FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

PETER KNEES

DIE PERFEKTE MUSIKEMPFEHLUNG – PERFEKT FÜR WEN?

ABOUT ME

- Music Information Retrieval researcher
    - Music search engines and interfaces
    - Music recommender systems
    - Recently: smarter tools for music creation
- PhD and PostDoc at JKU Linz (2005-2016)
- Since 2017: Assistant Professor at TU Wien

SESSION CONTENT

- Music Recommender Systems
  - Sources of data
  - Collaborative filtering
  - Content analysis
- Recommendation use cases
- Biases of platforms

RECOMMENDER CLASSIFICATION SCHEME

(based on users/community)	*(based on item's content)*	*(product finder)*
Collaborative Filtering (CF)	Content-based Recommenders	Knowledge-based Recommenders

Context-aware Recommenders

(based on the usage context)

Hybrid Recommenders

(a mixture of different approaches)

DATA FUELS RECOMMENDERS

Interaction Data

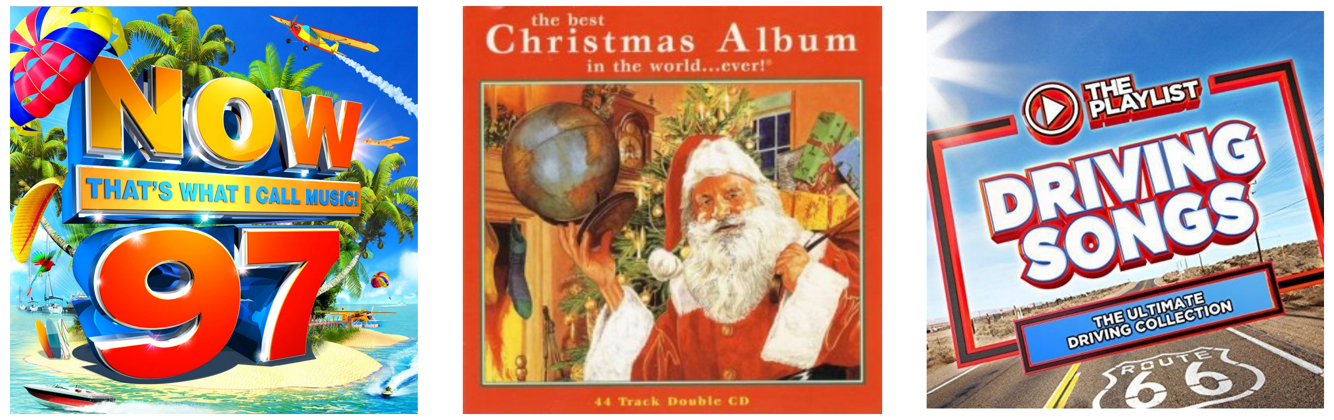
- Listening logs, listening histories
- Feedback ("thumbs"), purchases

User-generated

- Tags, reviews, stories

Curated collections

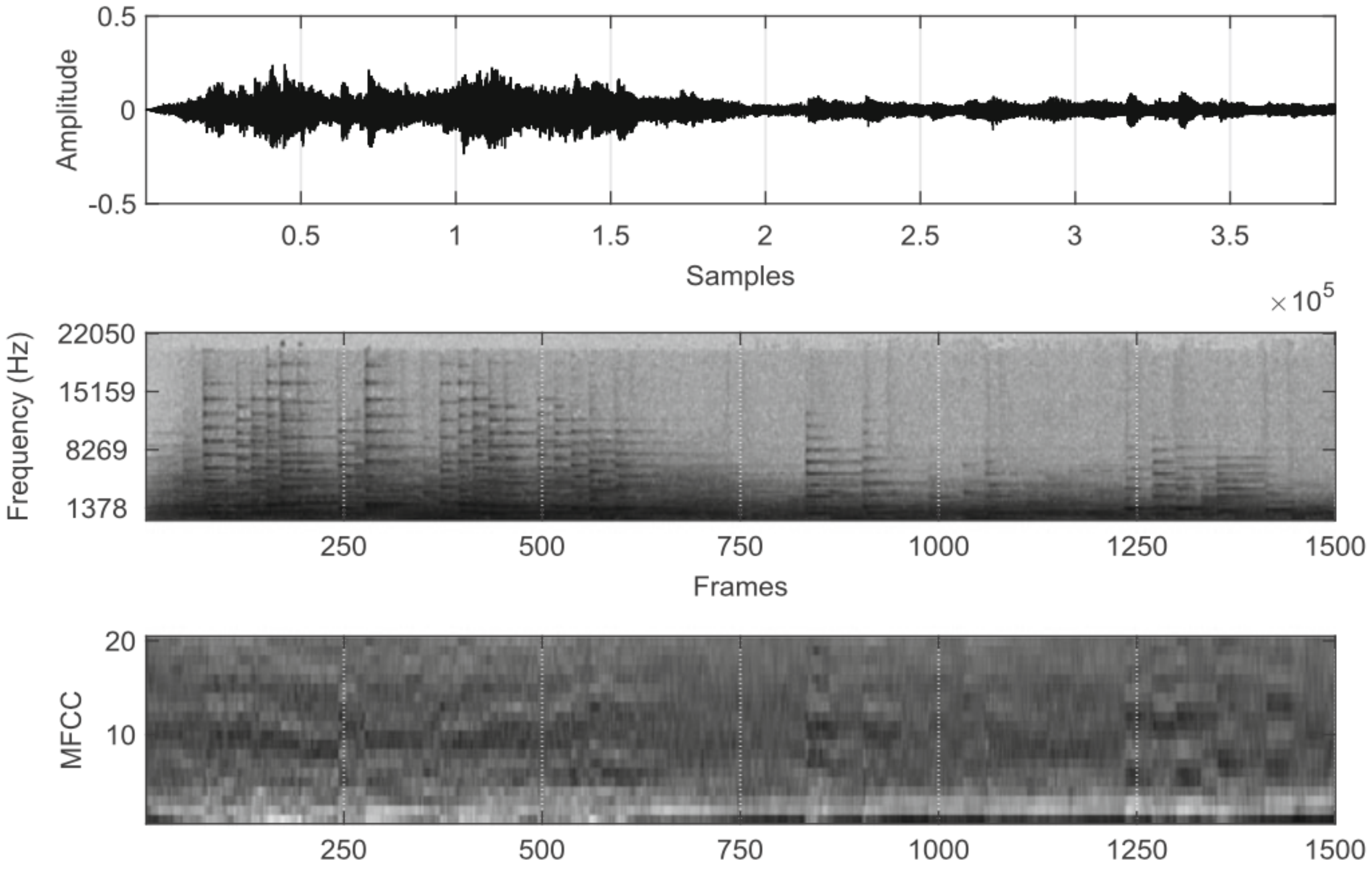
- Playlists, radio channels
- CD album compilations



DATA FUELS RECOMMENDERS

Content (audio, symbolic, lyrics)

- Machine listening/content analysis
- Human labelling

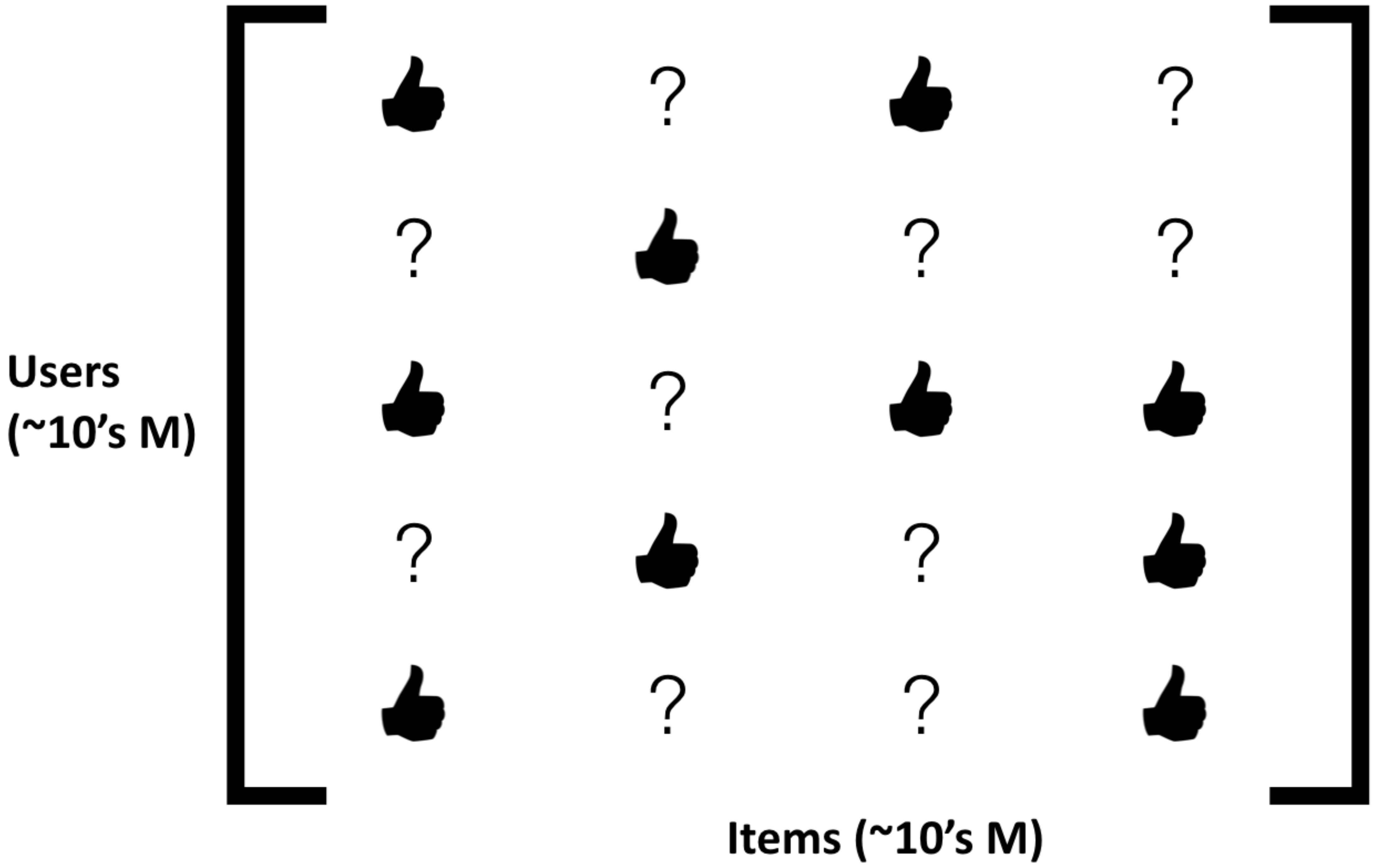


Meta-data

- Editorial
- Curatorial
- Multi-modal (album covers etc.)

