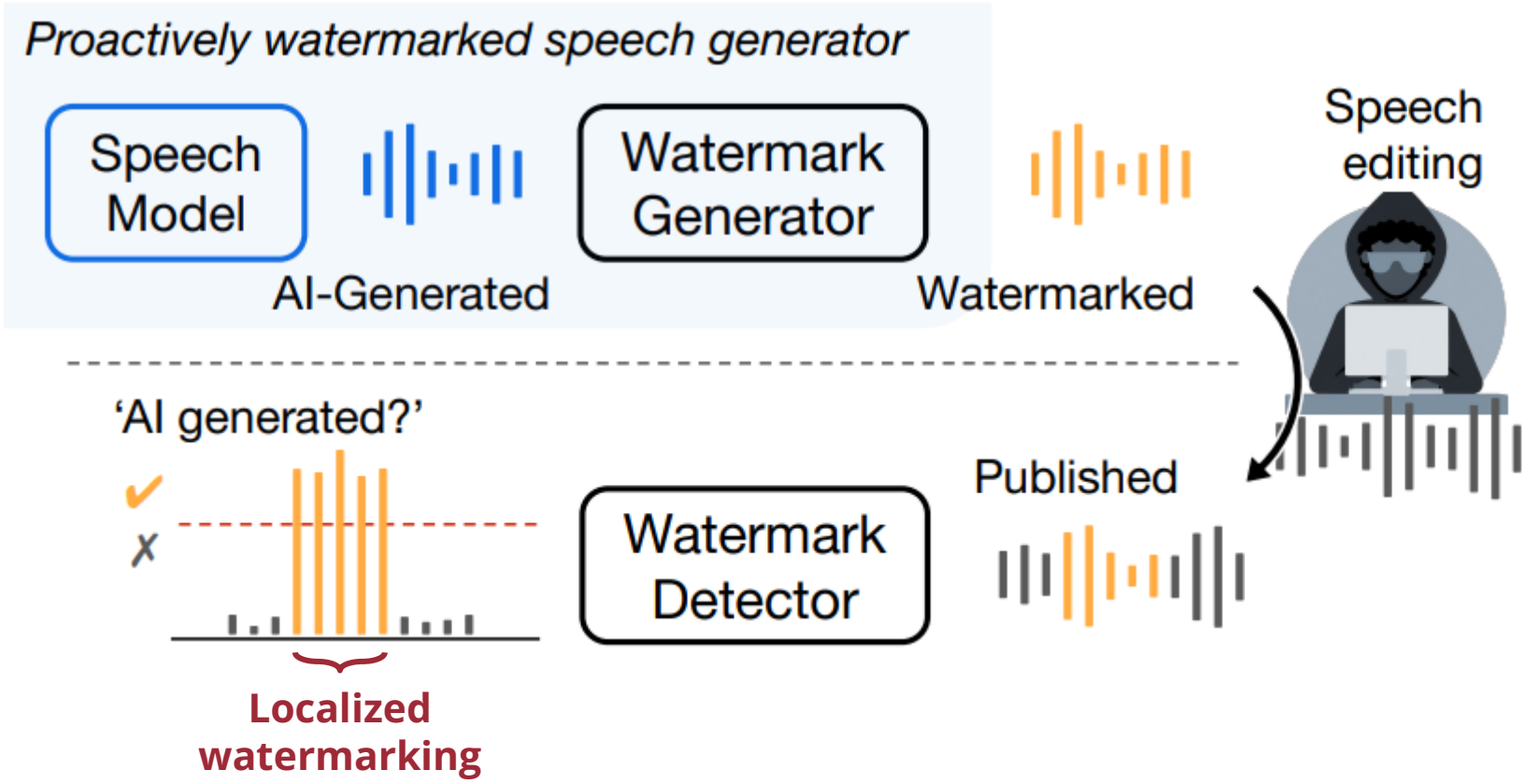
- Furkan Yesiler, Christopher Tralie, and Joan Serrà, "Version Identification in the 20s," *ISMIR Tutorials*, 2020.

Music Watermarking

Audio Watermarking

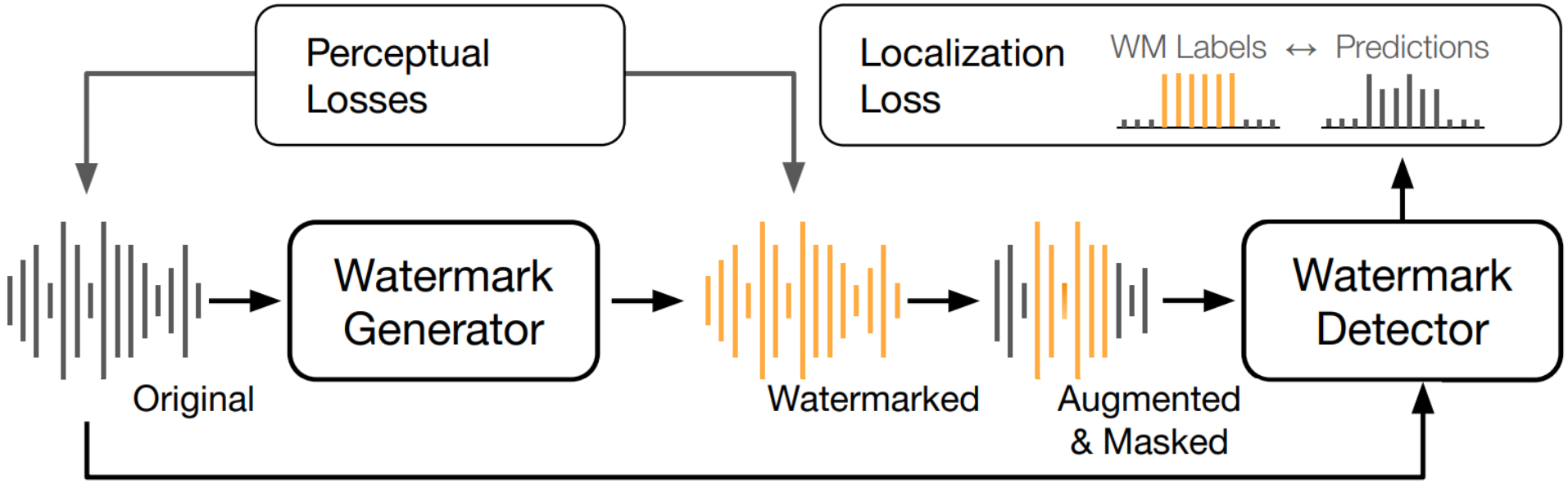


Audio Watermarking Against Generated Audio



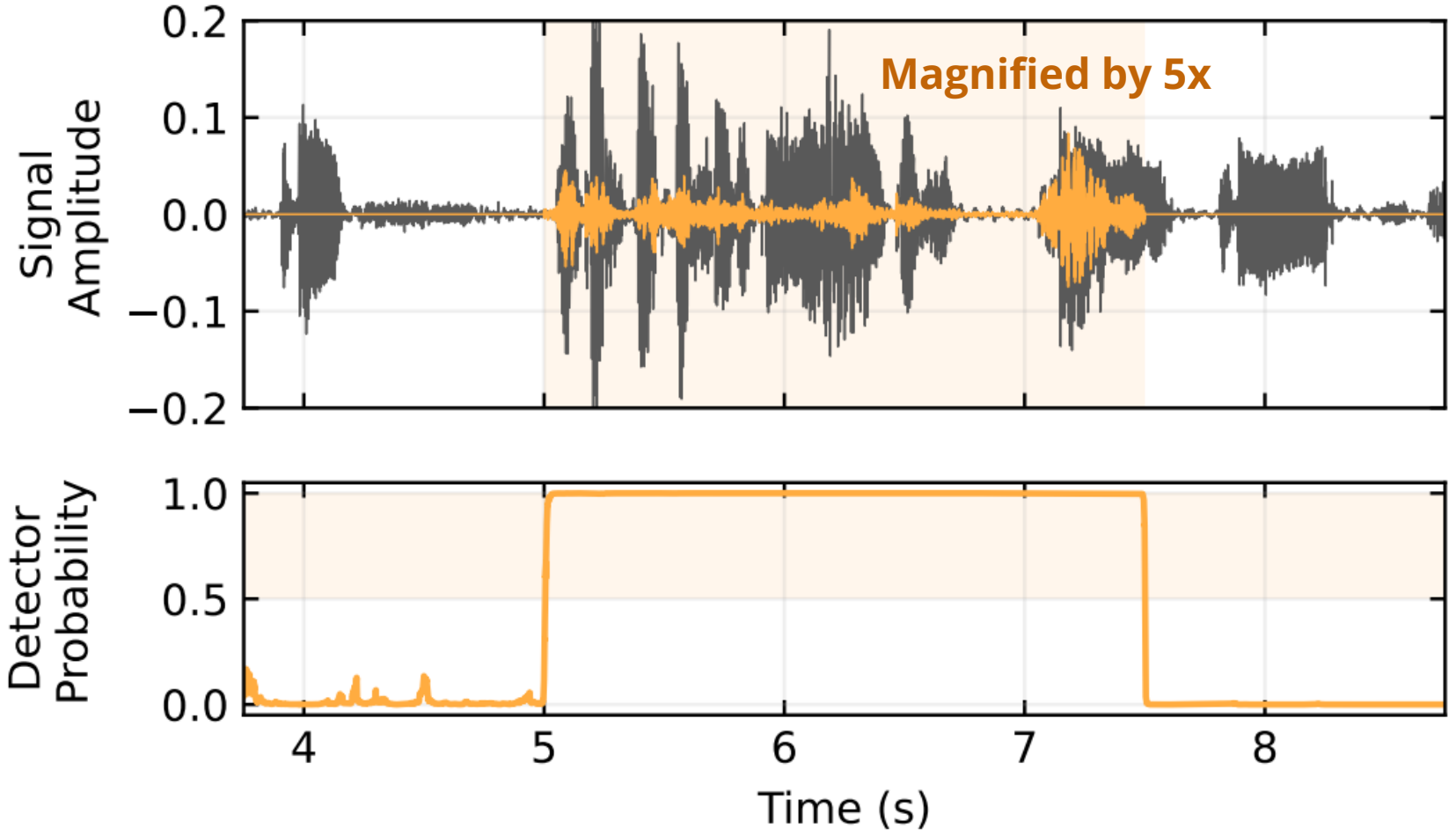
(Source: Roman et al., 2024)

AudioSeal (Roman et al., 2024)



(Source: Roman et al., 2024)

AudioSeal (Roman et al., 2024)

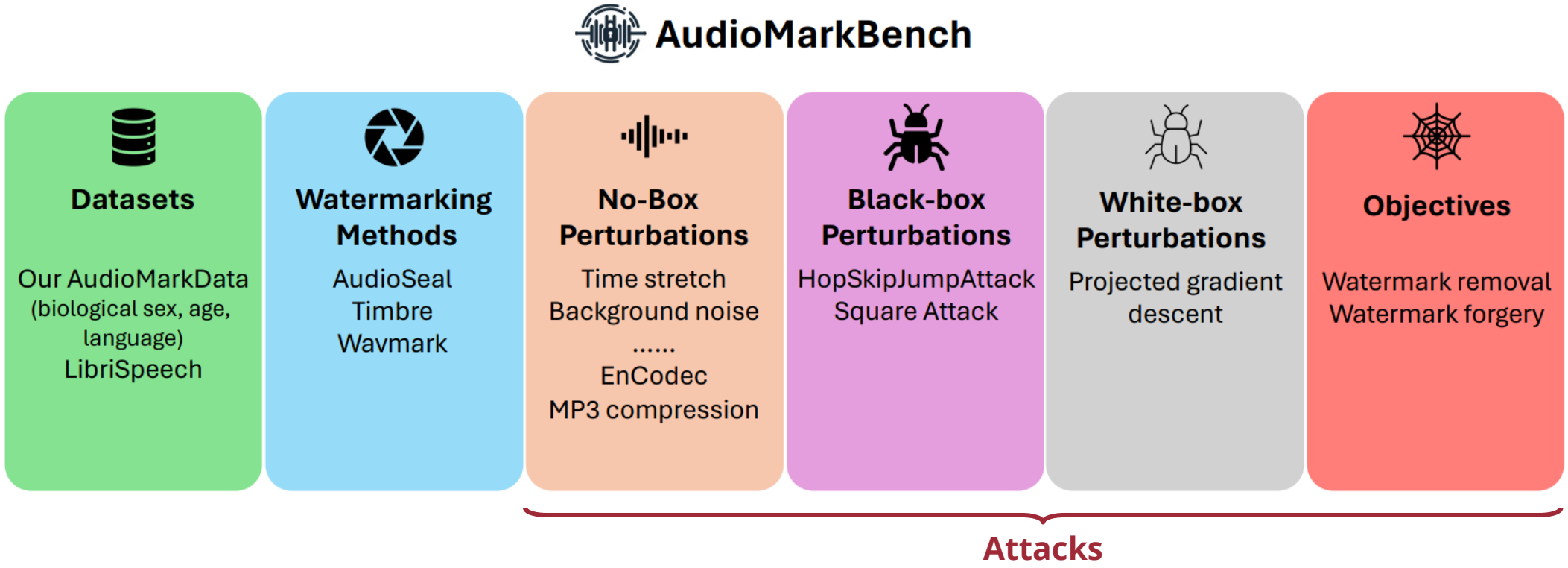


Original **Watermarked**

(Source: Roman et al., 2024)

Robin San Roman, Pierre Fernandez, Hady Elsahar, Alexandre Défossez, Teddy Furon, and Tuan Tran, "Proactive Detection of Voice Cloning with Localized Watermarking," *ICML*, 2024.

