TUTORIAL

RECSYS'17

Monday 08/28/2017

New Paths in Music Recommender Systems Research

Markus Schedl
JKU Linz

Peter Knees
TU Wien

Fabien Gouyon
Pandora Inc.

JOHANNES KEPLER UNIVERSITÄT LINZ

TU WIEN

pandora

The Band

Markus Schedl (@m_schedl)
Associate Professor
Johannes Kepler University Linz, Austria

Peter Knees (@peter_knees)
Assistant Professor
TU Wien, Austria

Fabien Gouyon (@fgouyon)
Principal Scientist at Pandora
President, International Society for Music Information Retrieval

Overview

- Introduction to Music Recommendation
- It's All About the Use Cases
- Use Case 1: Station/Playlist Generation
- Use Case 2: Context-Aware Music Recommendation
- Use Case 3: Recommendation in the Creative Process of Music Making
- What's Next?

The Deck

Latest version of slides available at:

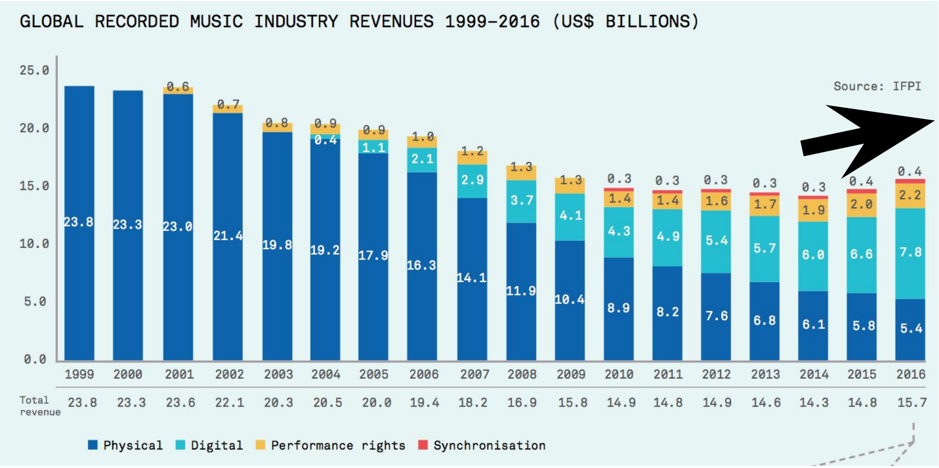
http://www.cp.jku.at/tutorials/mrs_recsys_2017/

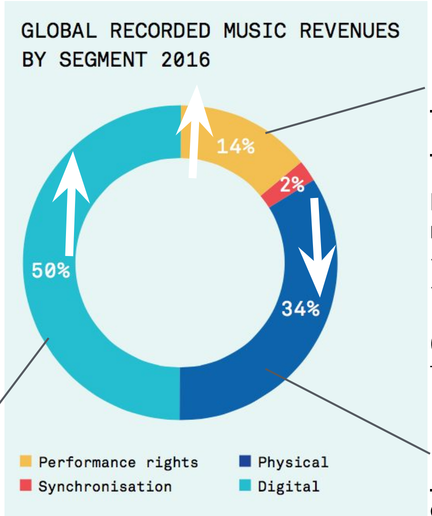
Overview paper available at:

http://www.cp.jku.at/tutorials/mrs_recsys_2017/overview_paper.pdf

Intro

Music Industry Changing Landscape





PERFORMANCE RIGHTS

Revenue from music reproduction:
- on AM/FM radio
- at public venues

(NB: Excluding perf. rights from Streaming)