COLLABORATIVE FILTERING

- Exploits **interaction data**
- *"People who listened to track A, also listened to track B"*
- Main underlying assumption: users that had similar taste in the past, will have similar taste in the future
- Typical methods
  - Comparing rows/columns in matrix
  - Matrix factorization



THE INTERACTION MATRIX

Can contain number of plays, listening time, rating, etc.

Listening	*Track 1*	*Track 2*	*Track 3*	*Track 4*
User 1	3		2	3
User 2	4	3	4	
User 3	3	2	1	4
User 4		5	4	1
User 5	5		3	

"user profile"

Similar users found, e.g. by comparing user profiles

FACTORS HIDDEN IN THE DATA

Assumption of matrix factorization-based recommender systems:

- Observed data are interactions of 2 factors: users and items

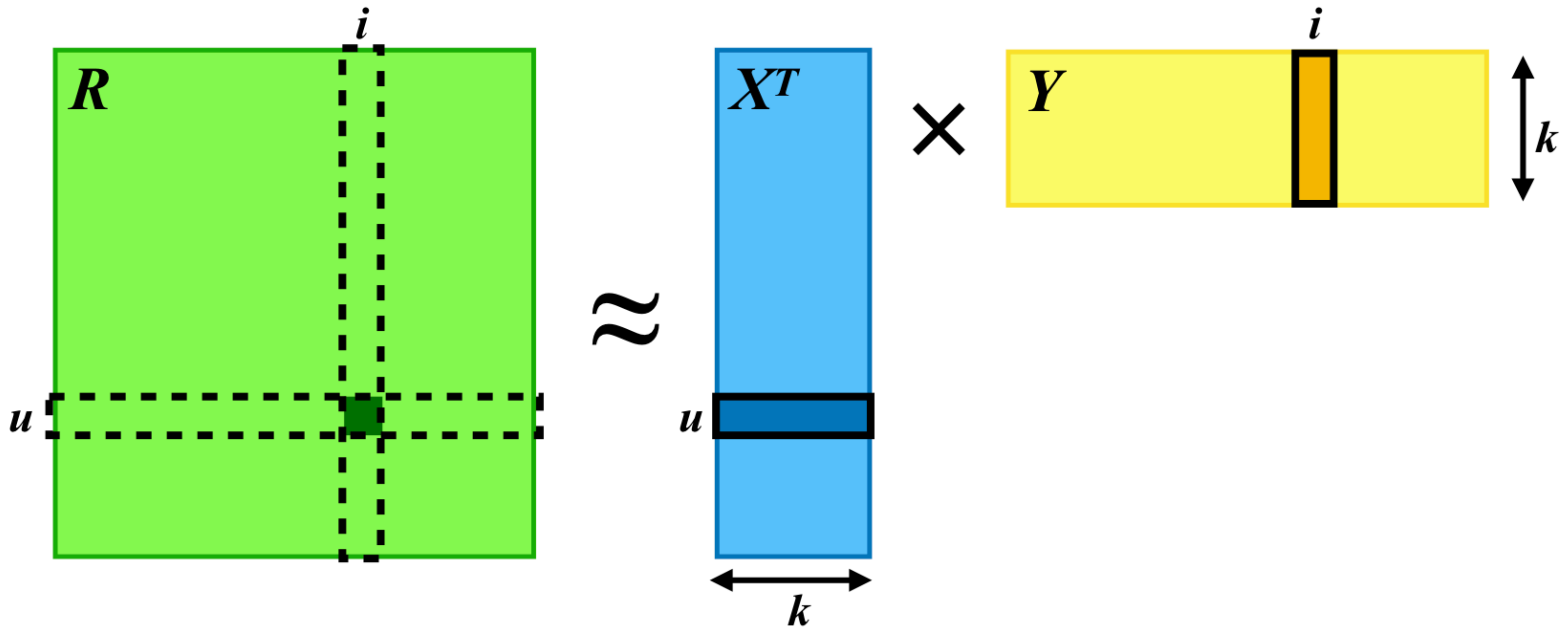
- Calculate latent factors for users and items from the data

MATRIX FACTORIZATION

- Decompose rating matrix into user and item matrices of lower dimension k
- Learning factors from given ratings using stochastic gradient descent

$$\min_{x_\star, y_\star} \sum_{r_{u,i} \text{ is known}} (r_{ui} - x_u^T y_i)^2 + \lambda(\|x_u\|^2 + \|y_i\|^2)$$

- Prediction of rating: inner product of vectors of user u and item i
- Factors not necessarily interpretable (just capture variance in data)



EXAMPLE OF COLLABORATIVE FILTERING OUTPUT

People who liked **Disturbed — The Sound of Silence**, also liked…

1. Bad Wolves – Zombie

2. Five Finger Death Punch – Bad Company

3. Disturbed – The Light

4. Metallica – Nothing Else Matters

Effect/issue: popularity biases

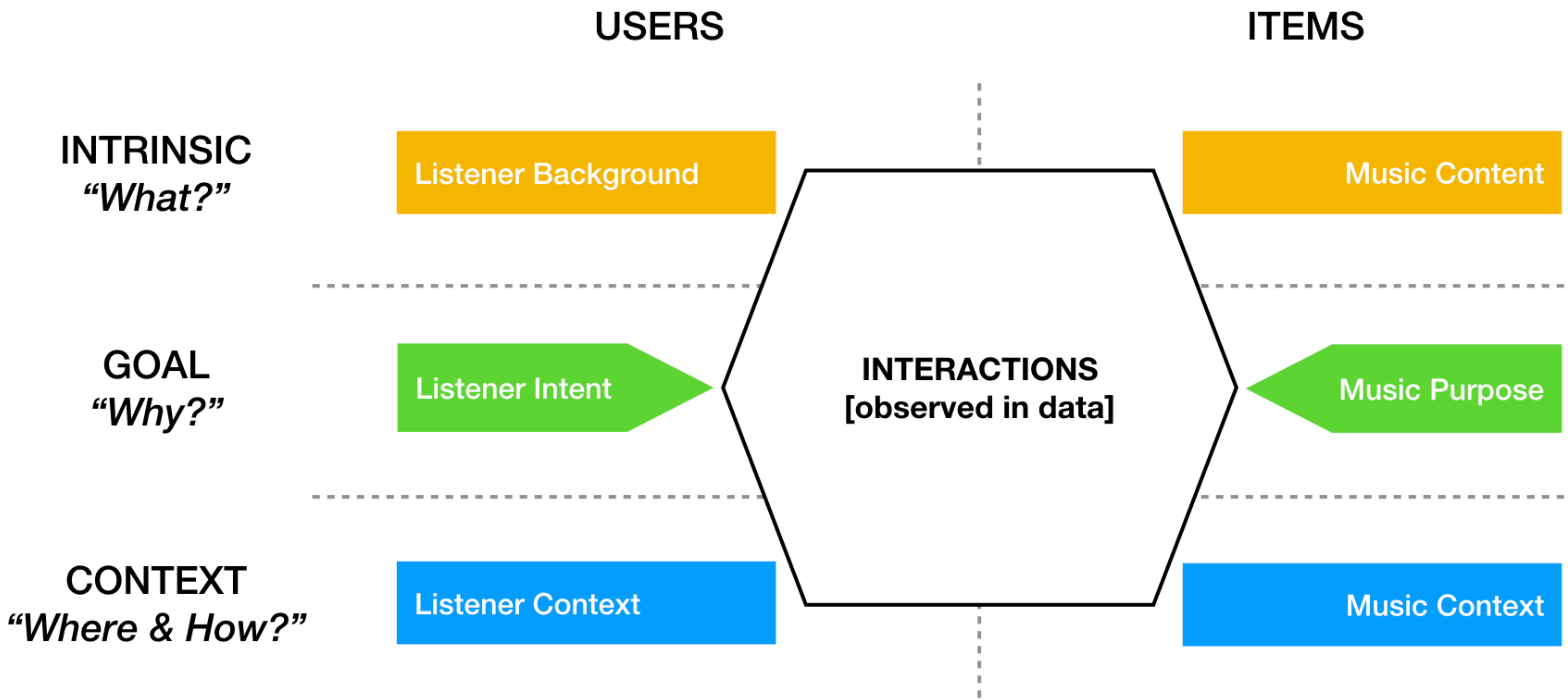
FACTORS HIDDEN IN THE DATA

Assumption of matrix factorization-based recommender systems:

- Observed data are interactions of 2 factors: users and items

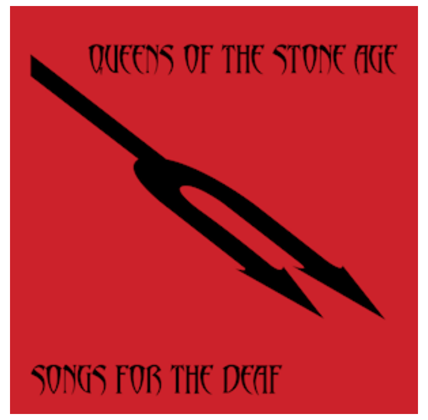
- Calculate latent factors for users and items from the data
- But it's a bit more complex…

FACTORS HIDDEN IN THE DATA



AUDIO CONTENT ANALYSIS: SELECTED FEATURES

Disturbed
The Sound of Silence



- Beat/downbeat → Tempo: 85 bpm
- Timbre
  e.g. for genre classification, "more-like-this" recommendations
- Tonal features
  e.g. for melody extraction, cover version identification

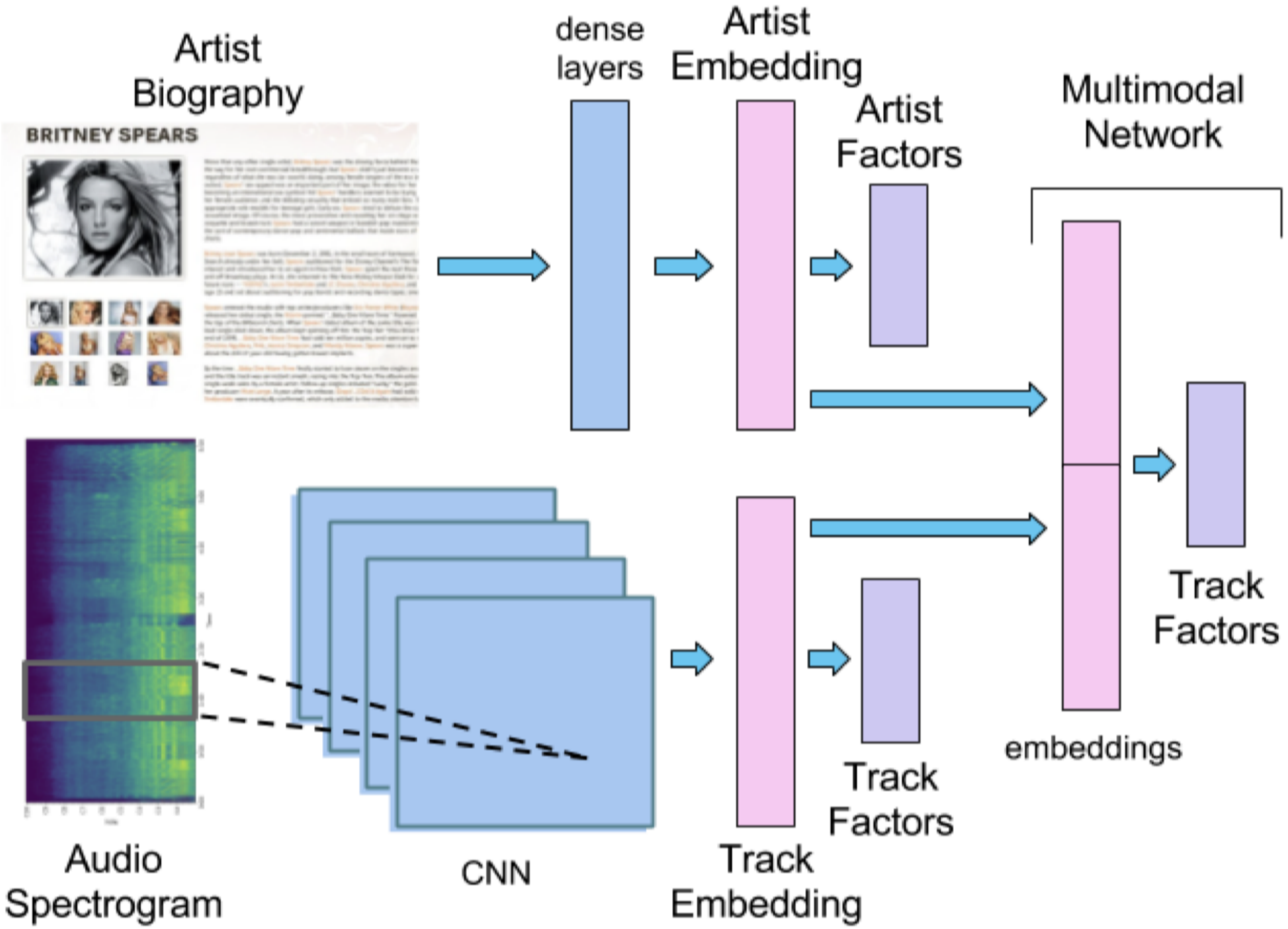
Different versions of this song:
Simon & Garfunkel - The Sound of Silence
Anni-Frid Lyngstad (ABBA) - En ton av tystnad
etc.

- Semantic categories via machine learning
  `not_danceable, gender_male, mood_not_happy`

Effect/issue: no popularity biases, but no personalization

A MIXTURE OF MANY THINGS

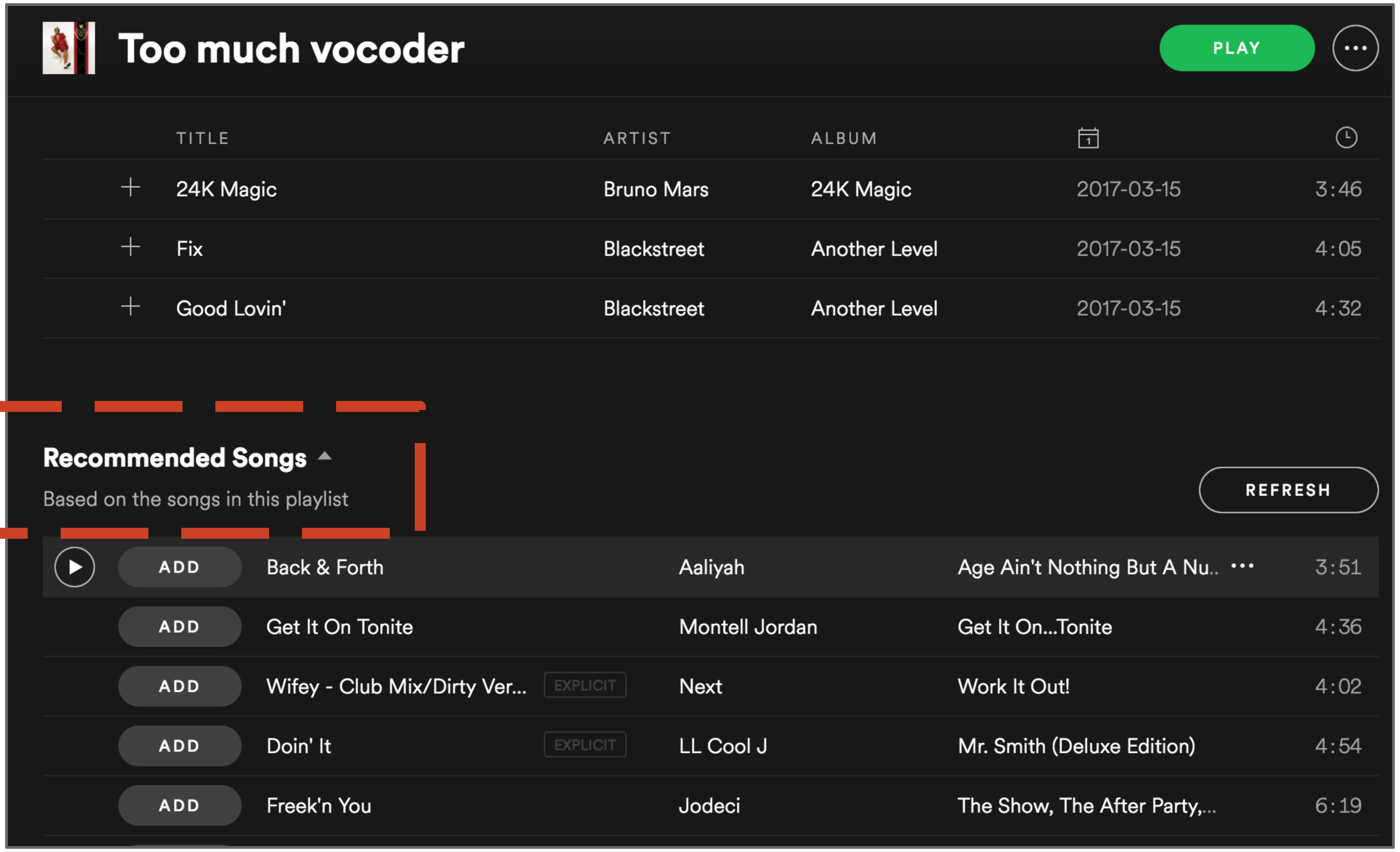
- Incorporation of different sources and complementary information
- Machine Learning to fit which recommender/information in which context
- E.g. to control for diversity, exploitation vs exploration, novelty, etc.
- Different types of recommenders and models for different features