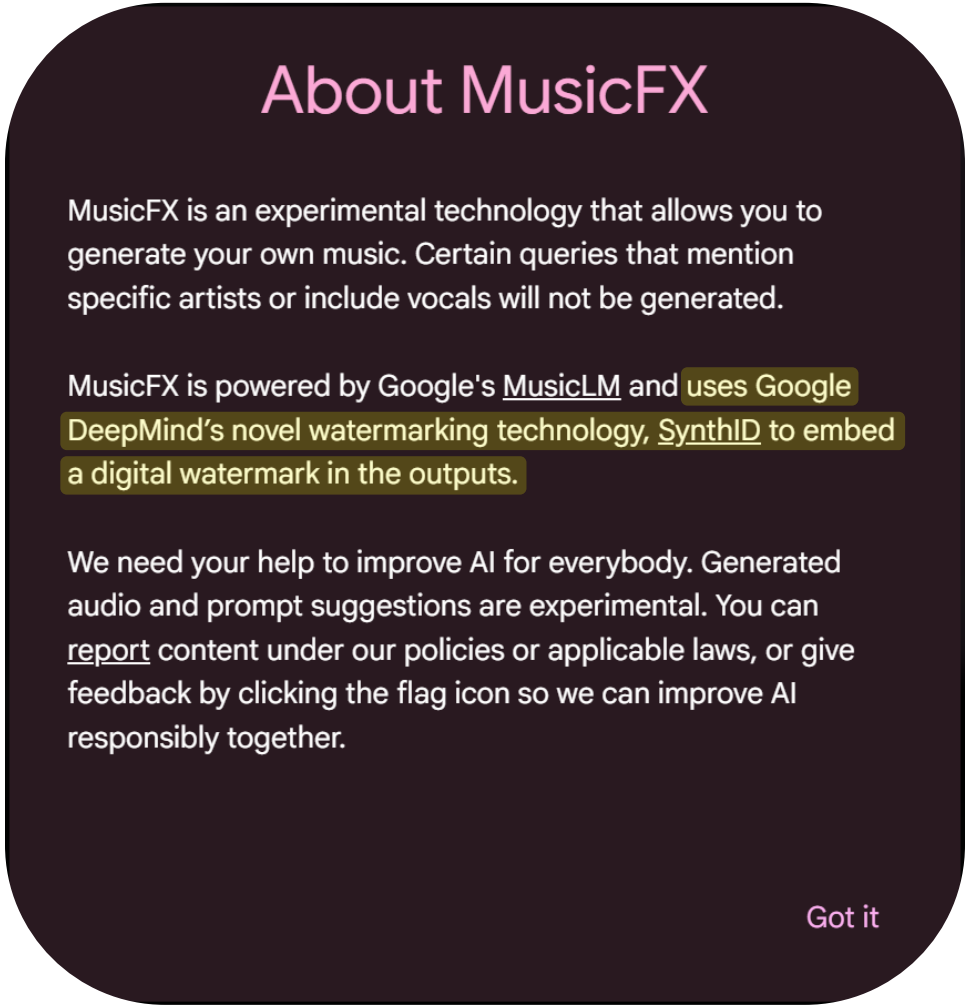
AudioMarkBench (Liu et al., 2024)



(Source: Liu et al., 2024)

Hongbin Liu, Moyang Guo, Zhengyuan Jiang, Lun Wang, and Neil Zhenqiang Gong, "AudioMarkBench: Benchmarking Robustness of Audio Watermarking," *NeurIPS Datasets and Benchmarks Track*, 2024.

MusicFX's SynthID (Google)



labs.google/fx/tools/music-fx
deepmind.google/technologies/synthid/

Music Recommendation

Music Recommendation

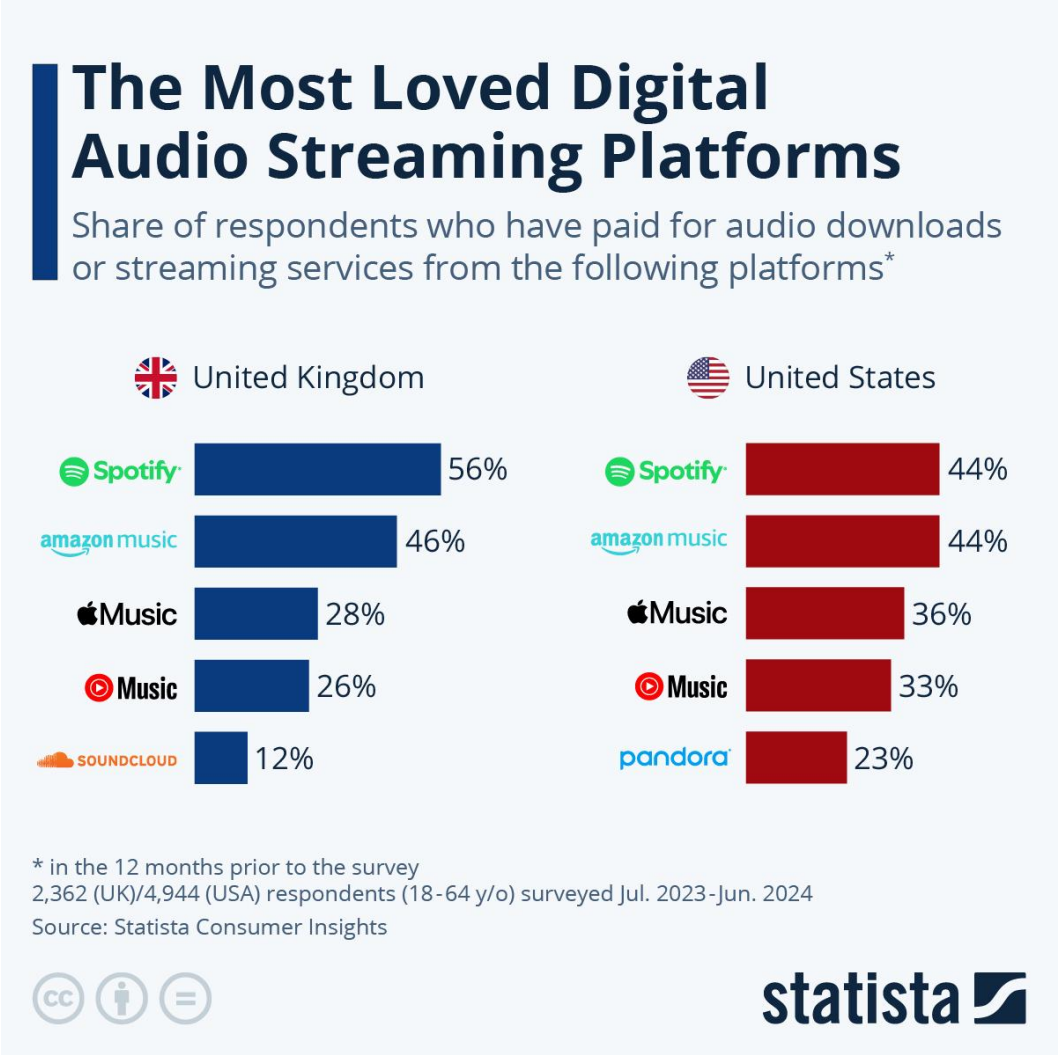
Spotify

amazon music

Music

YouTube Music

pandora

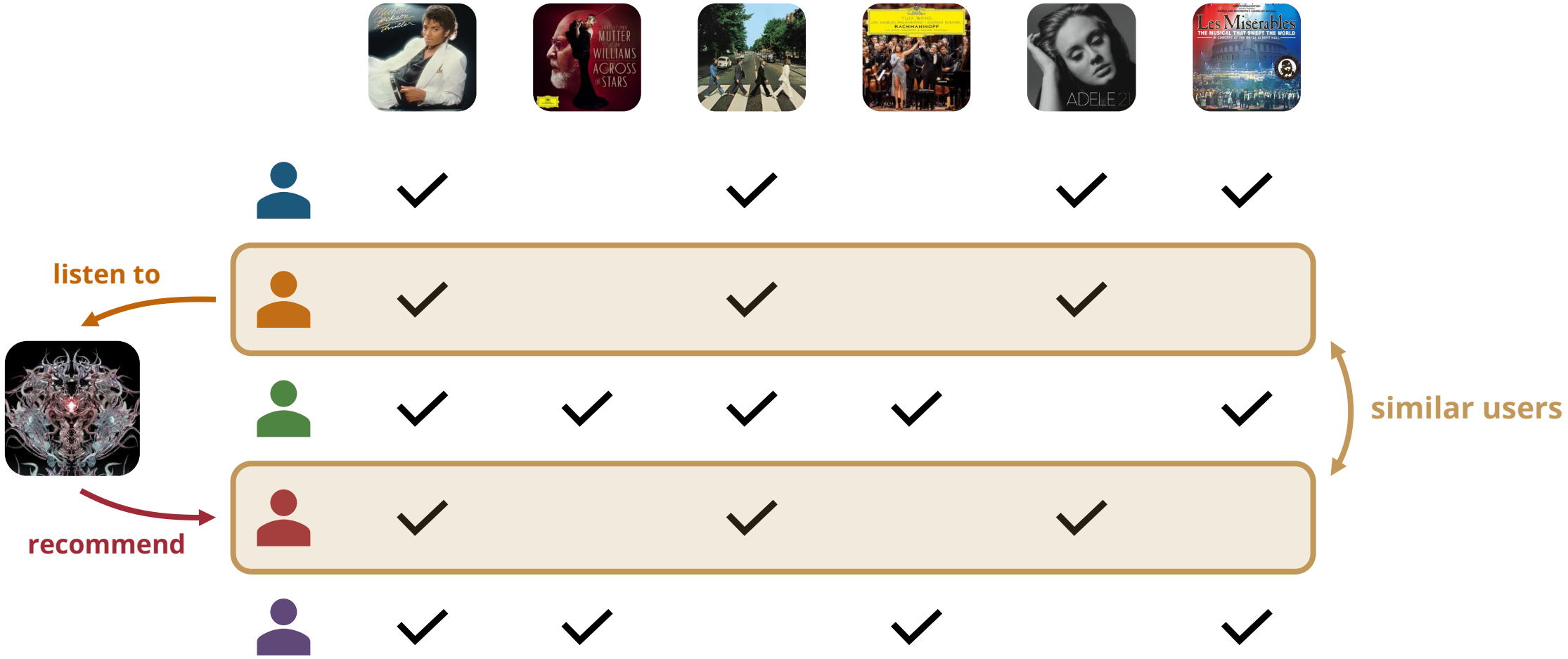


Anna Fleck, "The Most Loved Digital Audio Streaming Platforms," *Statista*, September 30, 2024.