PHYSICAL

e.g. CDs

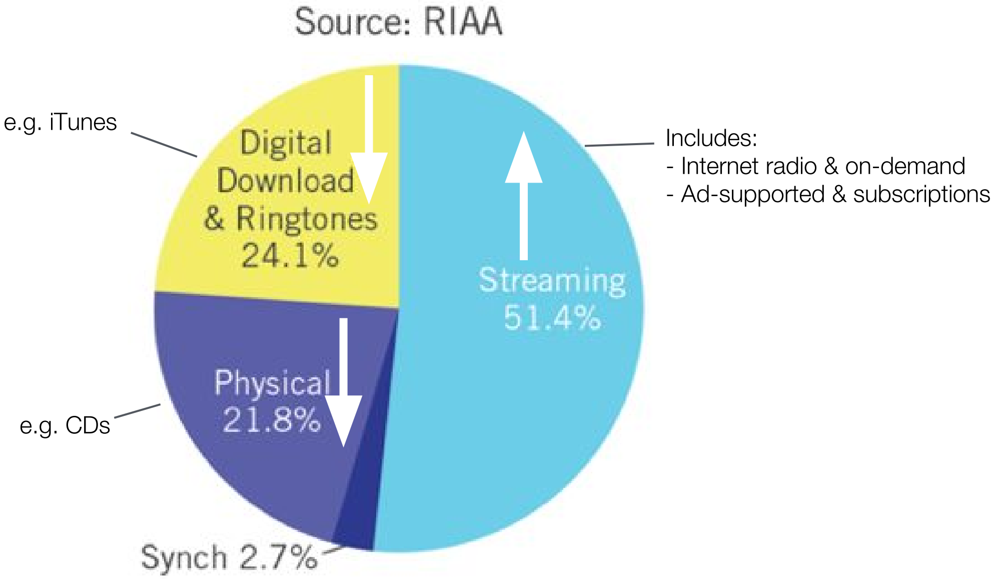
DIGITAL

59% of which is Streaming, i.e.:
- Internet radio & on-demand
- Ad-supported & subscriptions

Also includes **Downloads** (e.g. iTunes) - which are **declining**

Music Industry Changing Landscape

Revenue breakdown by media, in US, 2016:
(Performance rights not shown)



Physical were 50% in 2010
Streaming was 9% in 2011

2016, in US, of those consuming music:

- 75% used streaming
- 20% bought CDs

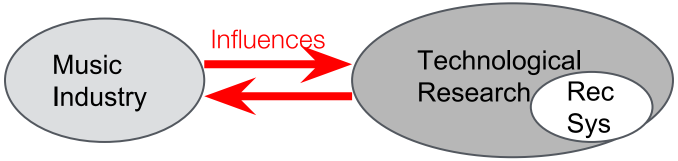
2016: US music industry saw **biggest gains** since 1998

Music Industry Changing Landscape

- Growing industry
- Accelerating transition: Physical → Streaming

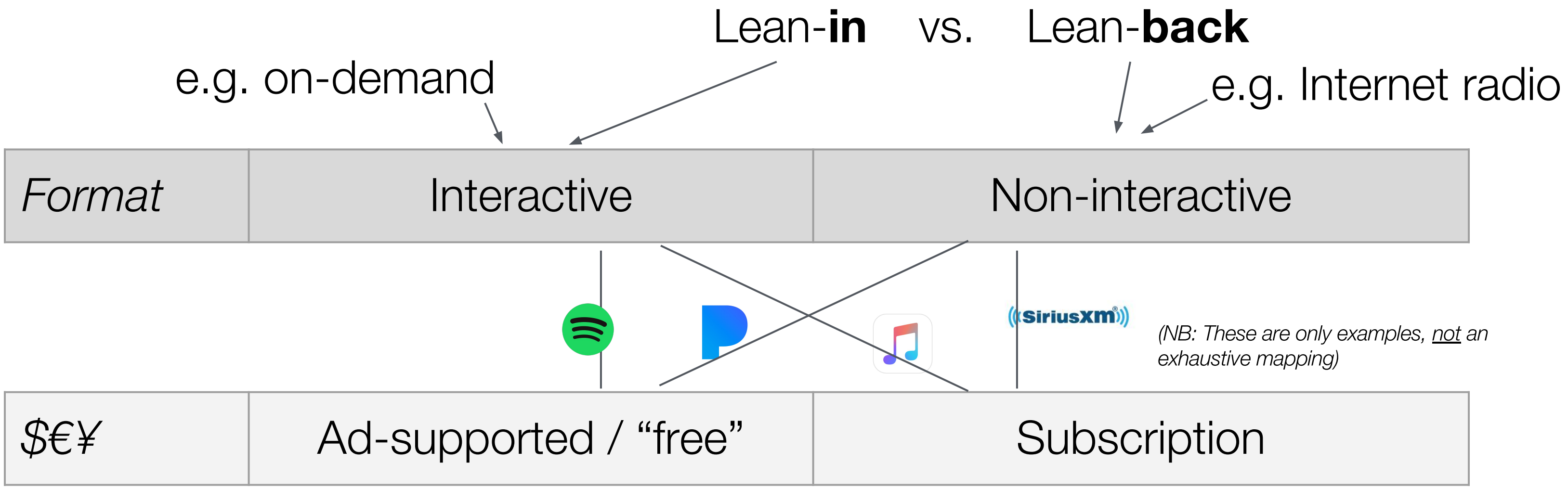
Not just a format transition, but a fundamental revolution.
Moving **away from ownership, towards access**.