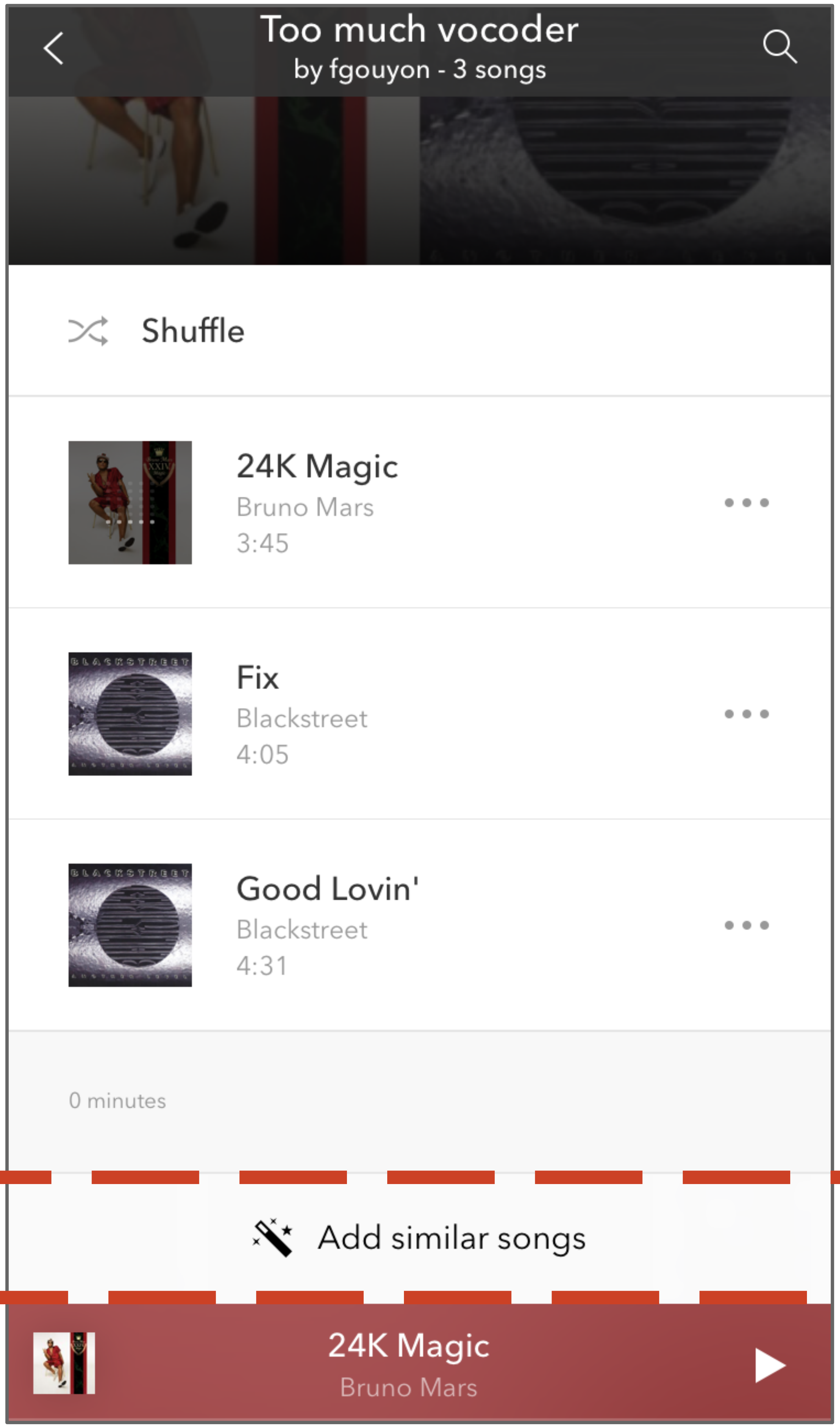


Oramas et al., RecSys DLRS 2017

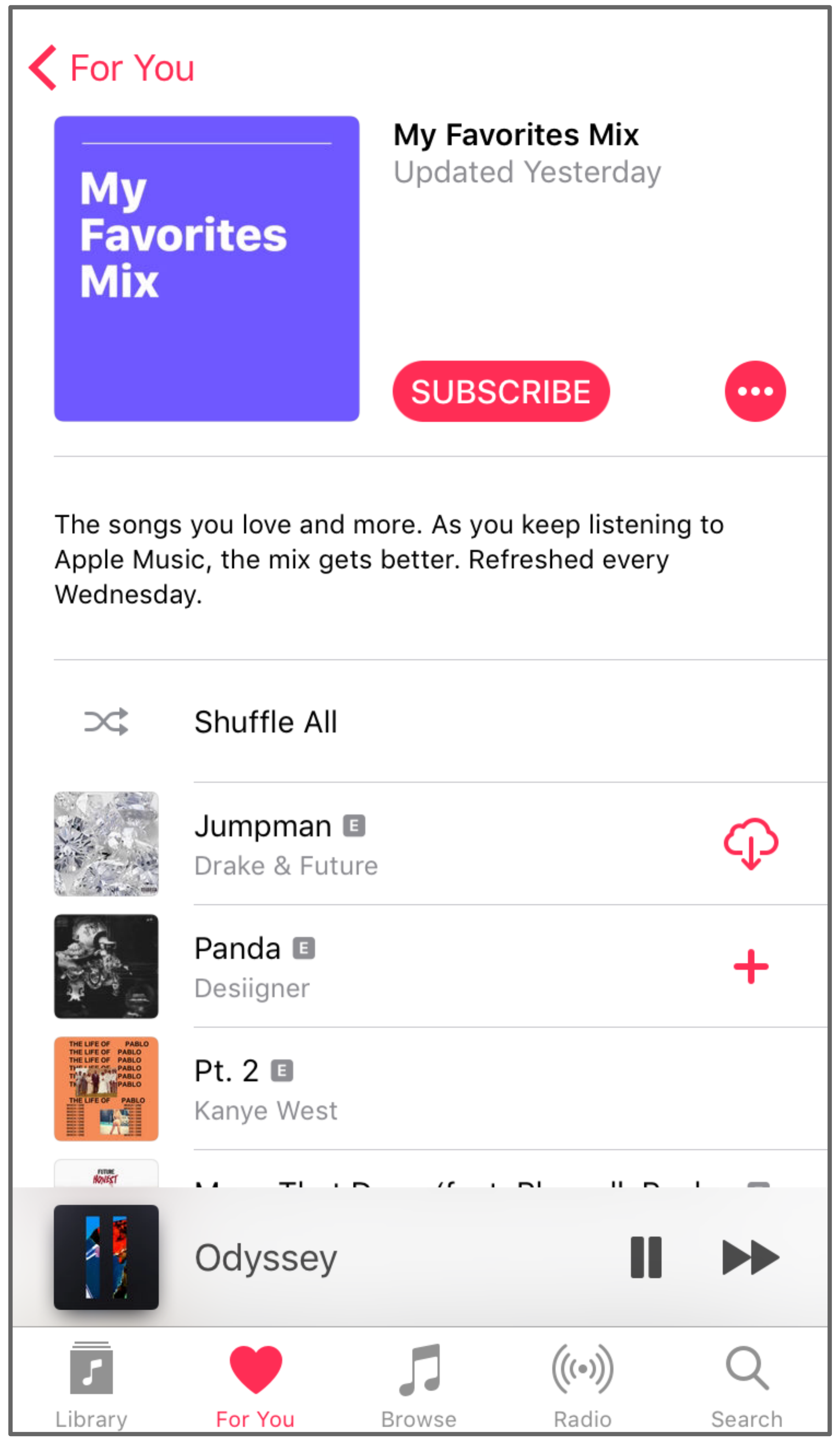
FOCUS ON: LEAN-IN EXPERIENCE

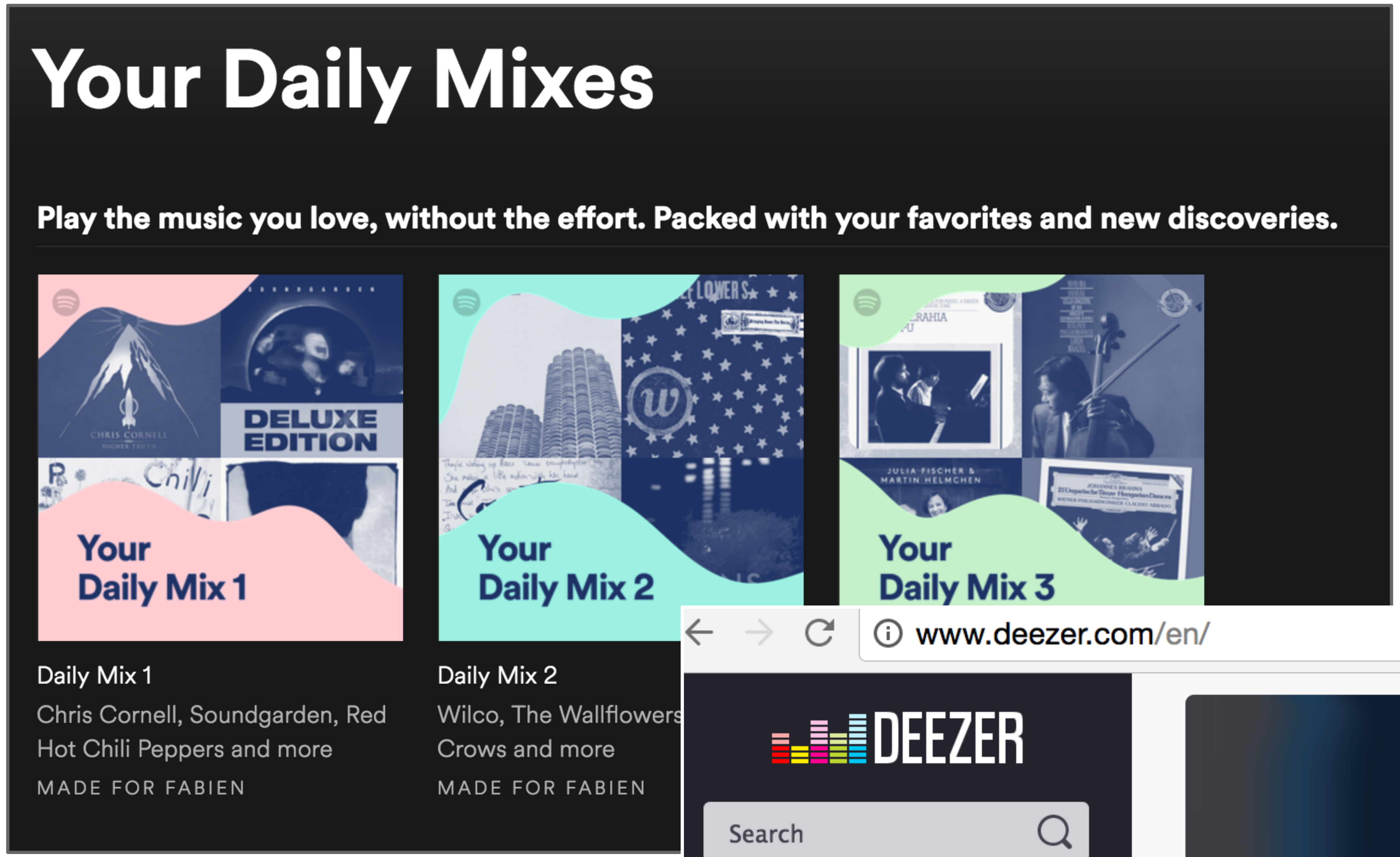