Music Recommendation

Collaborative Filtering

User-based Collaborative Filtering



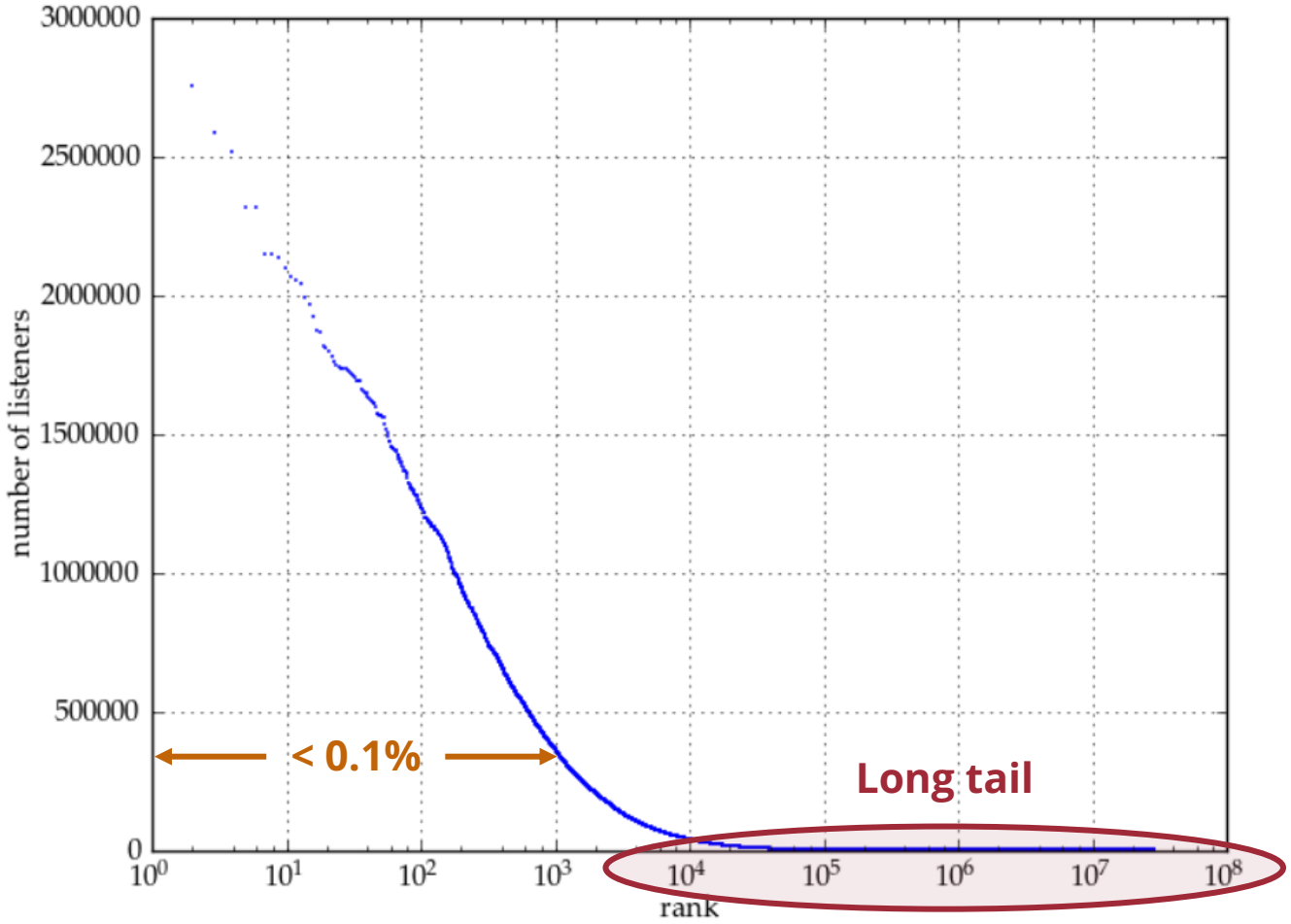
Collaborative Filtering

Item-based Collaborative Filtering

Item-based Collaborative Filtering

Challenges of Music Recommendation

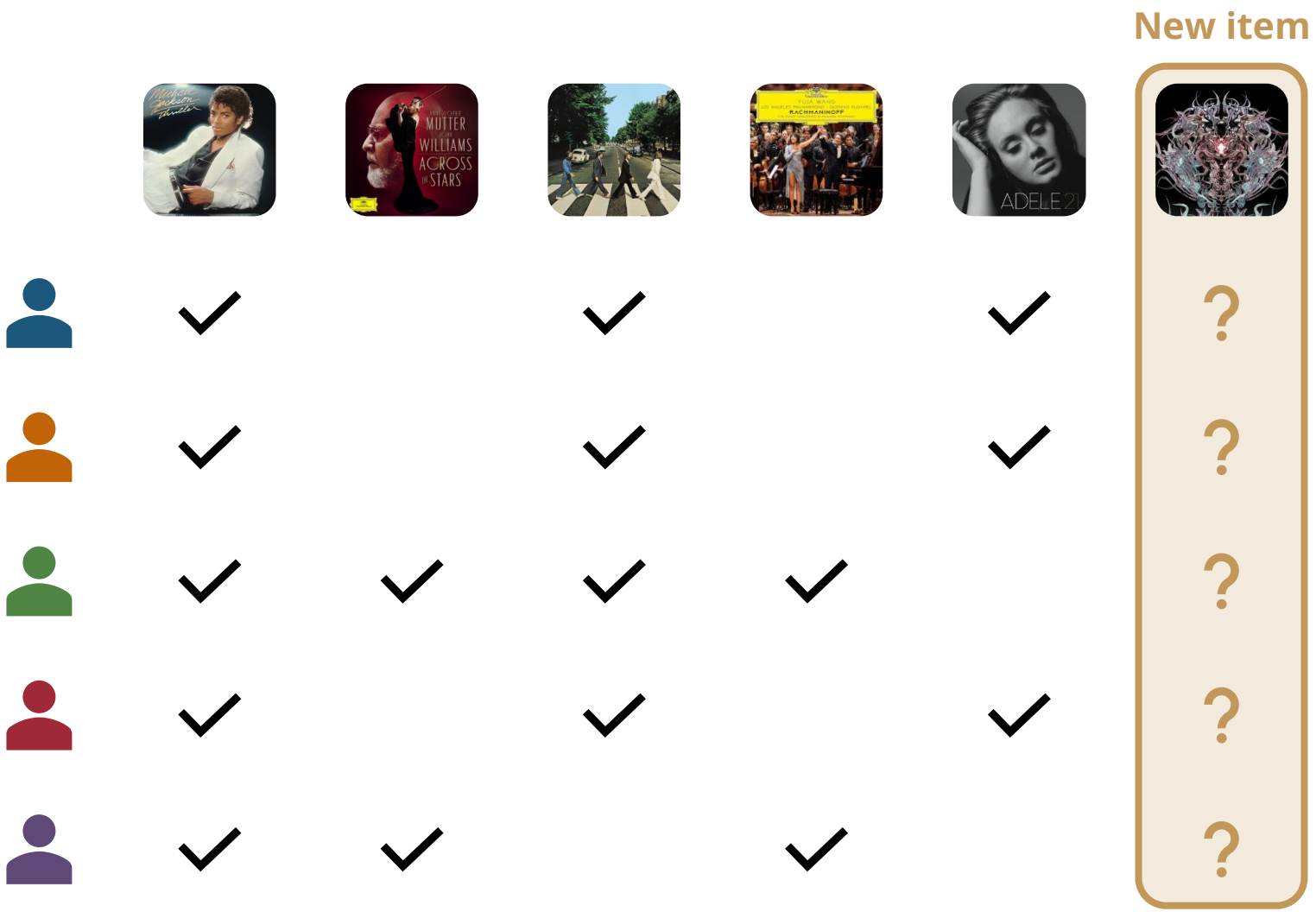
The Long Tail Problem



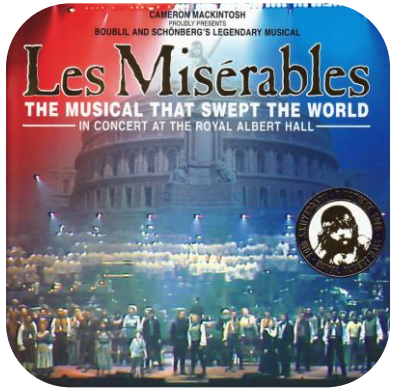
(Source: Levy & Bosteels, 2010)

Mark Levy and Klaas Bosteels, "Music Recommendation and the Long Tail," *WOMRAD*, 2010.

Cold Star Problem: New Items



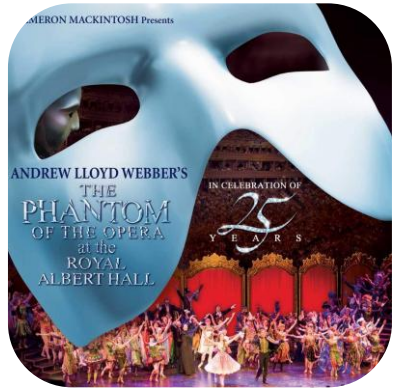
Content-based Filtering



Musical

Live concert version

Concert at Royal Albert Hall



Musical

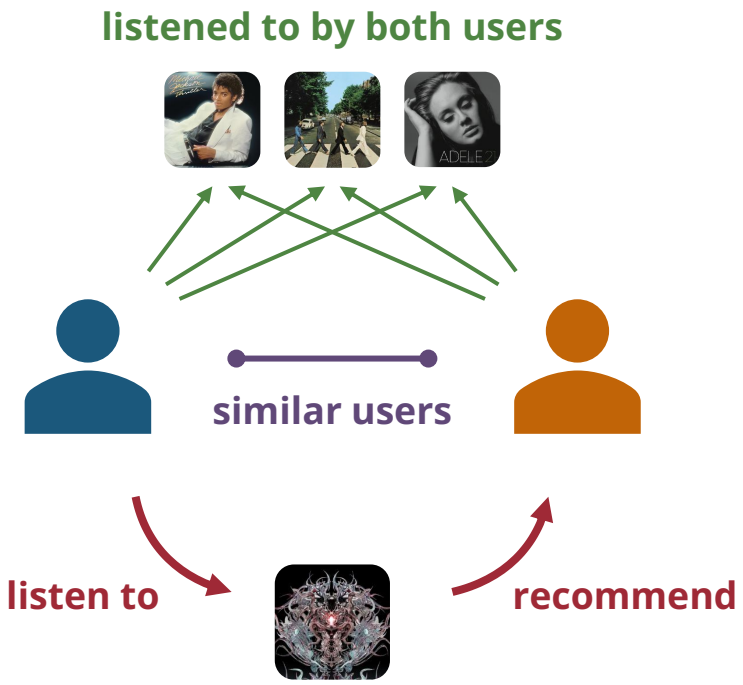
Live concert version

Concert at Royal Albert Hall

similar songs

Collaborative Filtering vs Content-based Filtering

Collaborative filtering



Content-based filtering

Cold Star Problem: New Users

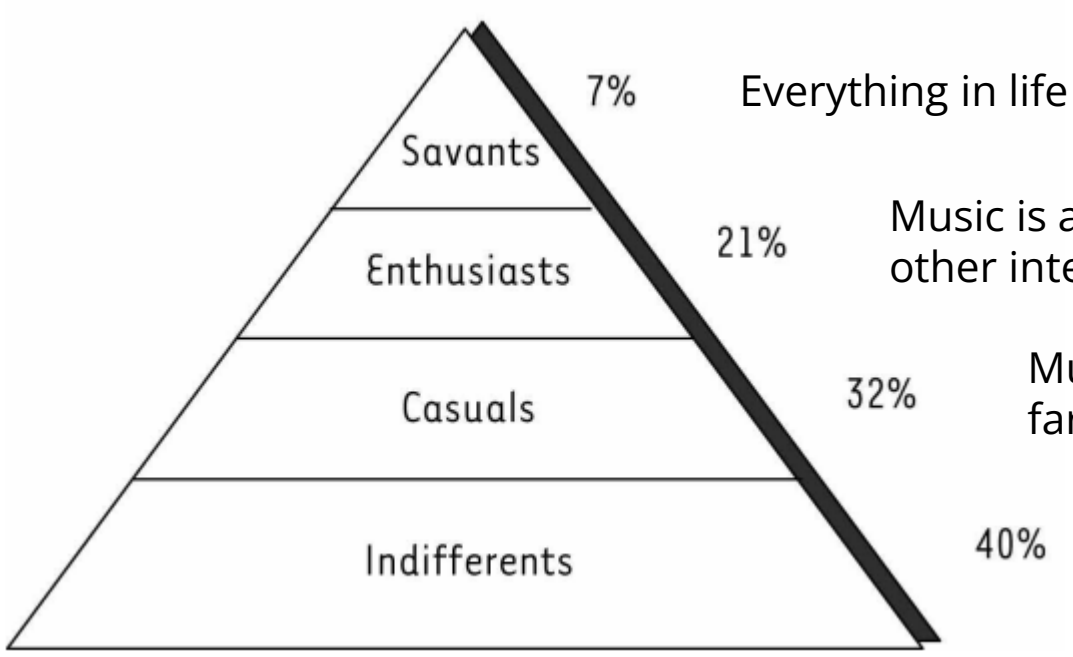
User Profile Modelling

Data type	Example
Demographic	Age, marital status, gender etc.
Geographic	Location, city, country etc.
Psychographic	*Stable*: interests, lifestyle, personality etc. *Fluid*: mood, attitude, opinions etc.

(Source: Song et al., 2012)

Yading Song, Simon Dixon, and Marcus Pearce, "A Survey of Music Recommendation Systems and Future Perspectives," *CMMR*, 2012.

User Listening Experience Modelling



Everything in life seems to be tied up with music

Music is a key part of life but is also balanced by other interests

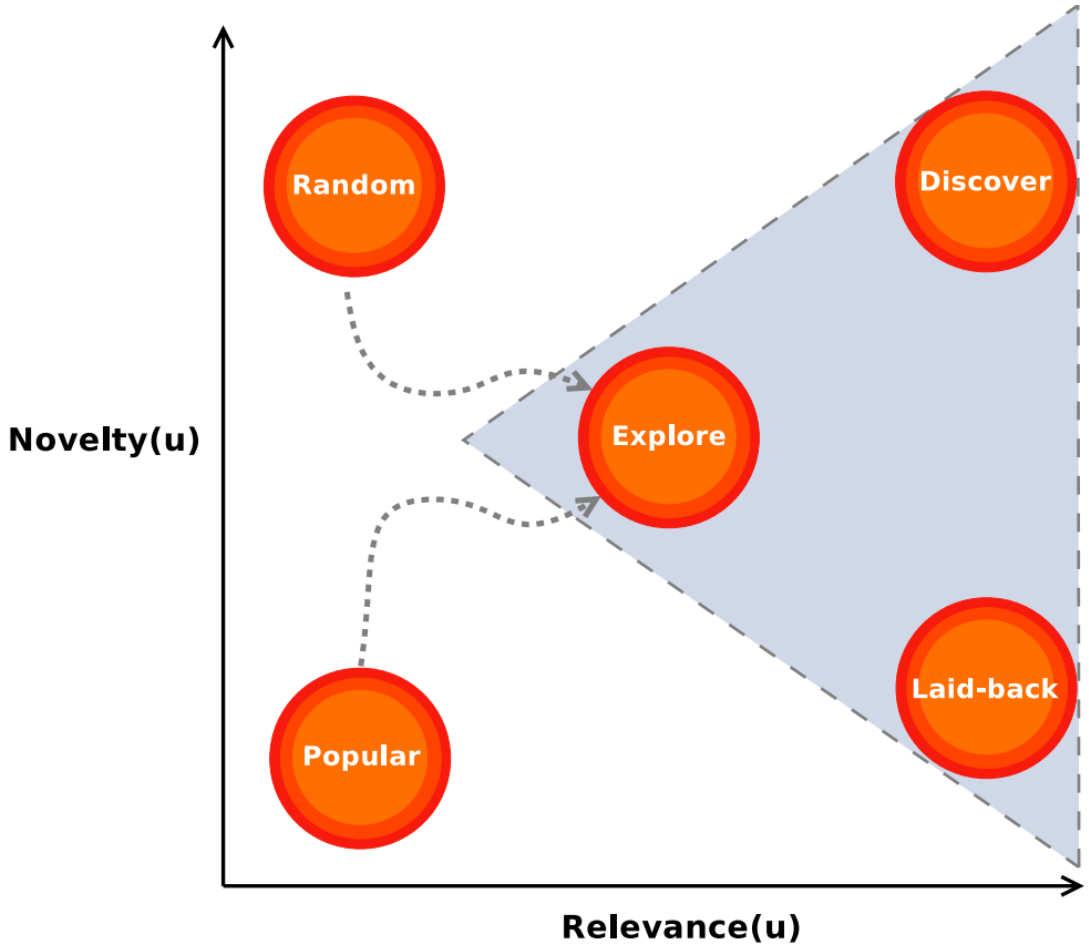
Music plays a welcome role, but other things are far more important

Would not lose much sleep if music ceased to exist

(Source: Jennings, 2007)

David Jennings, "Net, Blogs and Rock 'n' Roll: How Digital Discovery Works and What it Means for Consumers," *Nicholas Brealey Publishing*, 2007.

Novelty vs Relevance

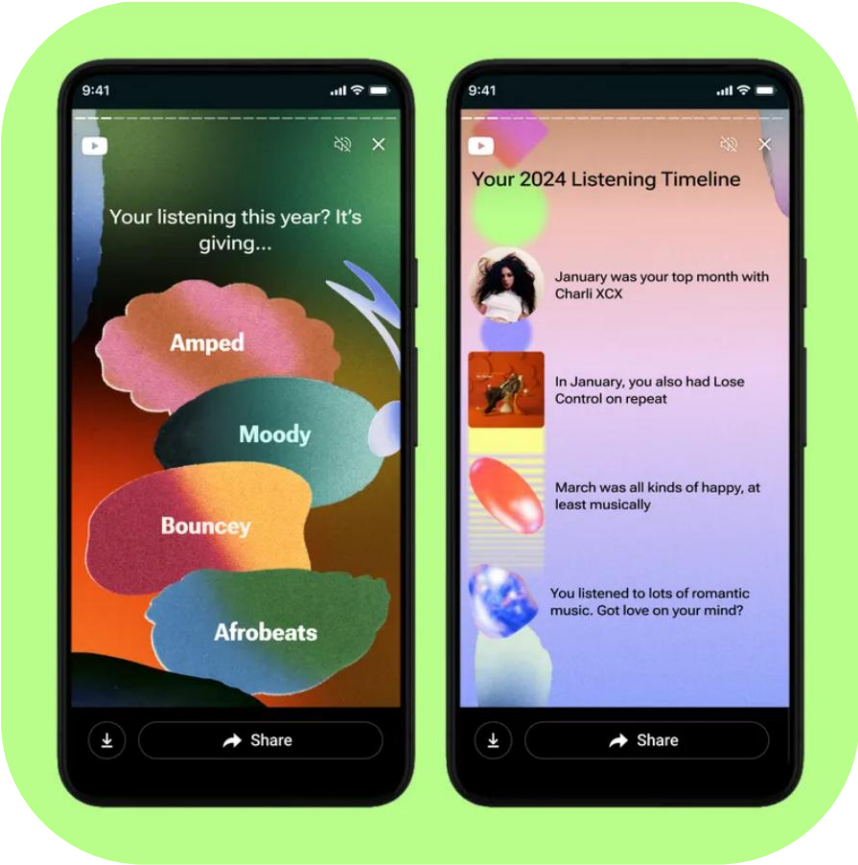


(Source: Celma, 2010)

Òscar Celma, "Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space," Springer, 2010.