→ Change of paradigm for RecSys: Recommending an **experience**, not just a product/item



Music Industry Changing Landscape

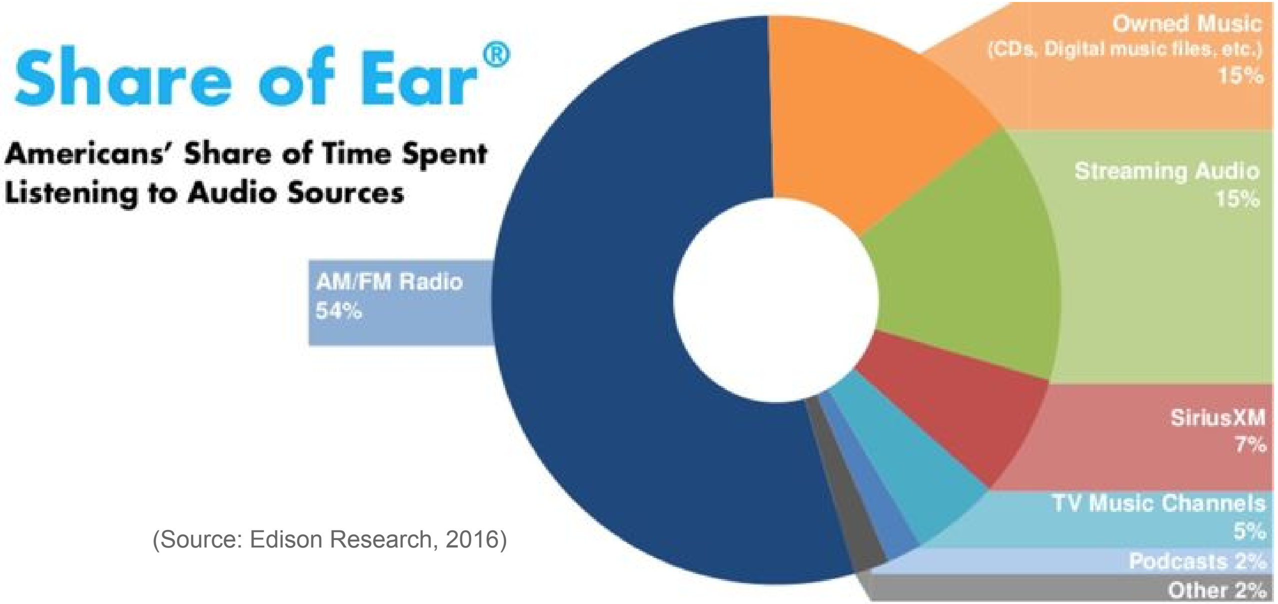
- **“Access”** can have different meanings
- New listening format still **not well-defined**… The (battle-)field is wide open.
- Lots of recent developments



→ Potential RecSys impact

Music Industry Changing Landscape

Looking at where $€¥ comes from is not the full picture…
… <u>time spent listening</u>, by media, tells a different story:



(Source: Edison Research, 2016)

Music Industry Changing Landscape

- Streaming “taking over” physical & downloads
- But competing with AM/FM radio, too

The Quest for “Discovery”

Ongoing quest for defining listening format calls for:

- Innovative Discovery features
- <u>Right balance between lean-in & lean-back experiences</u>

What makes music recommendation special?