Lean in: Building Playlists

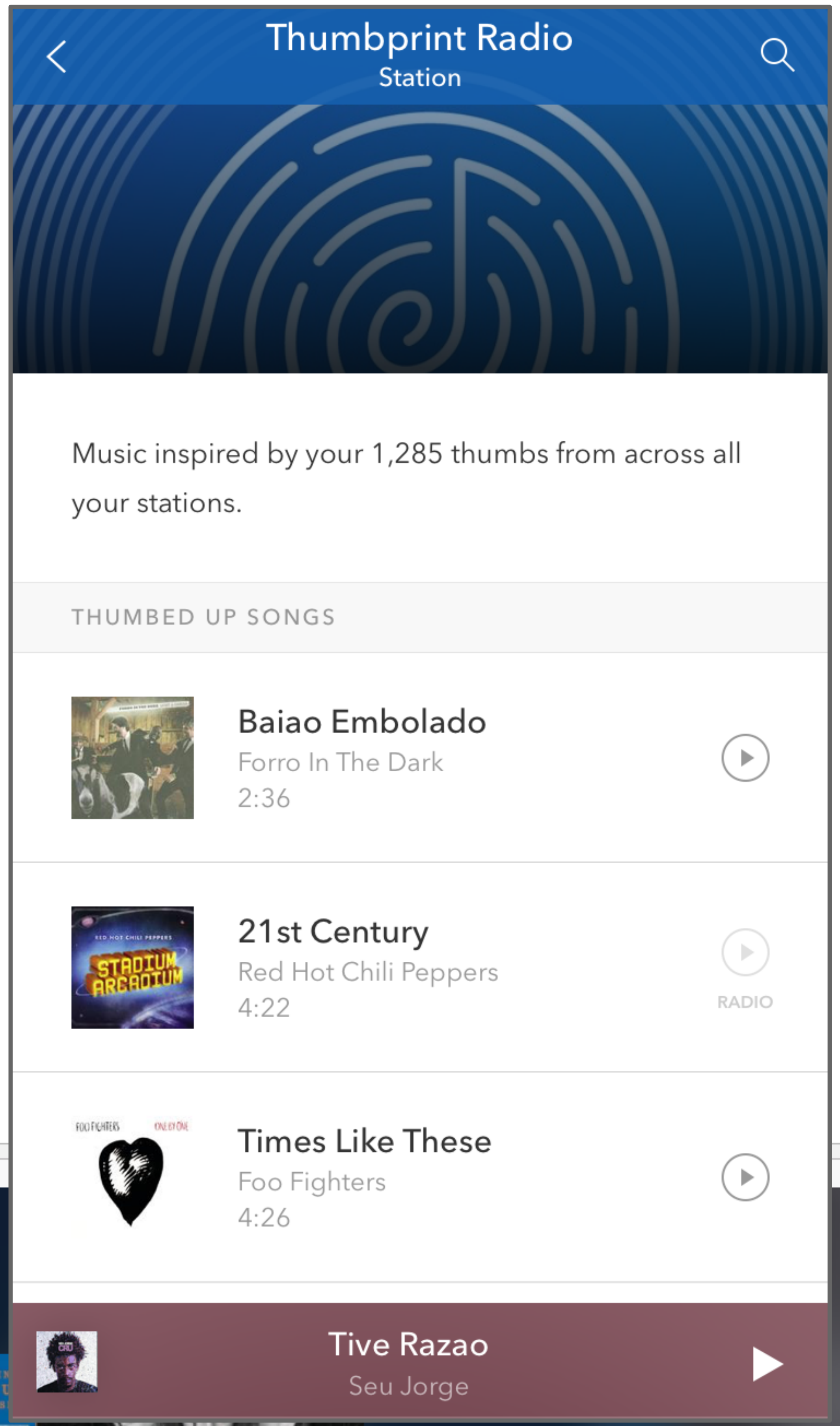


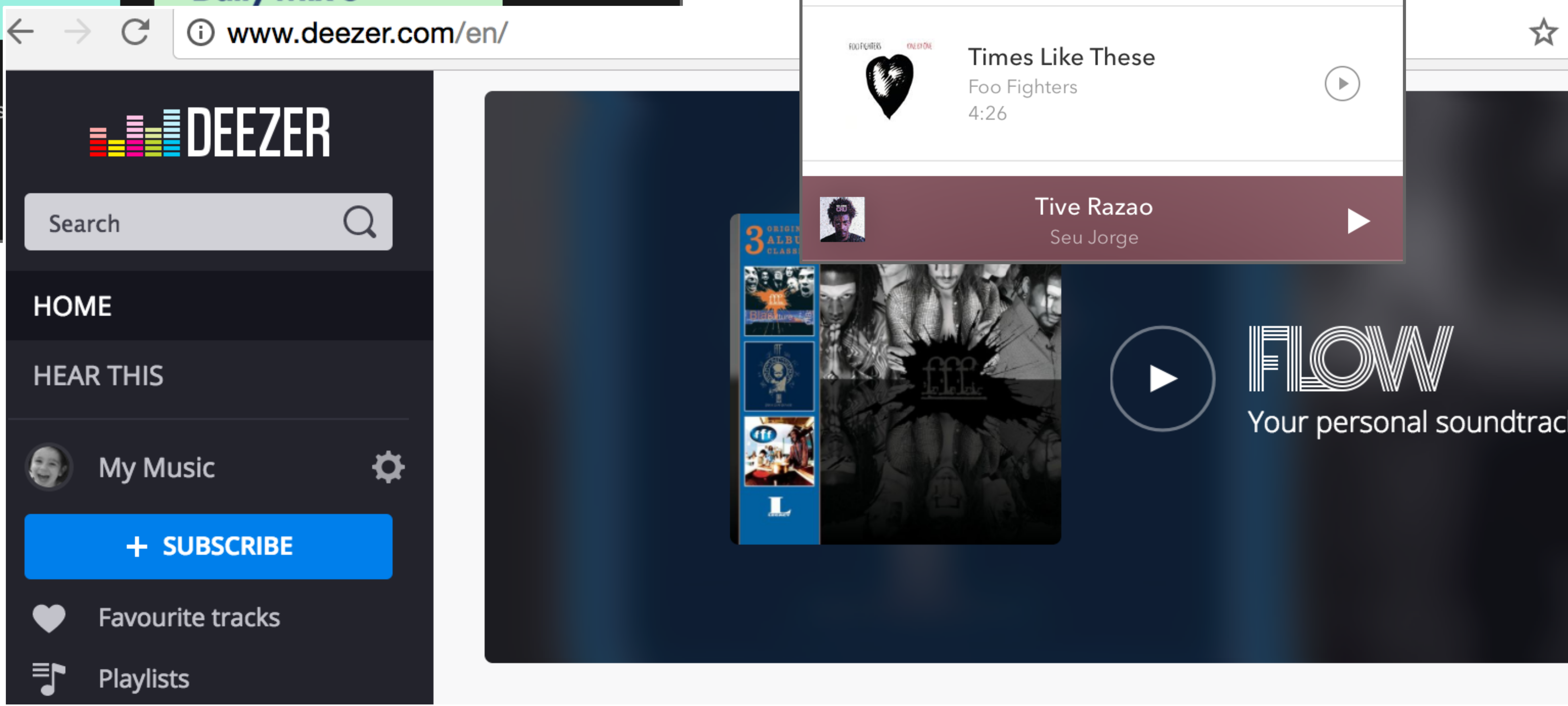


FOCUS ON: RE-DISCOVERY



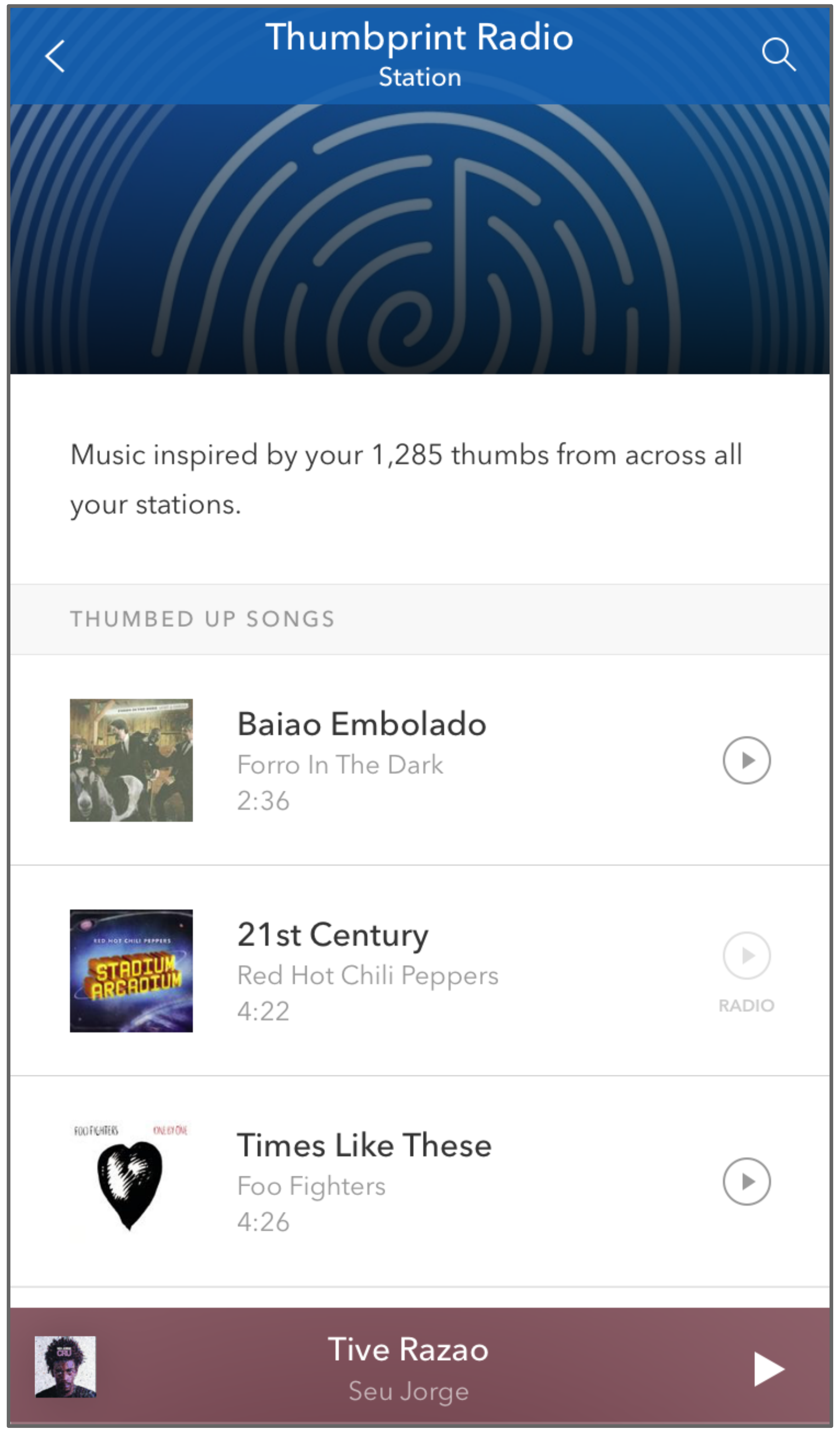