Listening Behavior Analysis

YouTube's Music Recap



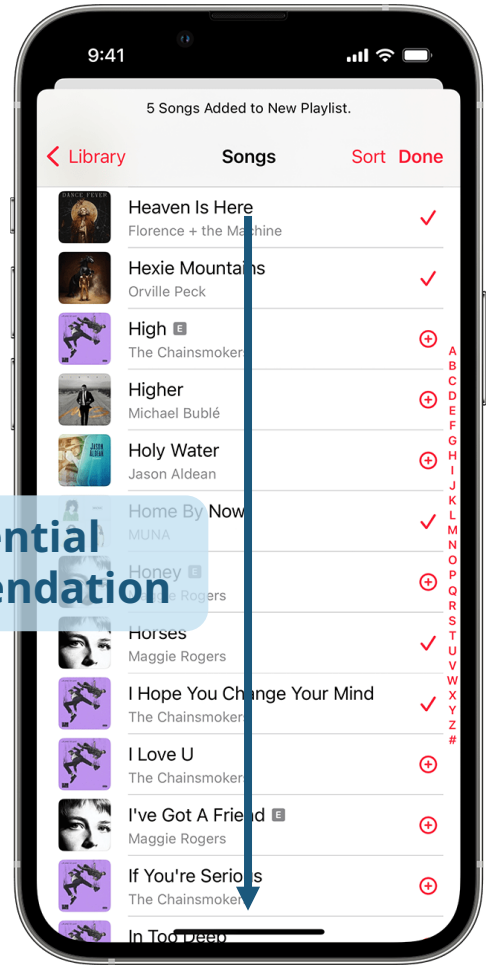
(Source: YouTube)

Spotify's Listening Personality

(Source: Spotify)

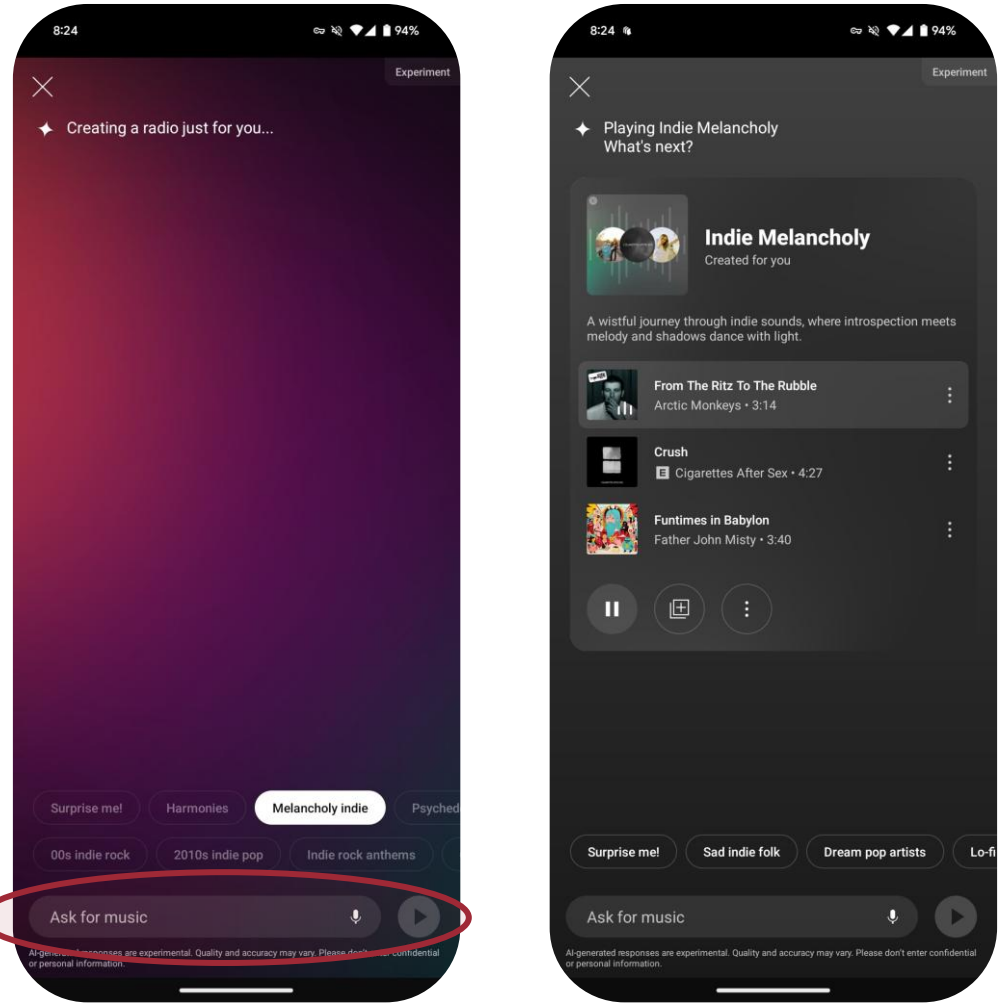
blog.youtube/news-and-events/2024-music-recap-youtube/
engineering.atspotify.com/2023/01/whats-a-listening-personality/

Music Playlist Generation



Sequential recommendation

YouTube Music's Ask Music (2024)

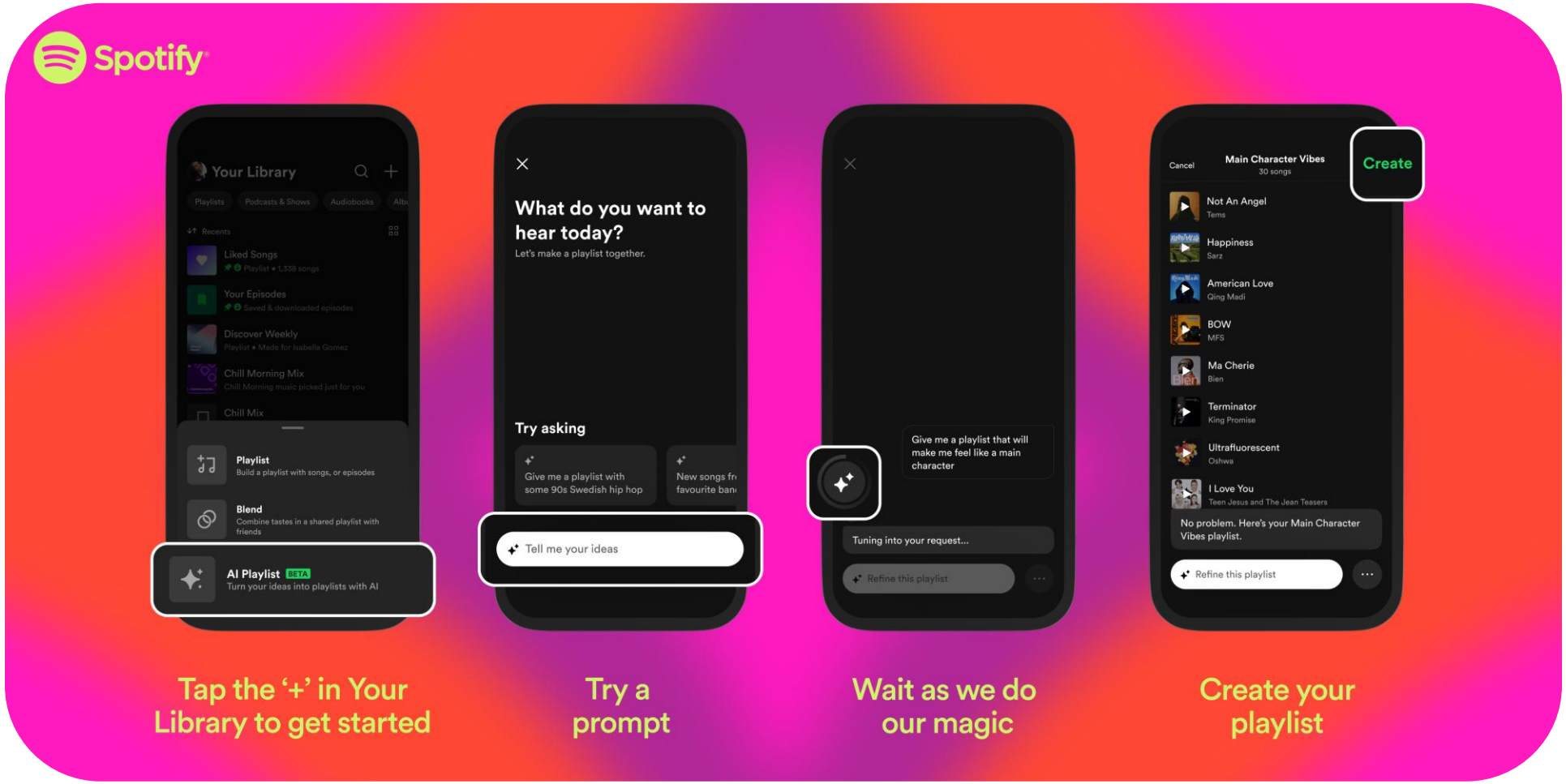


Text prompt

(Source: Android Police)

androidpolice.com/youtube-music-ask-music-ai-playlist-assistant/

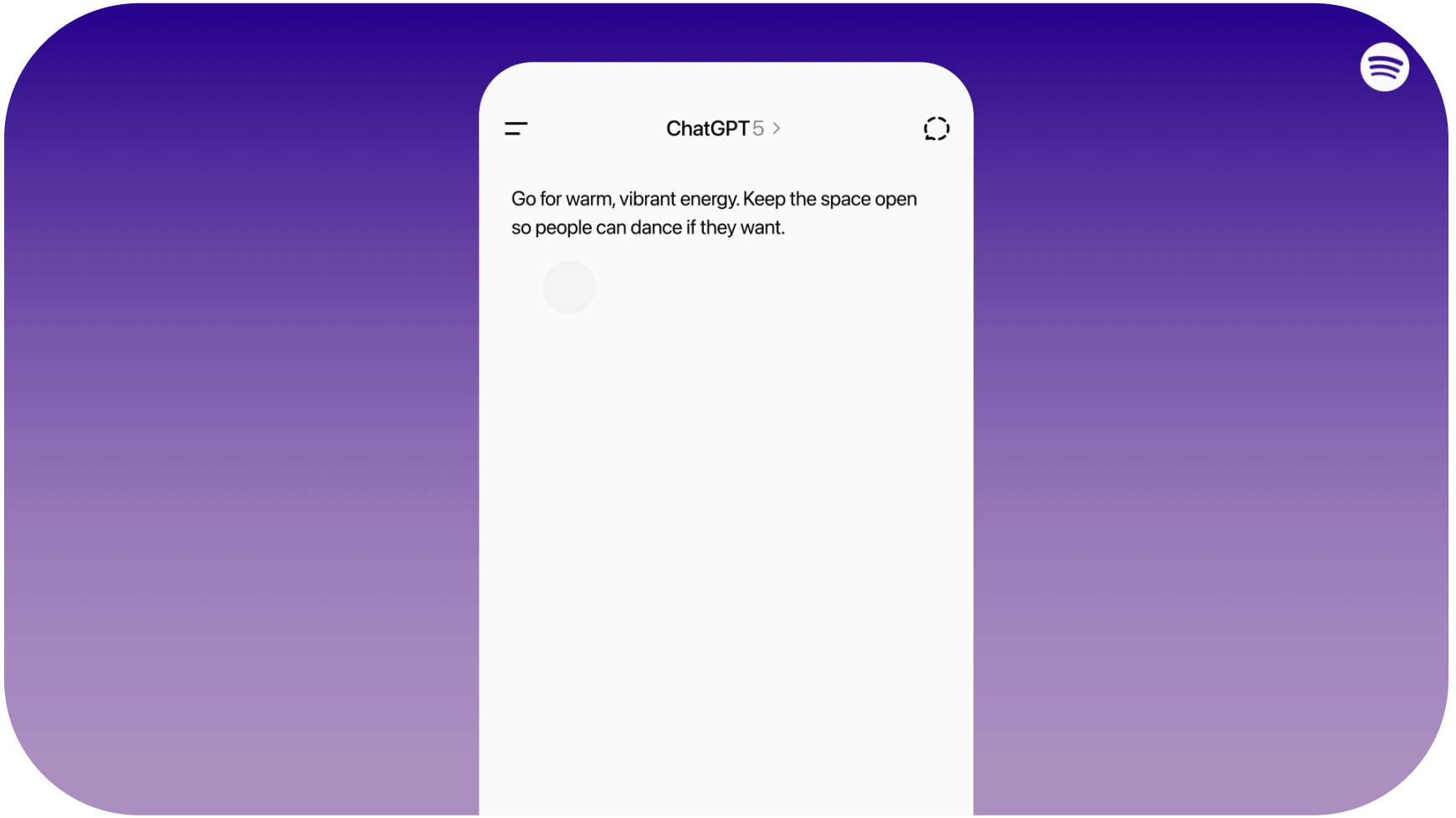
Spotify’s AI Playlist (2024)



(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/

Spotify’s Personalized Picks in ChatGPT (2025)



(Source: Spotify)

newsroom.spotify.com/2025-10-06/spotify-personalized-prompts-chatgpt

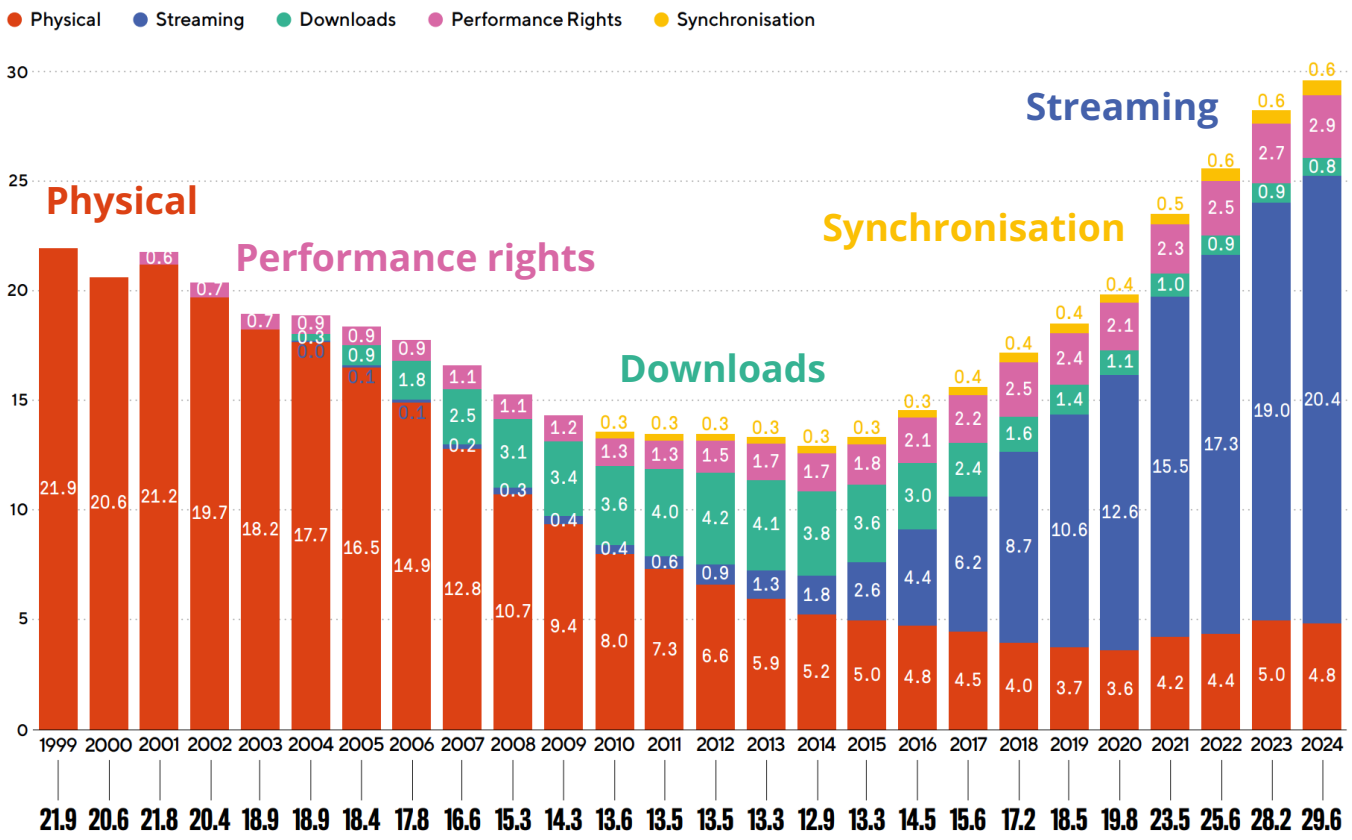
Resources on Music Recommendation

- Markus Schedl, Peter Knees, and Fabien Gouyon, "Overview and New Challenges of Music Recommendation Research in 2018," *ISMIR Tutorials*, 2018.
- Markus Schedl, Hamed Zamani, Ching-Wei Chen, Yashar Deldjoo, and Mehdi Elahi, "Current Challenges and Visions in Music Recommender Systems Research," *International Journal of Multimedia Information Retrieval*, 7:95–116, 2018.