- Duration of items (3+ vs. 90+ minutes in movies)
  → lower commitment necessary, items more “disposable”,
  bad recommendations maybe not as severe
- Sequential consumption
- Re-recommendation may be appreciated (in contrast to movies, TV shows)
- Often consumed passively (while working, background music, etc.)
- Different consumption locations/settings: static (e.g., via stereo at home) vs. variable (e.g., via headphones during exercise)
- Listening intent and context are crucial

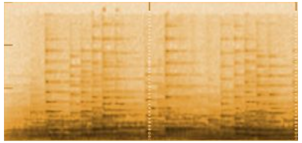
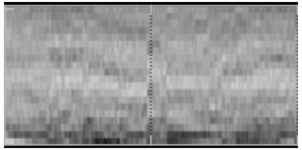
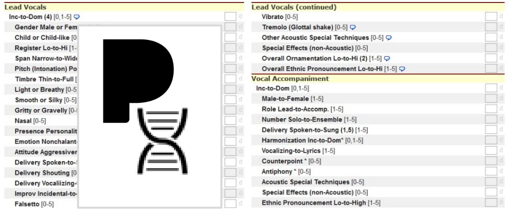
What makes music recommendation special?

- Importance of social component
- Highly emotionally connoted (in contrast to products, e.g. home appliances)
- Music often used for self-expression
- Various actors for recommendations (listeners, producers, performers, etc.)
- Various types of items (songs, albums, artists, audio samples, concerts, venues, fans, etc.)
- Magnitude of available data items/catalogs

Lots of Data and Data Sources

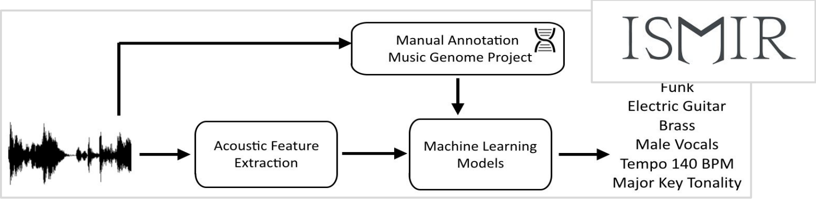
Content (audio, symbolic, lyrics)

- Machine listening/content analysis
- Human labelling

Meta-data

- Editorial
- Curatorial
- Multi-modal
  (e.g., album covers, booklets)

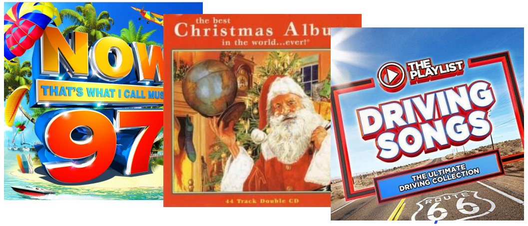


Lots of Data and Data Sources

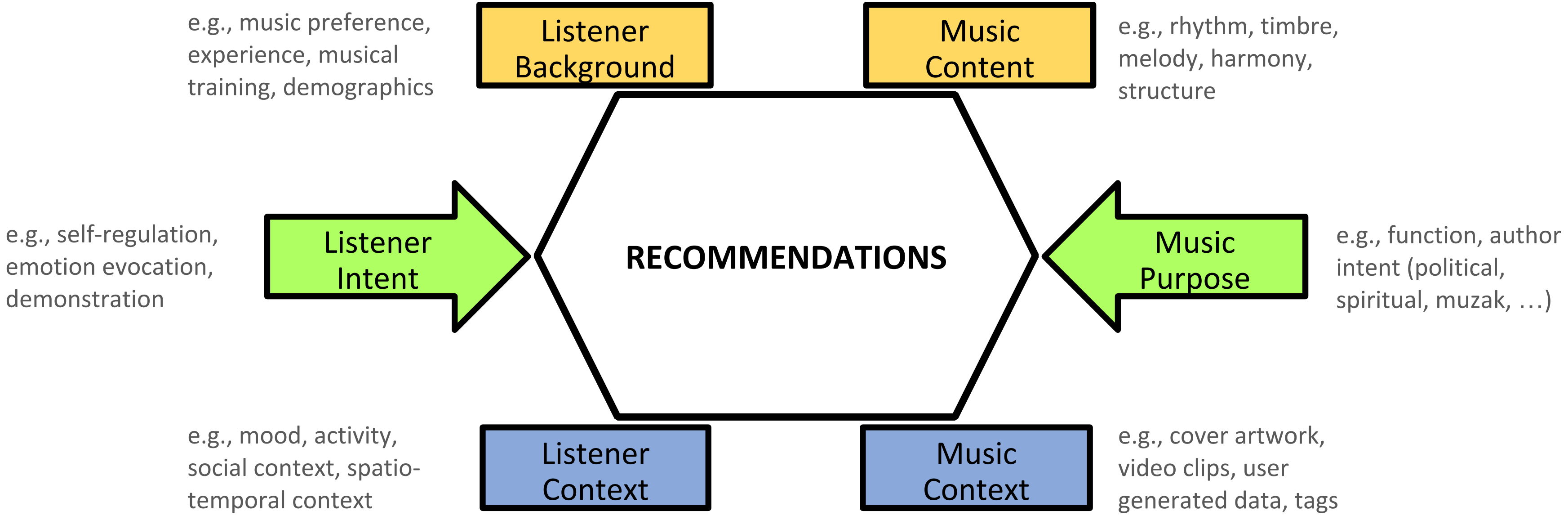
User-generated
- "Community meta-data"
- e.g., tags, reviews