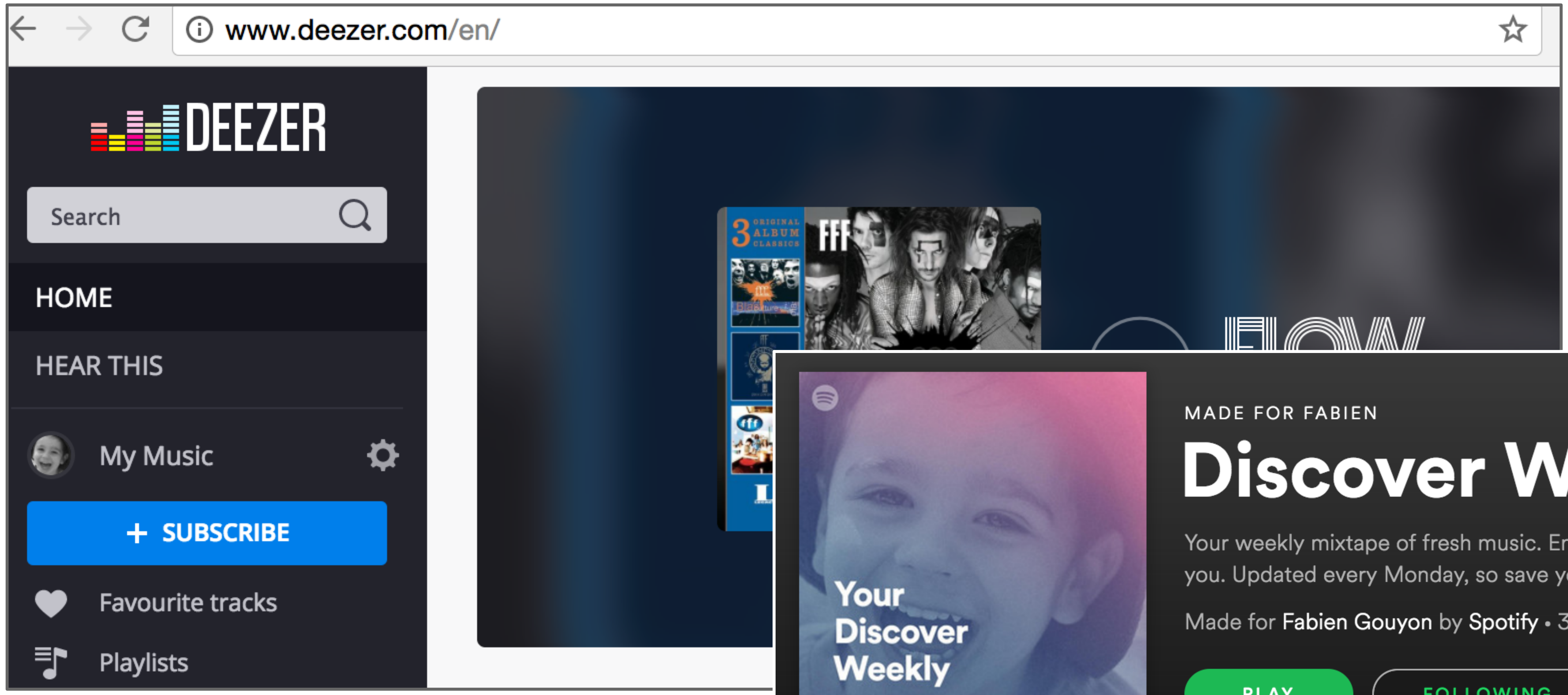


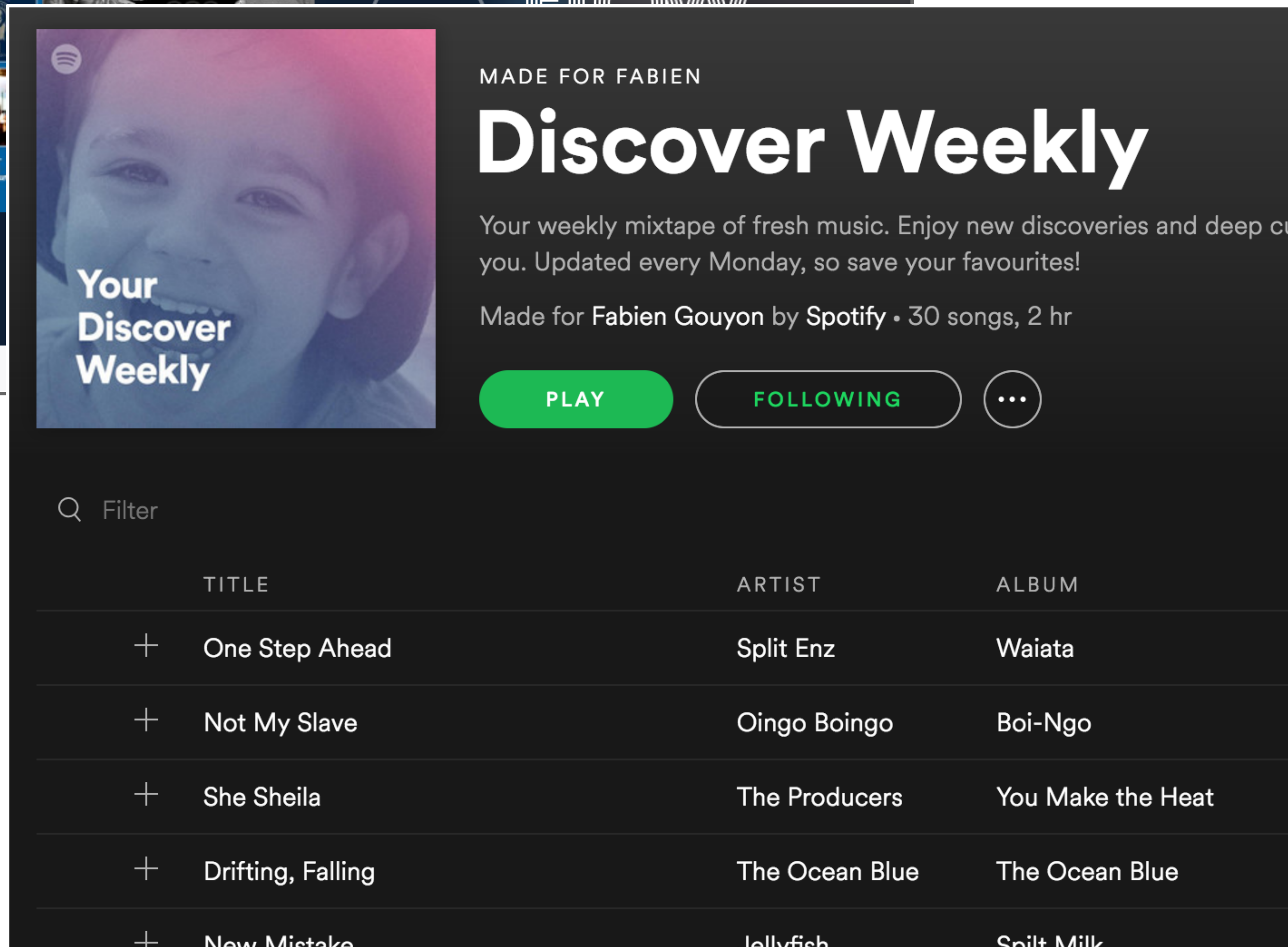


Focus on stuff you know you like
Personalized, leaning towards exploit

FOCUS ON: HYPER-PERSONALIZED DISCOVERY