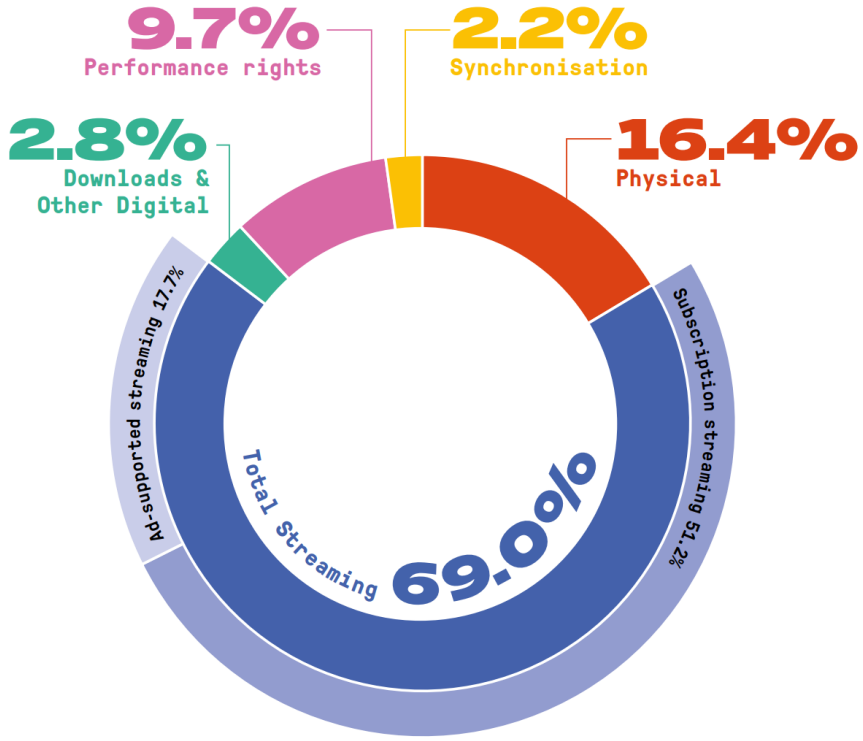
Recap

Global Recorded Music Industry Revenues

Global recorded music industry revenues 1999-2024 (US$ billions)

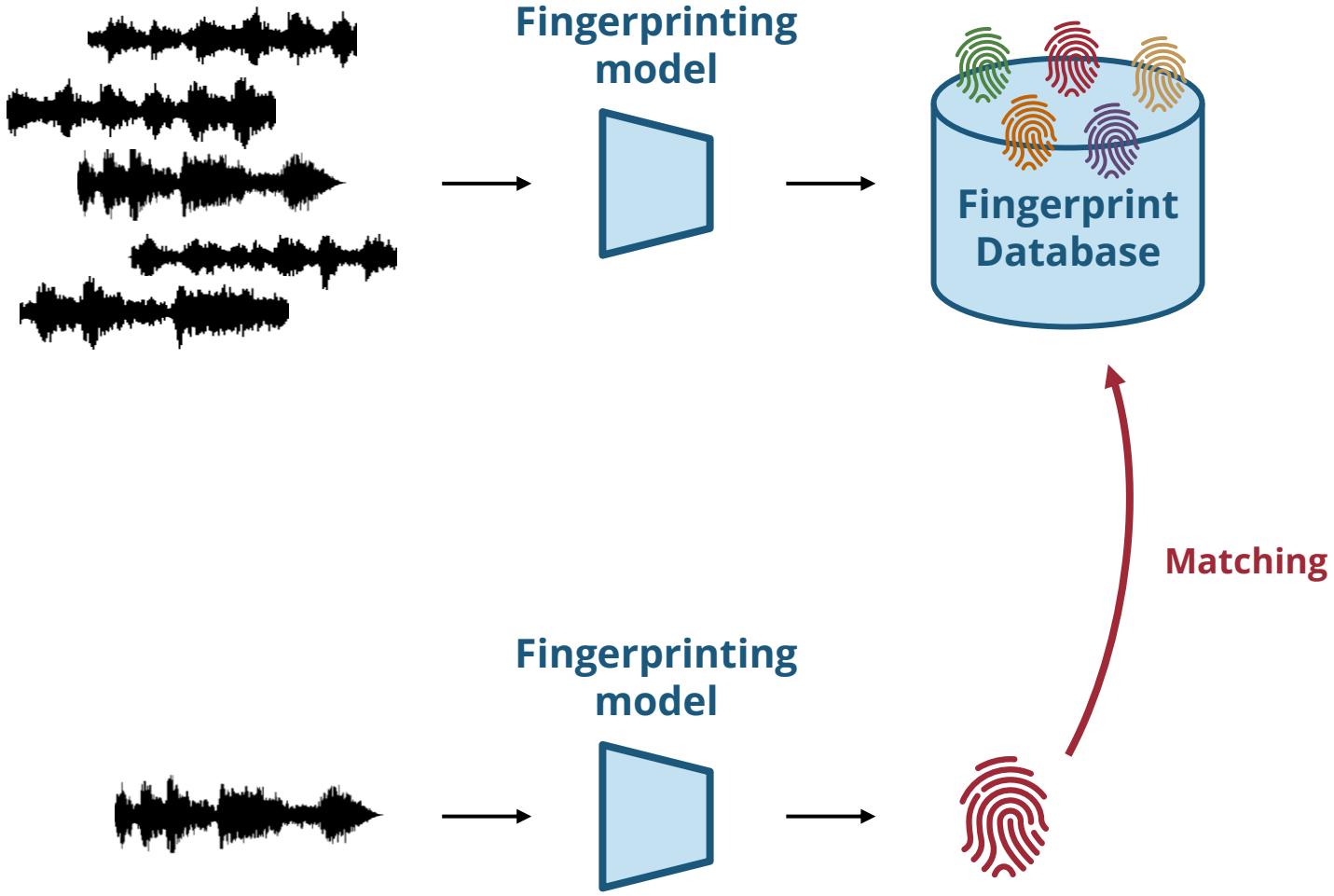


Global recorded music revenue share by format - 2024



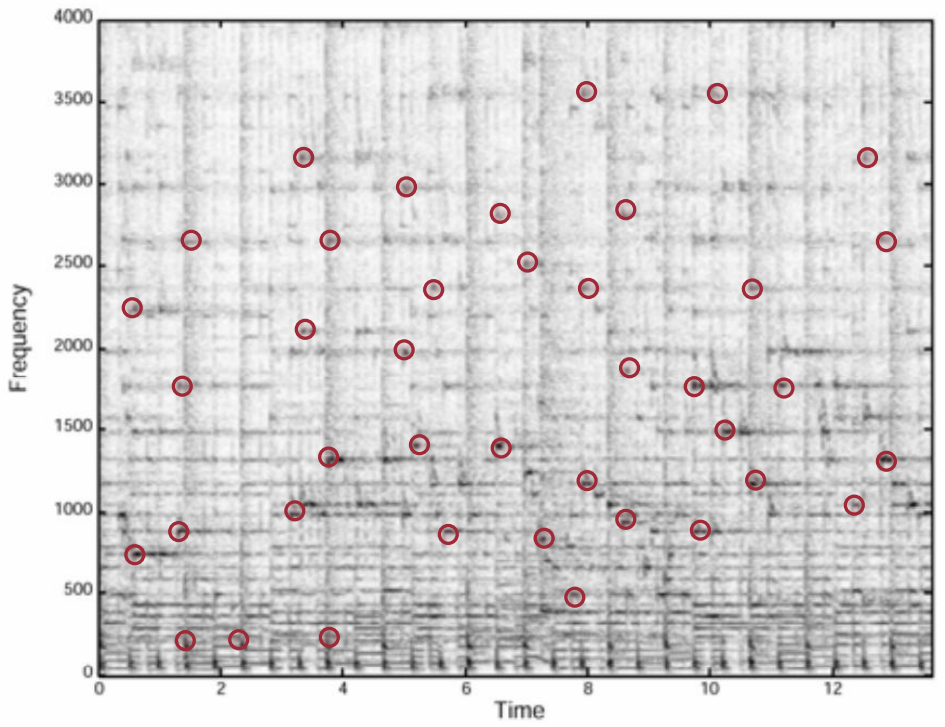
Global Music Report 2025, IFPI, 2025

Audio Fingerprinting for Audio Identification

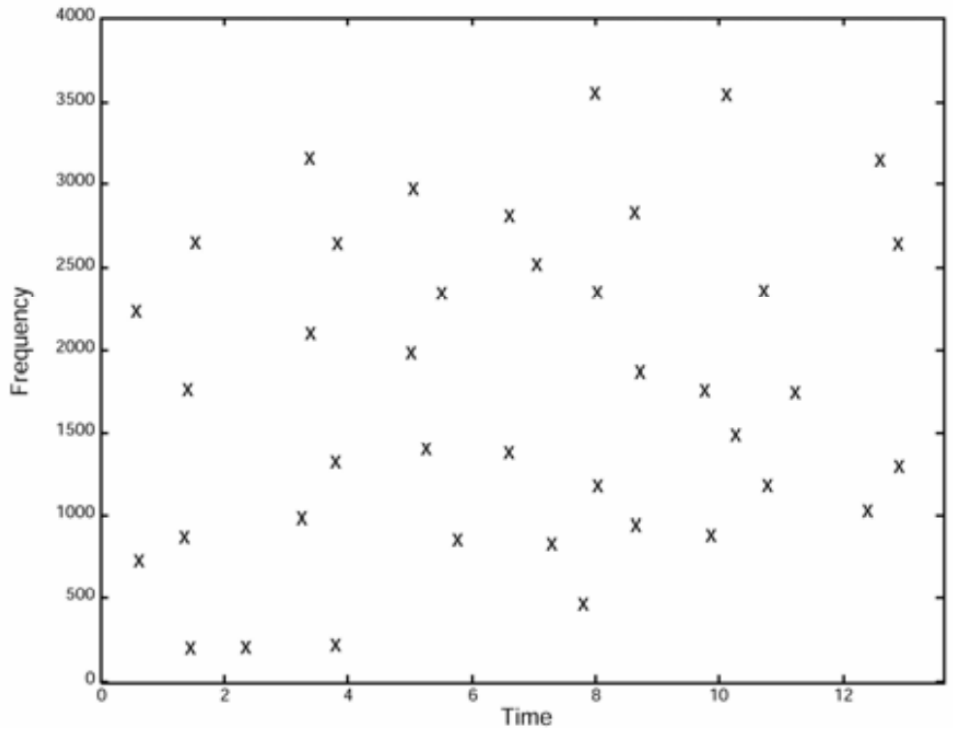


Peak-based Audio Fingerprinting (Wang et al., 2003)

Spectrogram



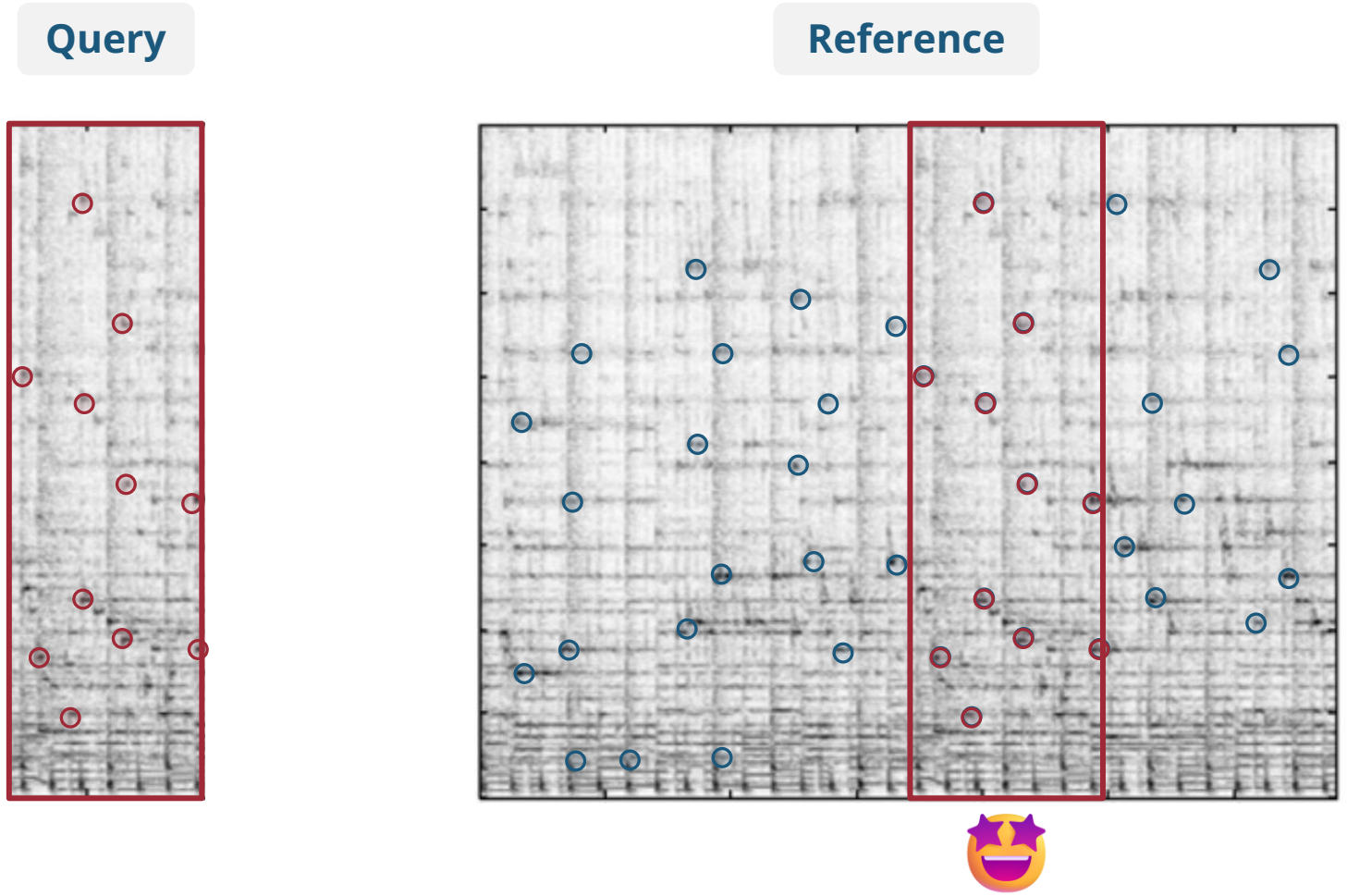
Constellation Map



(Source: Wang et al., 2003)

Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

Peak-based Audio Fingerprinting (Wang et al., 2003)



(Source: Wang et al., 2003)

Avery Li-Chun Wang, "An Industrial-Strength Audio Search Algorithm," *ISMIR*, 2003.