Interaction Data
- Listening logs/shared listening histories
- Feedback ("thumbs")
- Purchases

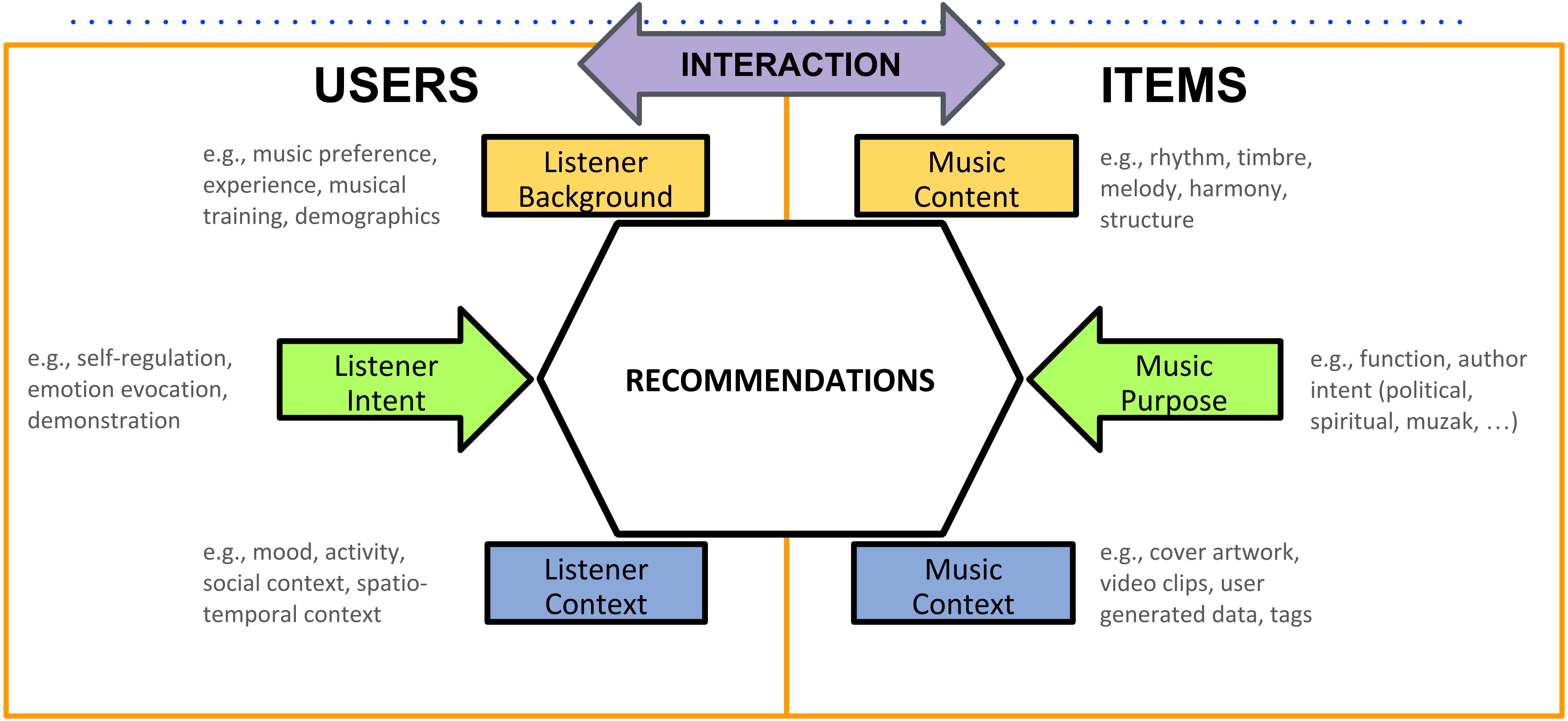
Curated collections
- Playlists, radio channels
- CD album compilations