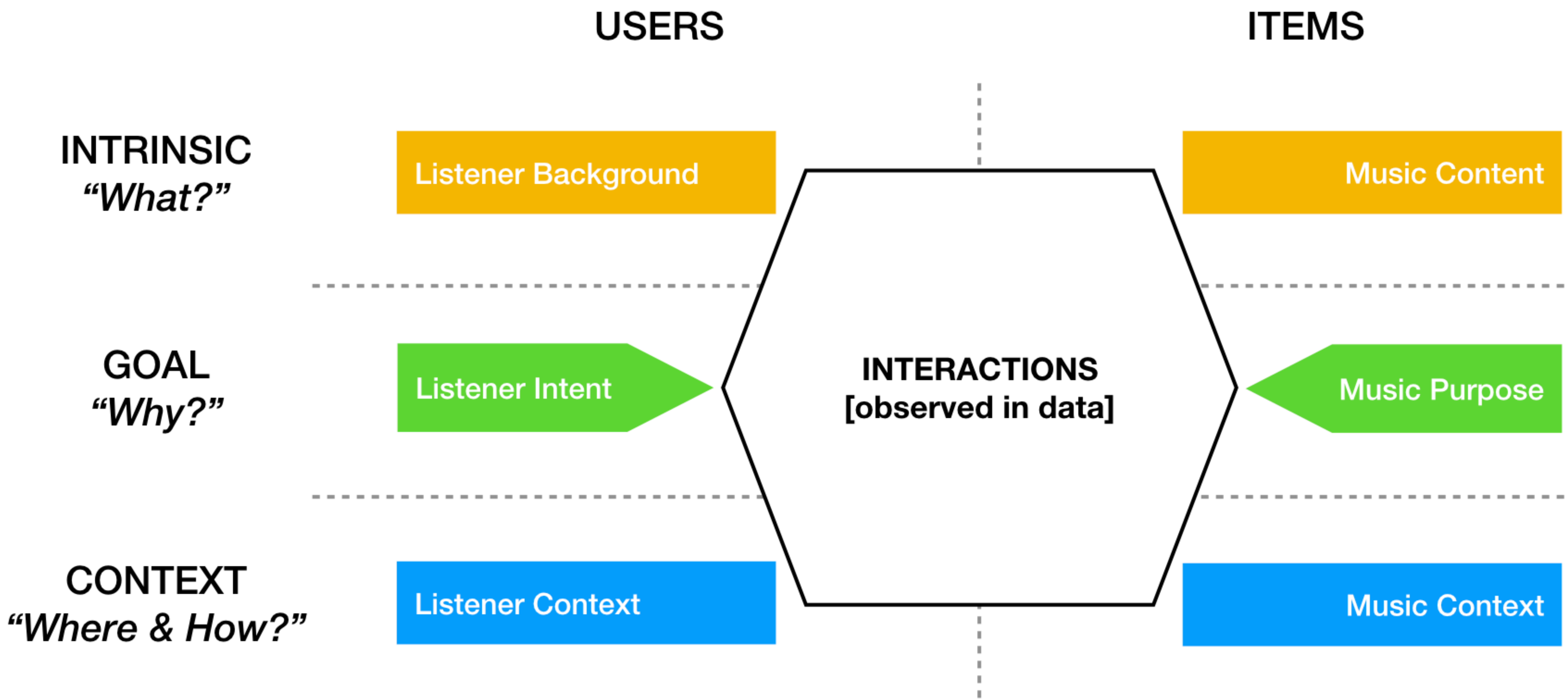




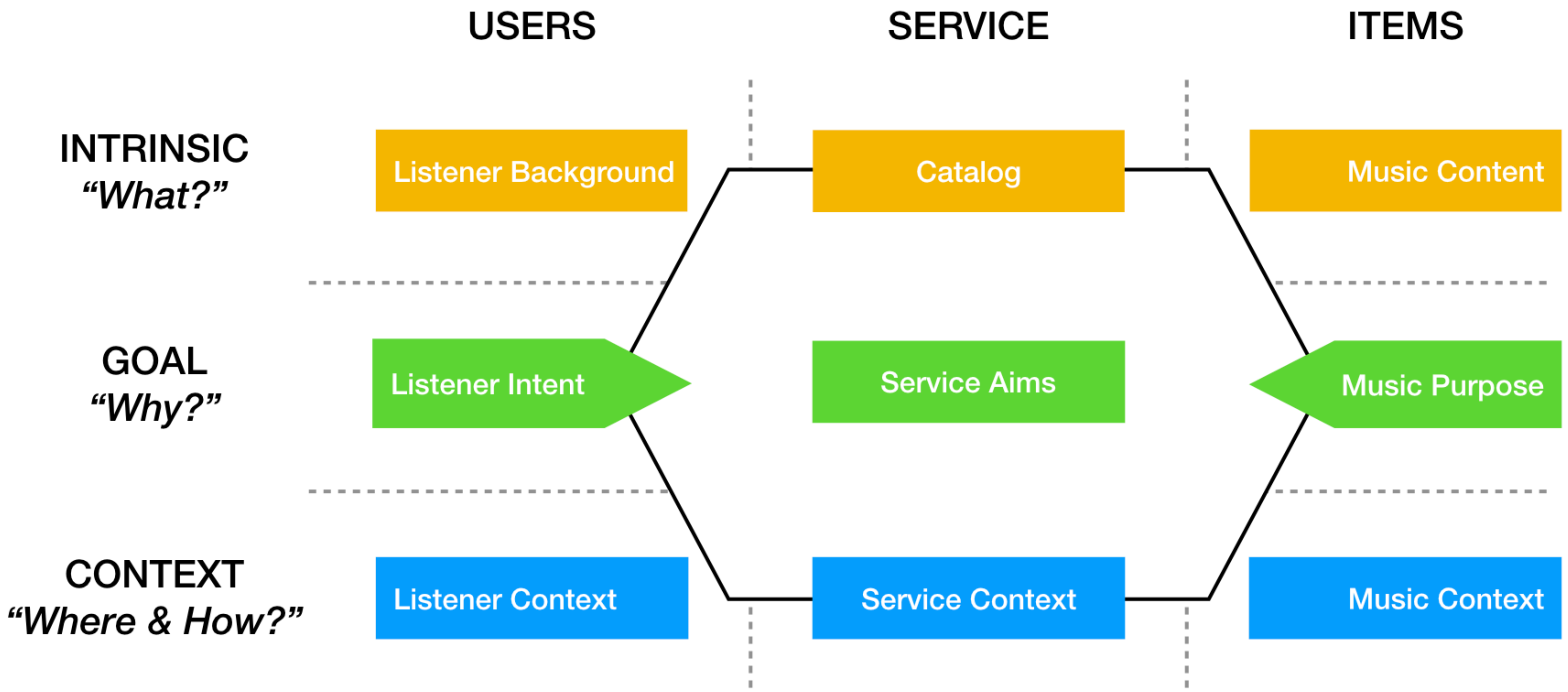
About discovering new stuff. Intended to feel like it's curated. Just. For. Me.