Version Identification: Beyond Fingerprinting & Covers

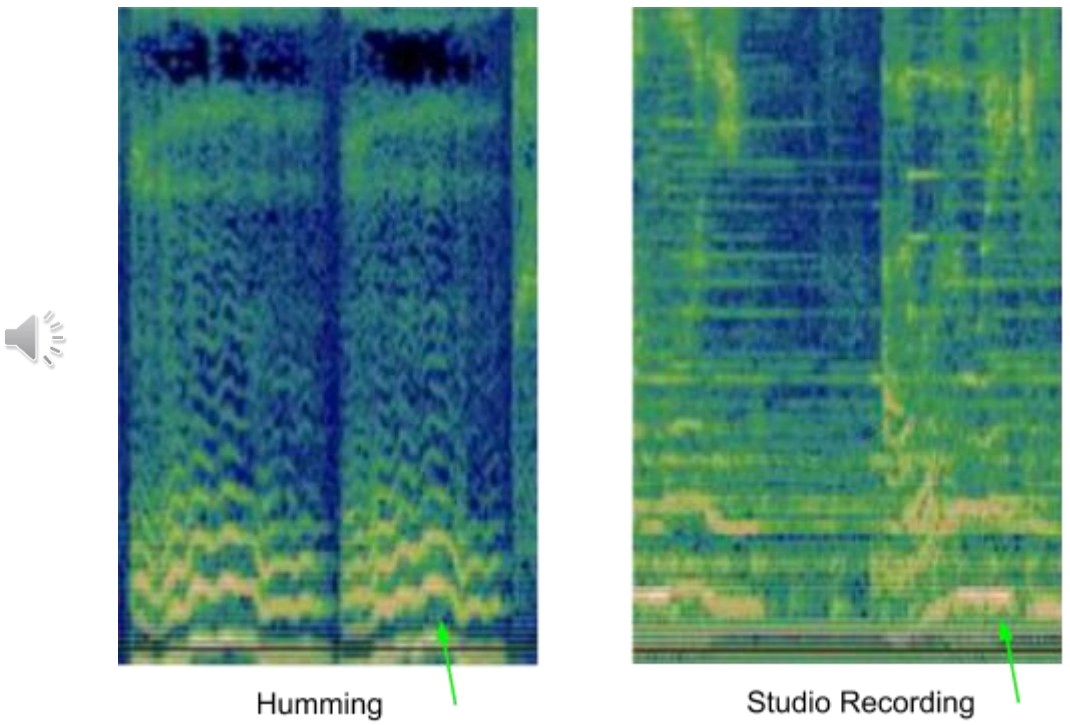
	Version Type																	
Musical Characteristic	Duplicate	Remaster	Radio Edit	Translation	Performance	Demo	Parody	Within-Genre	Karaoke	Live	Standard	Mashup	Acoustic	Medley	Remix	Cross-Genre	Arrangement	Quotation
Melody	0	0	0	0	0	1	0	1	2	1	1	0	1	1	1	2	2	2
Harmony	0	0	0	0	0	1	0	1	0	0	2	0	1	1	2	2	2	3
Tempo	0	0	0	0	2	1	1	1	0	2	1	3	2	2	3	2	2	3
Timing	0	0	0	0	2	1	1	1	0	2	1	3	2	2	2	3	3	3
Structure	0	0	1	0	1	1	1	1	1	2	2	3	2	3	3	2	3	3
Lyrics	0	0	1	3	0	1	3	0	3	1	0	0	0	1	1	1	1	2
Key	0	0	0	0	1	1	1	1	1	1	1	2	2	2	2	3	3	3
Timbre	0	0	0	0	1	1	1	1	2	2	3	2	3	2	3	3	3	3
Noise	0	1	1	1	3	3	2	3	2	3	3	2	3	3	2	3	3	3

Degree of Potential Difference	0	1	2	3
	Likely the same	May be variations	May be major differences	May be unrelated

(Source: Yesiler et al., 2021)

Furkan Yesiler, Guillaume Doras, Rachel M. Bittner, Christopher J. Tralie, and Joan Serrà, "Audio-based Musical Version Identification: Elements and Challenges," *IEEE Signal Processing Magazine*, 2021.

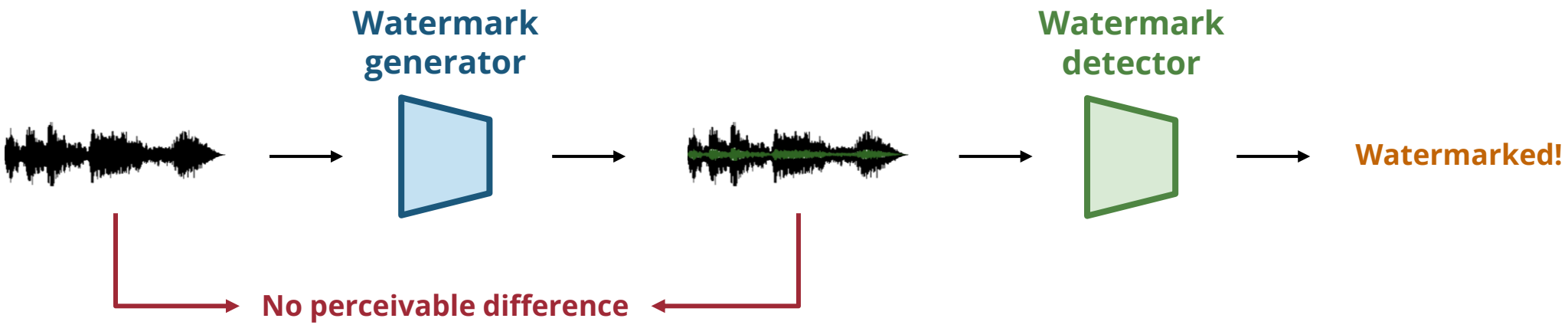
Google's Hum to Search (2020)



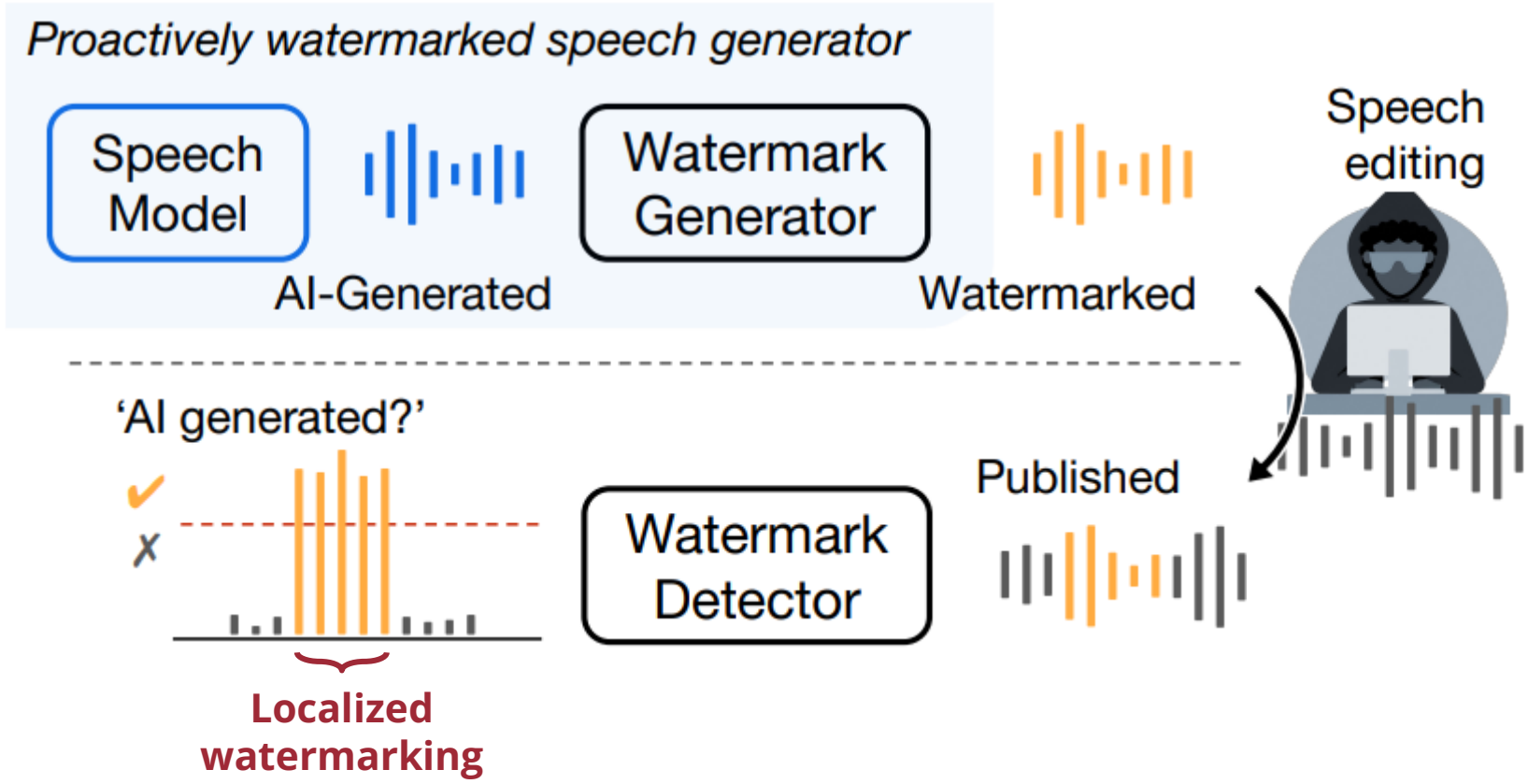
(Source: Google Research Blog)

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

Audio Watermarking



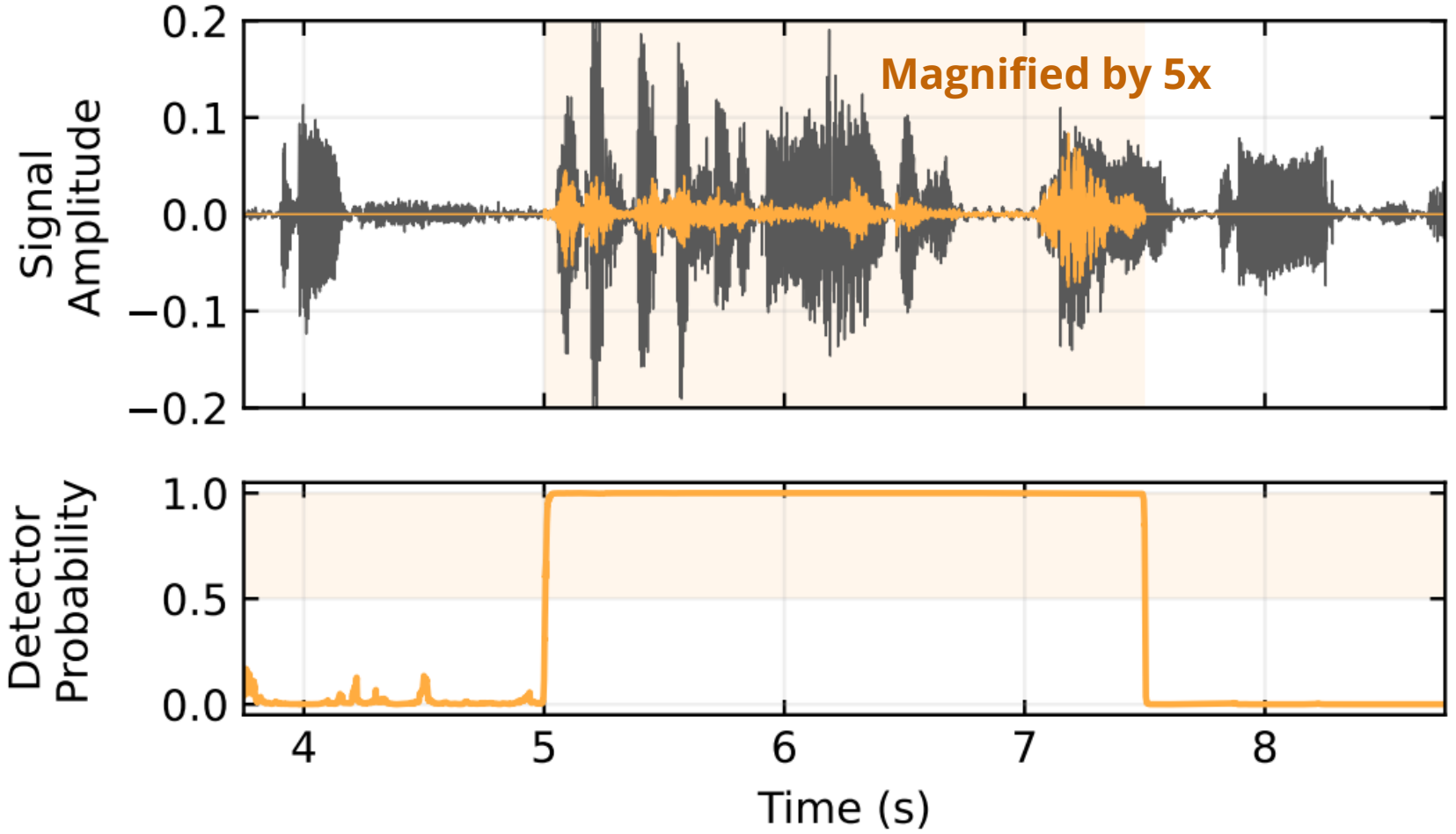
Audio Watermarking Against Generated Audio



(Source: Roman et al., 2024)

Robin San Roman, Pierre Fernandez, Hady Elsahar, Alexandre Défossez, Teddy Furon, and Tuan Tran, "Proactive Detection of Voice Cloning with Localized Watermarking," *ICML*, 2024.

AudioSeal (Roman et al., 2024)



Original **Watermarked**

(Source: Roman et al., 2024)

Robin San Roman, Pierre Fernandez, Hady Elsahar, Alexandre Défossez, Teddy Furon, and Tuan Tran, "Proactive Detection of Voice Cloning with Localized Watermarking," *ICML*, 2024.

Music Recommendation

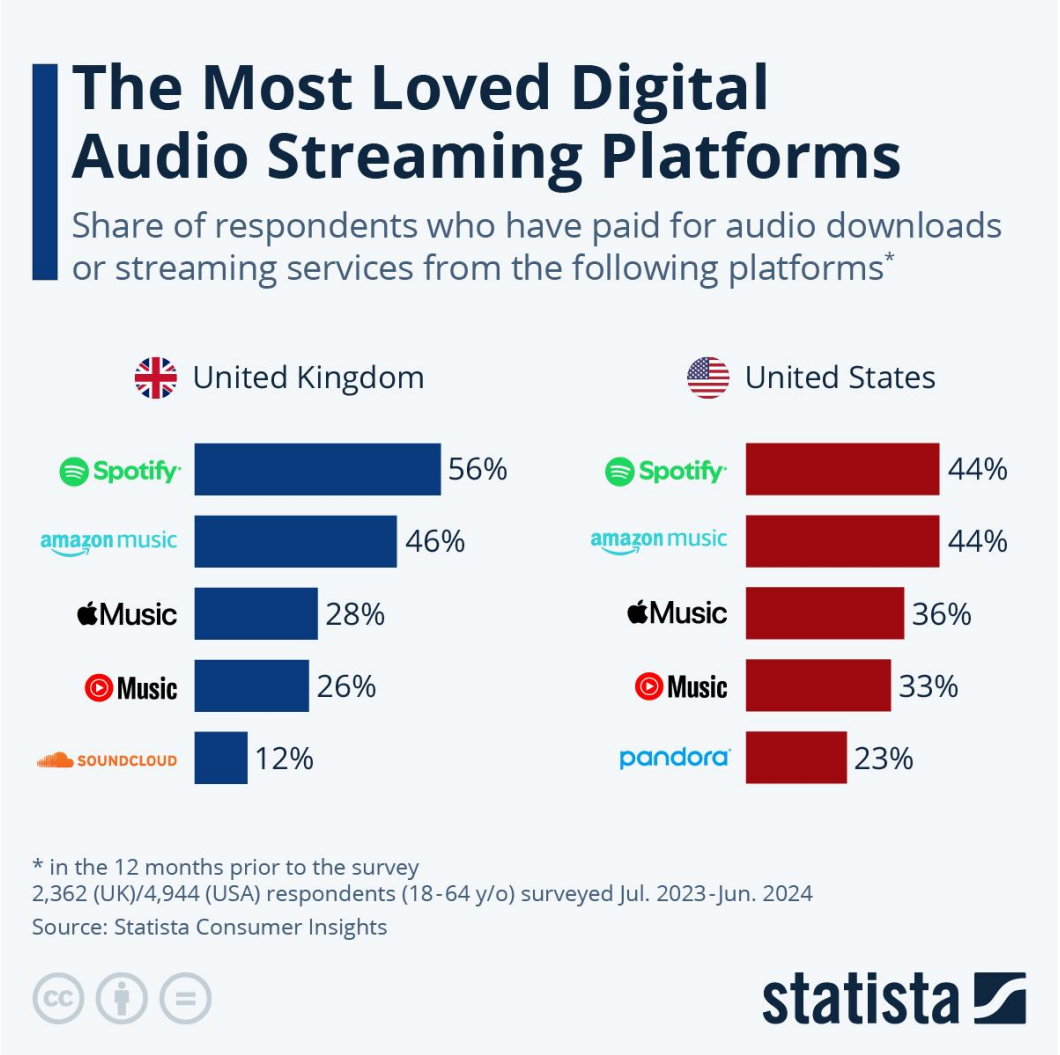
Spotify

amazon music

Music

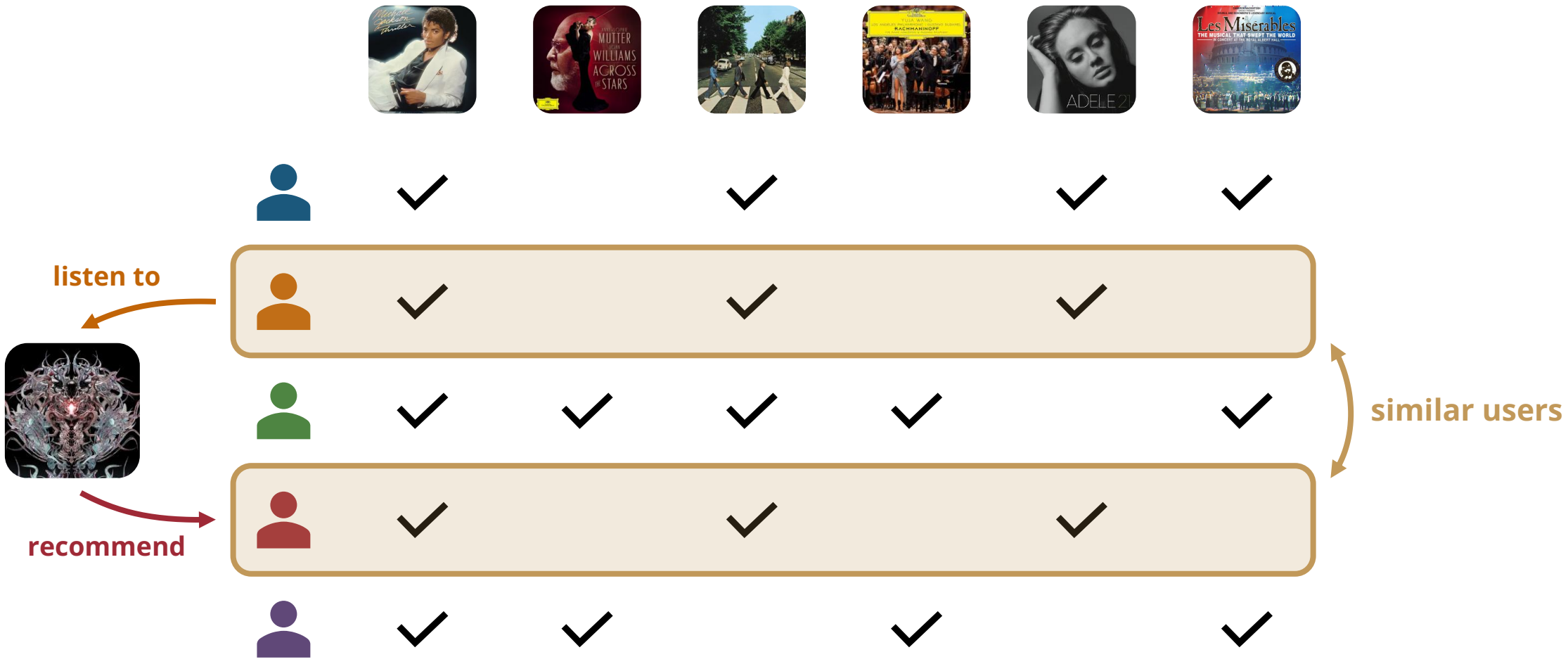
YouTube Music

pandora



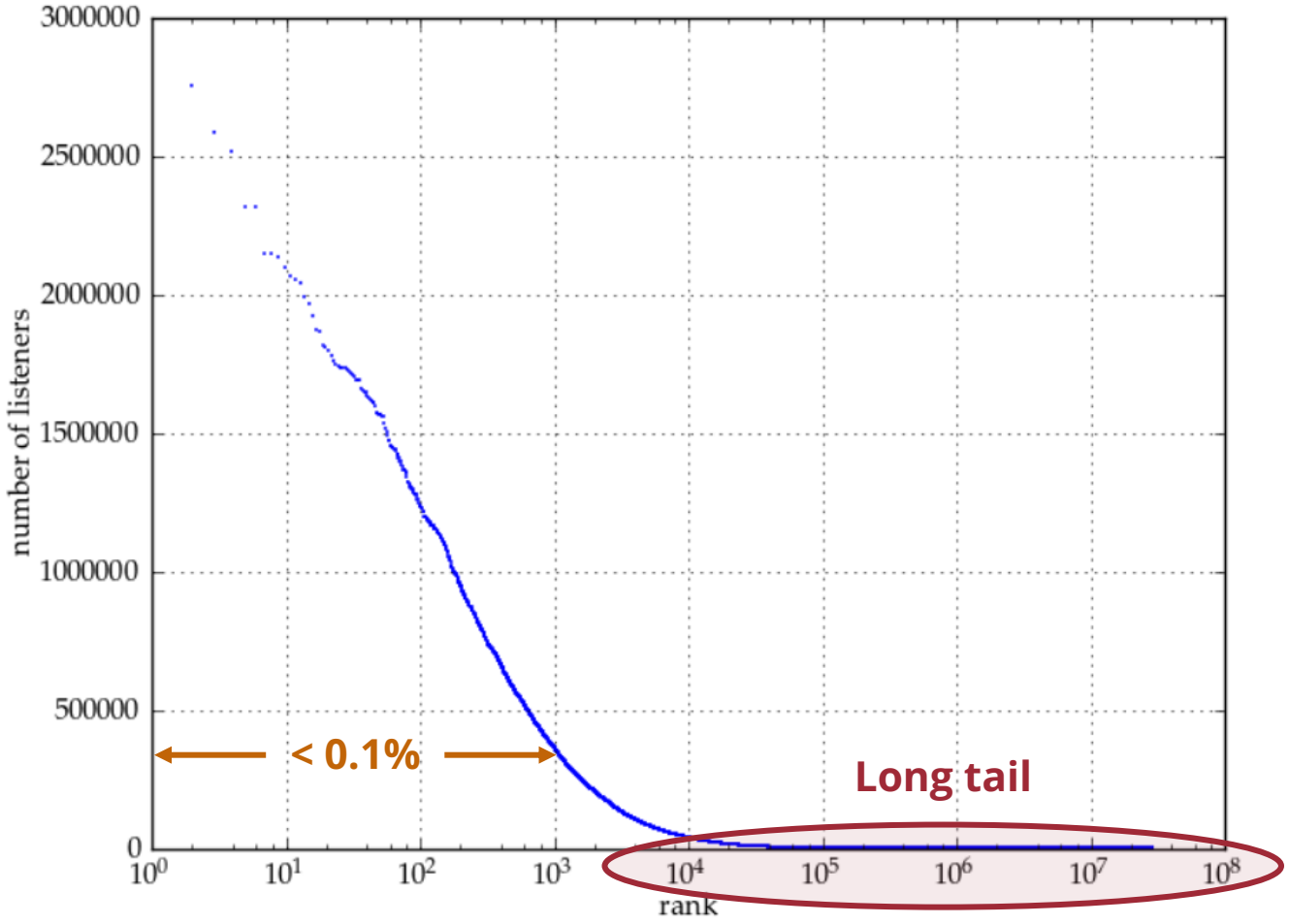
Anna Fleck, "The Most Loved Digital Audio Streaming Platforms," *Statista*, September 30, 2024.

User-based Collaborative Filtering



Item-based Collaborative Filtering

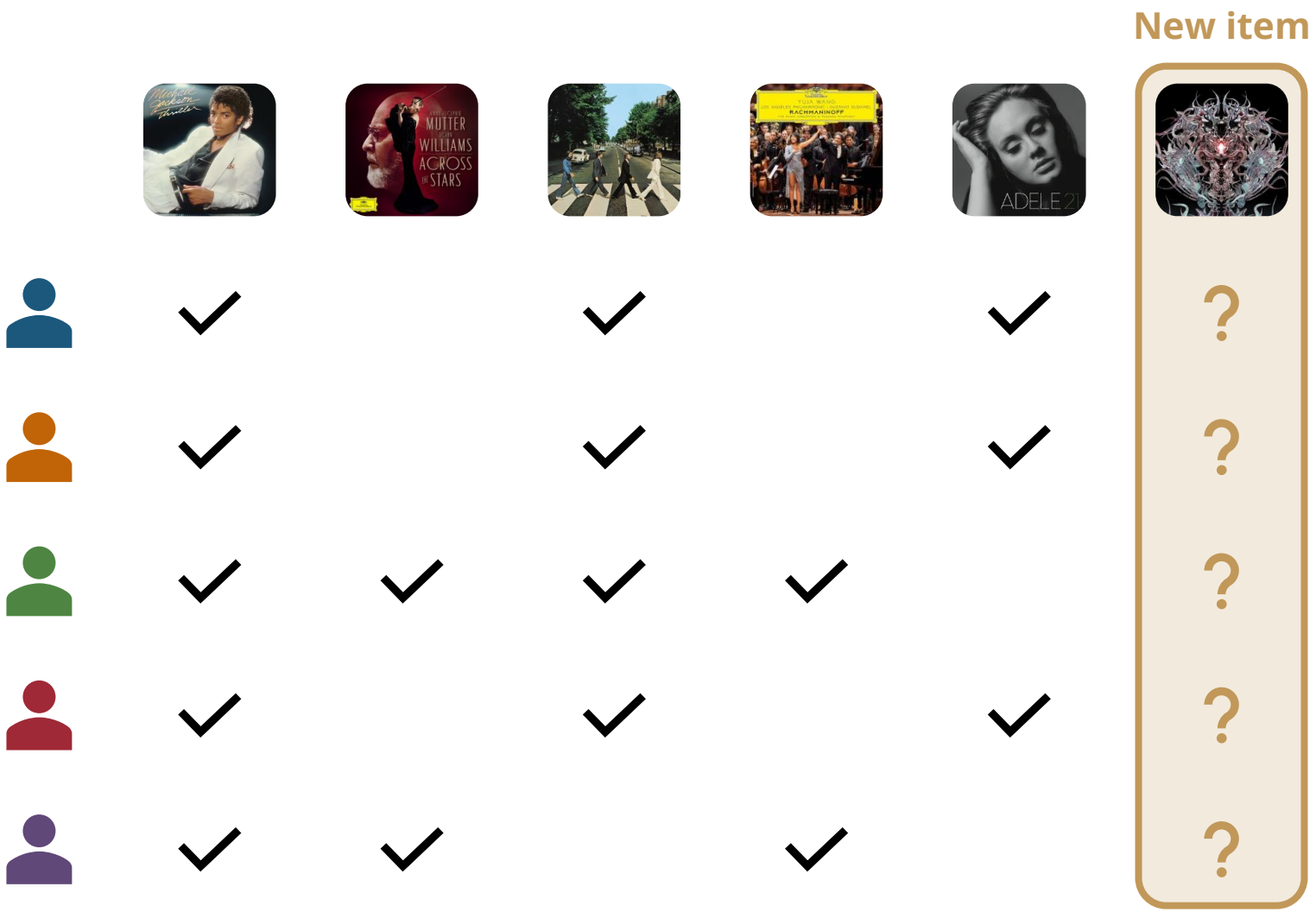
The Long Tail Problem



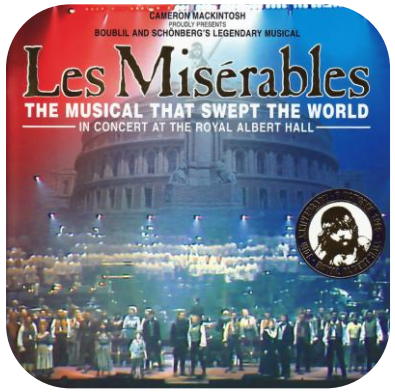
(Source: Levy & Bosteels, 2010)

Mark Levy and Klaas Bosteels, "Music Recommendation and the Long Tail," *WOMRAD*, 2010.