Leaning towards explore

ONE MORE THING…



FACTORING THE SERVICE INTO THE PICTURE



THE SERVICE INTRODUCES FURTHER BIASES

Catalog

- Which content is provided/recommended?
- e.g. Soundcloud recommends different content than Spotify

Service Aims

- Why is this service in place? What is the purpose/identified market niche?
- What are the identified use cases? (Discovery? Radio? Exclusives? Quality?)
- Do they push their own content (cf. Netflix)?

Service Context

- How do catalog and service aims depend on context?
- Are there licensing issues/restrictions in particular countries?
- Is the service context-aware? (e.g. app vs desktop/browser)

IMPLICATIONS

- Different methods with different biases incorporated
- Algorithmic design choices to deal with biases
- Service design choices and restrictions introduce biases
- Feedback loops amplify popularity biases
- Platforms are in control and can shape recommendations

POPULAR ACCESS MUSIC PROFILE STATS CONCERTS **PROMOTION** UPLOAD BETA MASTERIN

How can I promote my music on Spotify?

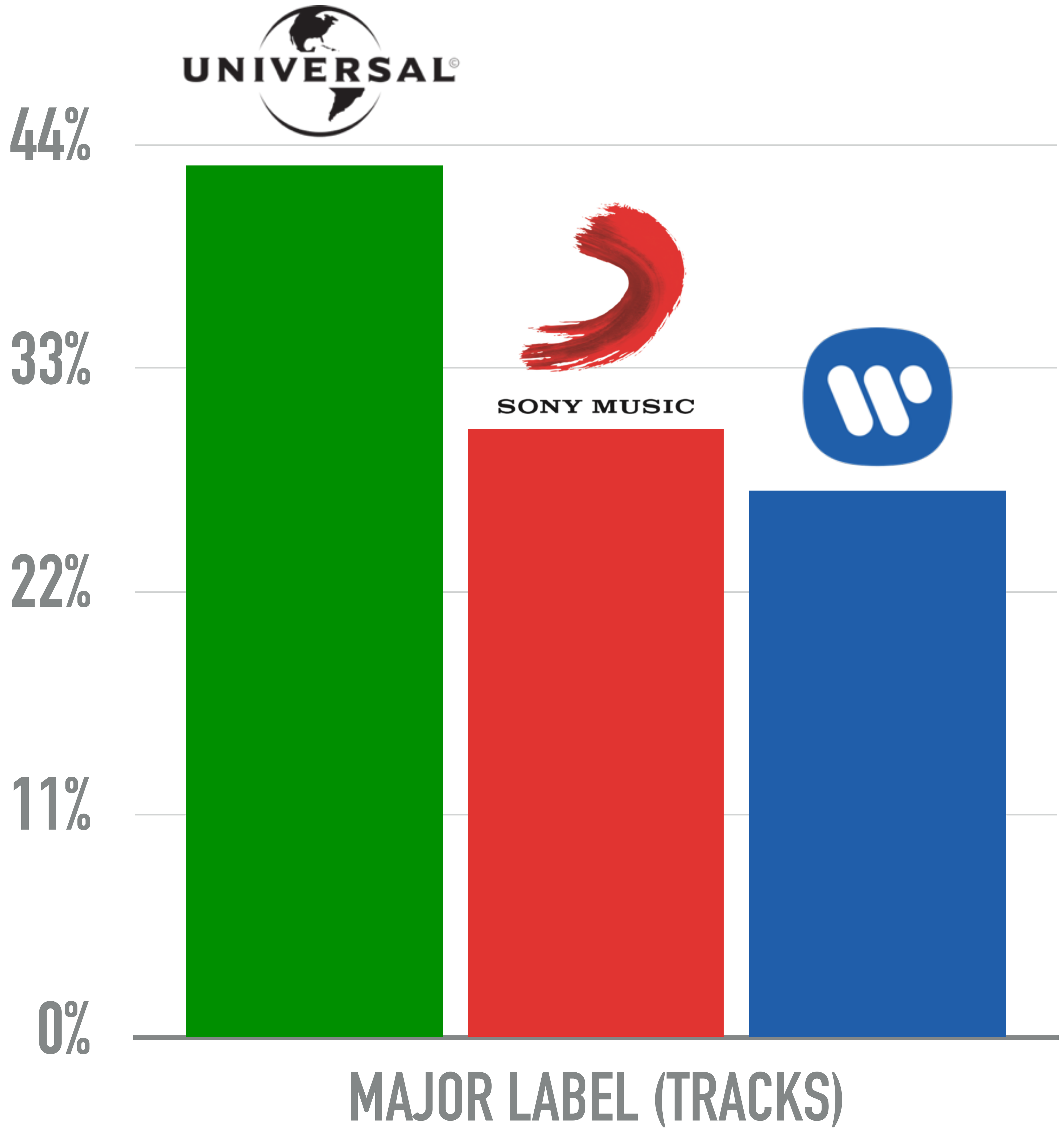
How do I get my music on a Spotify playlist?

How do I submit music to your Editorial team?

https://artists.spotify.com/faq/promotion

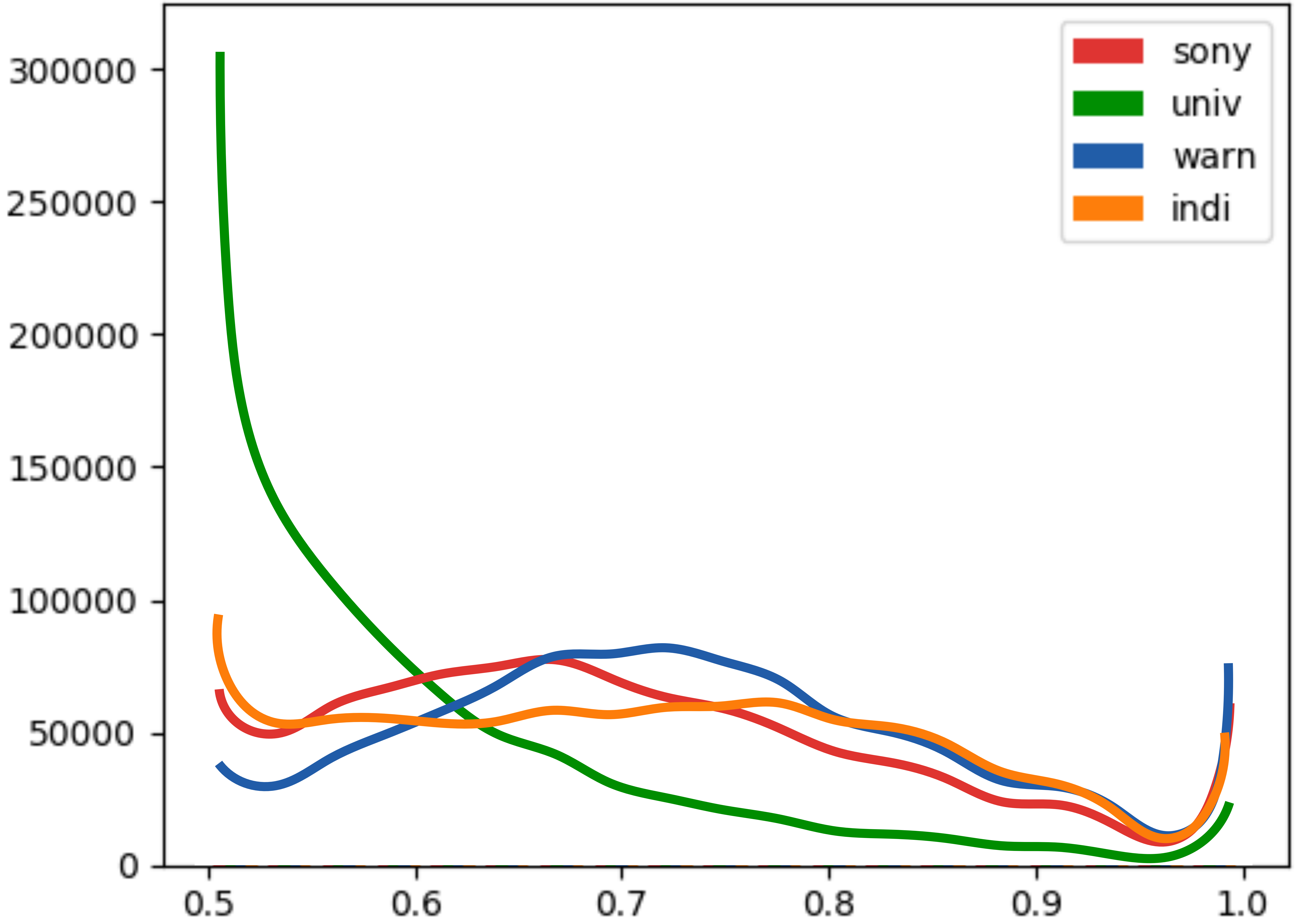
ALL ABOUT THE MUSIC …?

- Analysis of Spotify playlist dataset
- 1 million US playlists
- Webcrawler to identify record label of tracks
- Information for about 50% of tracks



ALL ABOUT THE MUSIC …?

- Investigating playlist diversity wrt. labels
- Entropy-based
- Left: pure
  Right: diverse
- Any comments?



CONTACT

FAKULTÄT FÜR !NFORMATIK

Faculty of Informatics

- Dr. Peter Knees
  Assistant Professor
  TU Wien, Institute of Information Systems Engineering

peter.knees@tuwien.ac.at

https://www.ifs.tuwien.ac.at/~knees/

@peter_knees

https://www.linkedin.com/in/peterknees/