Cold Star Problem: New Items



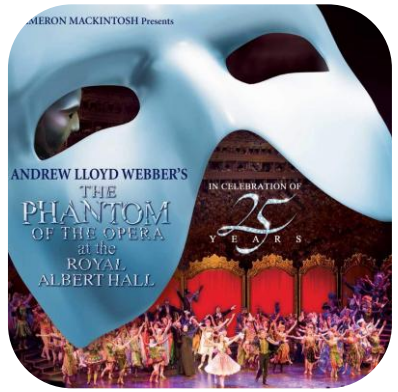
Content-based Filtering



Musical

Live concert version

Concert at Royal Albert Hall



Musical

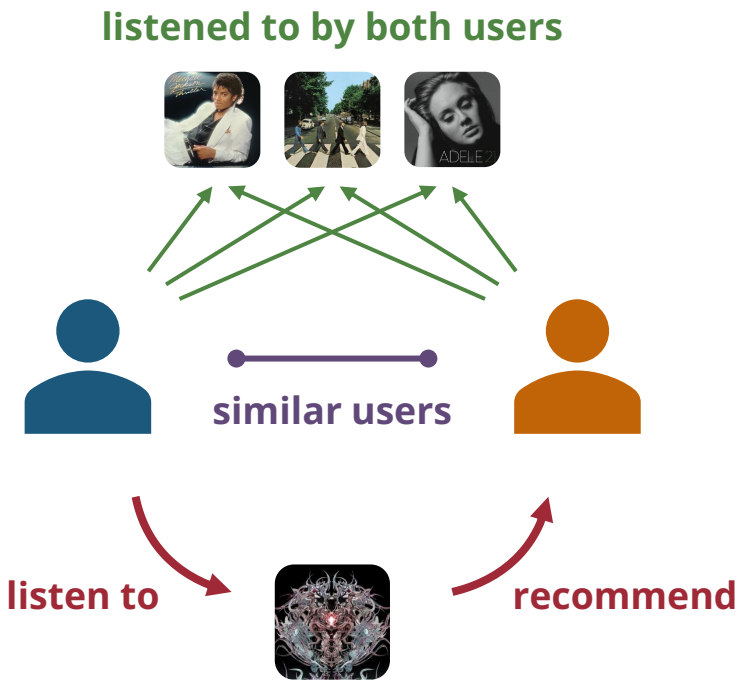
Live concert version

Concert at Royal Albert Hall

similar songs

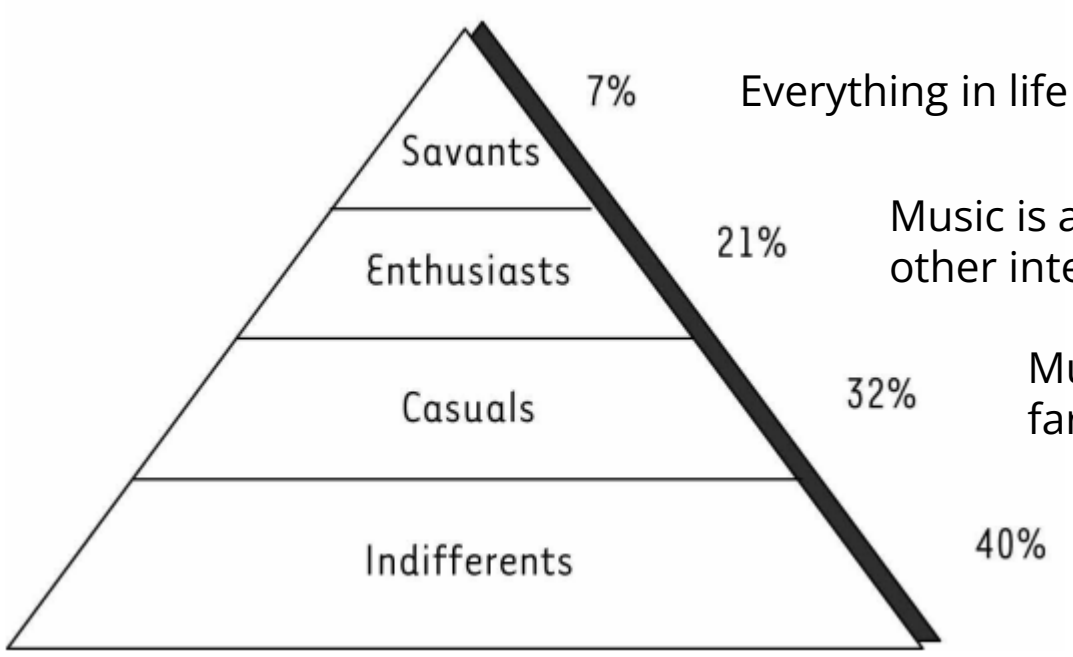
Collaborative Filtering vs Content-based Filtering

Collaborative filtering



Content-based filtering

User Listening Experience Modelling



Everything in life seems to be tied up with music

Music is a key part of life but is also balanced by other interests

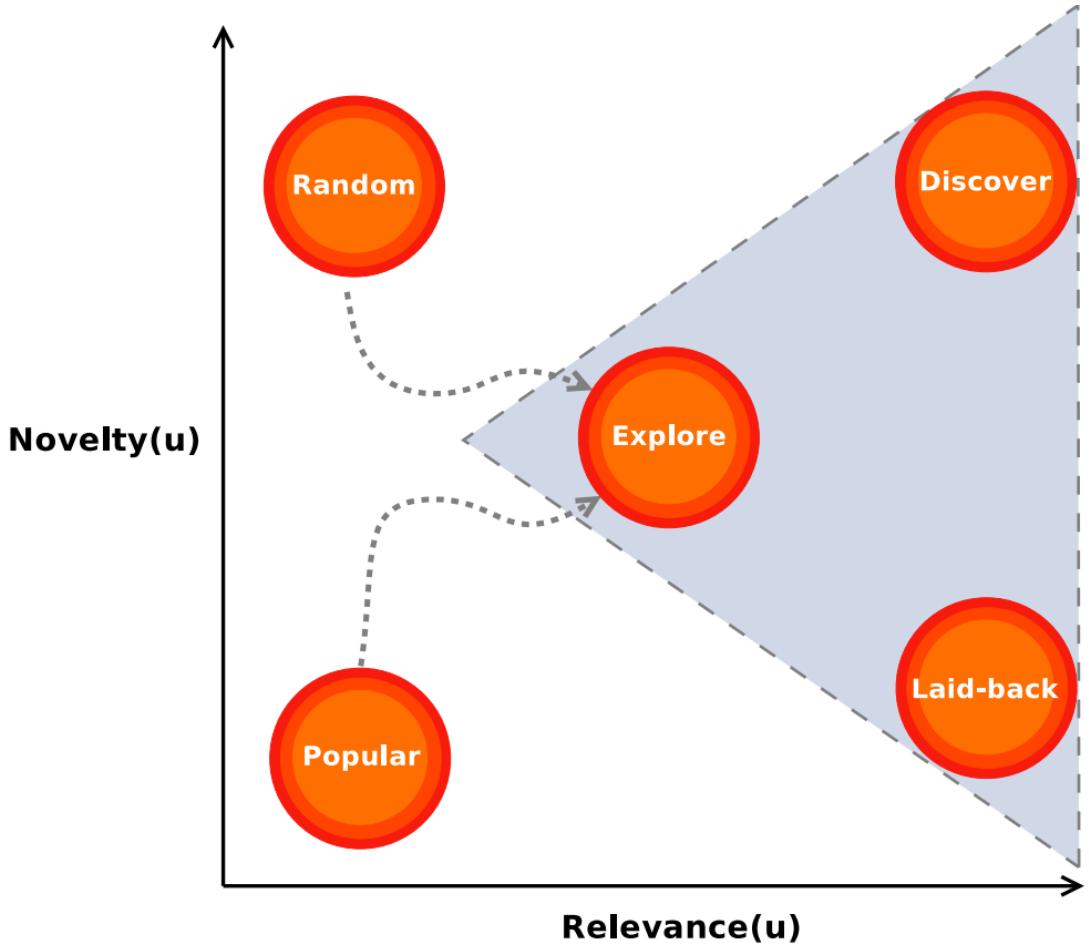
Music plays a welcome role, but other things are far more important

Would not lose much sleep if music ceased to exist

(Source: Jennings, 2007)

David Jennings, "Net, Blogs and Rock 'n' Roll: How Digital Discovery Works and What it Means for Consumers," *Nicholas Brealey Publishing*, 2007.

Novelty vs Relevance

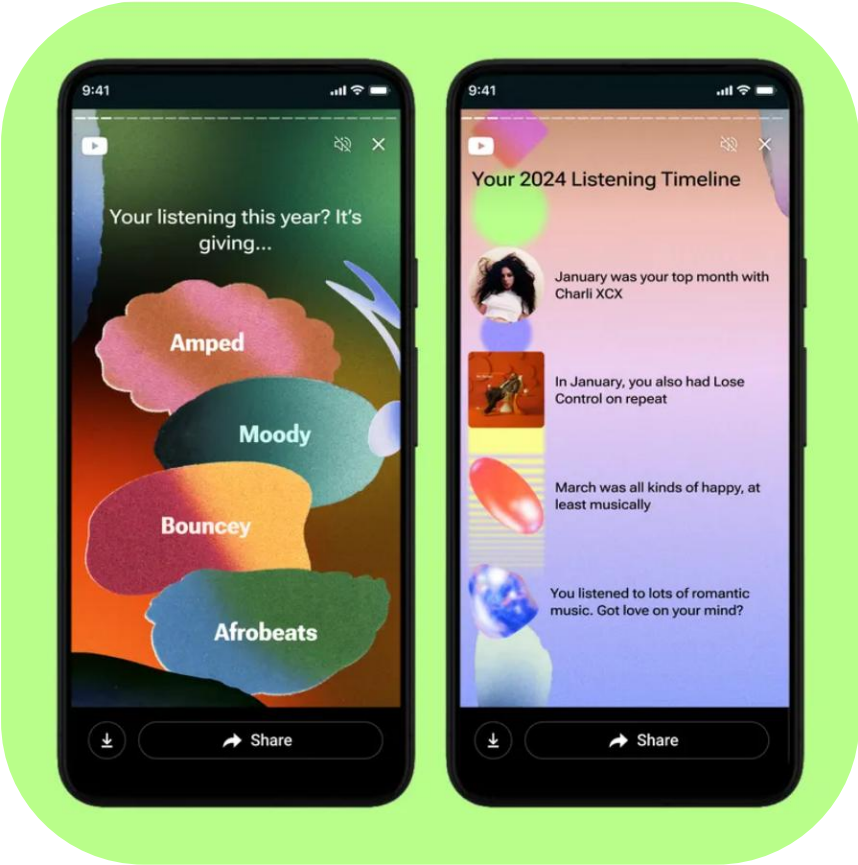


(Source: Celma, 2010)

Òscar Celma, "Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space," Springer, 2010.

Listening Behavior Analysis

YouTube's Music Recap



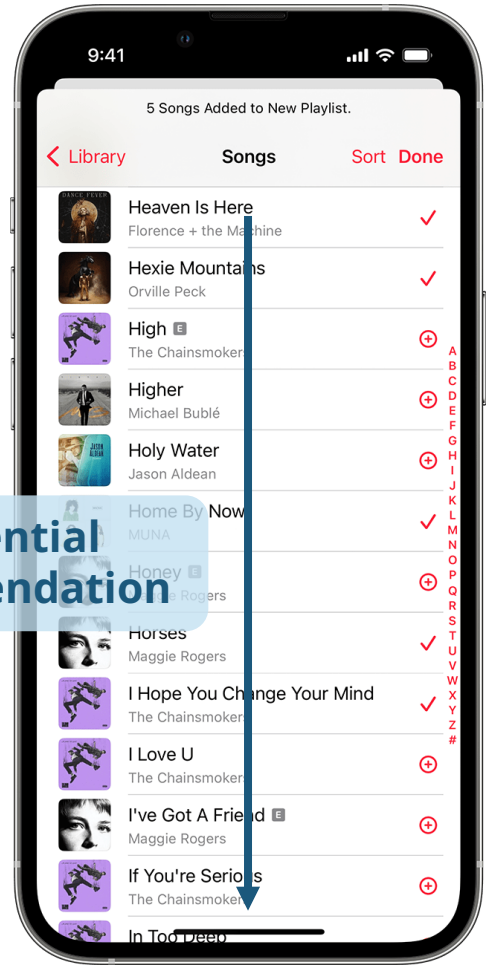
(Source: YouTube)

Spotify's Listening Personality

(Source: Spotify)

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

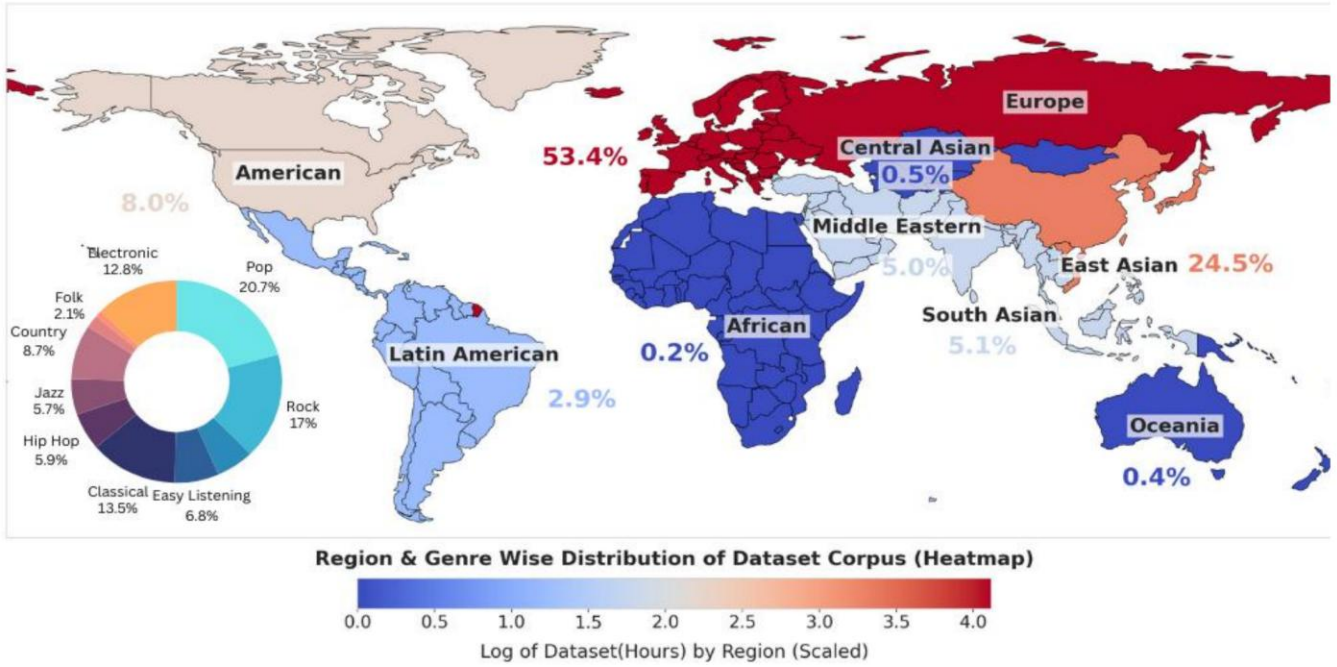
Music Playlist Generation



Sequential recommendation

Next Lecture

Discussion



(Source: Mehta et al., 2024)