Factors Hidden in the Data



Factors Hidden in the Data



Audio Content Analysis

- In contrast to e.g., movies: **true content-based recommendation!**
- Features can be extracted from any audio file
  → no other data or community necessary
  → no cultural biases (no popularity bias, no subjective ratings etc.)
- Learning of high-level semantic descriptors from low-level features via machine learning
- Deep learning becoming more relevant
  (representation learning and temporal modeling → CNNs, RNNs)

[Casey et al., 2008] *Content-based music information retrieval: Current directions and future challenges,* Proc IEEE 96 (4).

[Müller, 2015] *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*, Springer.

Audio Content Analysis: Selected Features

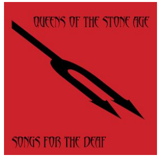
Disturbed
The Sound of Silence

Sound example

- Beat/downbeat → Tempo: 85 bpm (*madmom*)

Sound example

- Timbre (→ MFCCs)
e.g. for genre classification,
"more-of-this" recommendations



- Tonal features (→ Pitch-class profiles)
e.g. for melody extraction (*Essentia*),
cover version identification

Sound example

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 (*Essentia*):
not_danceable, gender_male, mood_not_happy

Toolboxes for Music Content Analysis

- **Essentia** (C++, Python): http://essentia.upf.edu
- **Librosa** (Python): https://github.com/librosa
- **Madmom** (Python): https://github.com/CPJKU/madmom
- **Marsyas** (C++): http://marsyas.info
- **MIRtoolbox** (MATLAB): https://www.jyu.fi/hytk/fi/laitokset/mutku/en/research/materials/mirtoolbox
- **jMIR** (Java): http://jmir.sourceforge.net
- **Sonic Visualiser** (MIR through VAMP plugins): http://sonicvisualiser.org

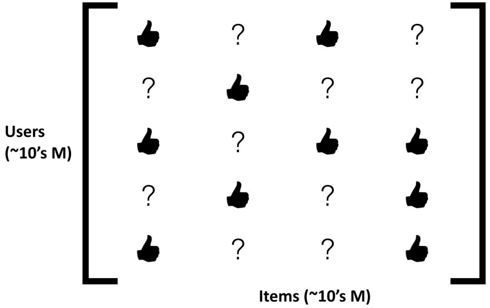
Text Analysis Methods (Basic IR)

- Text-processing of **user-generated content** and **lyrics**
  → captures aspects beyond pure audio signal
  → no audio file necessary
- Transform the content similarity task into a text similarity task (cf. “content-based” movie recommendation)
- Allows to use the full armory of text IR methods, e.g.,
  - Bag-of-words, Vector Space Model, TFIDF
  - Topic models, word2vec
- Example applications: Tag-based similarity, sentiment analysis (e.g., on reviews), mood detection in lyrics

[Knees and Schedl, 2013] *A Survey of Music Similarity and Recommendation from Music Context Data,* Transactions on Multimedia Computing, Communications, and Applications 10(1).

Collaborative Filtering

- Exploiting **interaction data**, stemming from “usage” of music
  → usually closer to “what users want”



- Implicit (e.g. plays) or explicit data (e.g. thumbs)
- Task: completion of user-item matrix
- Learning latent factors and biases
  cf. [Koenigstein et al. 2011]

$$b_{ui} = \mu + b_{u,type(i)} + b_{u,session(i,u)} + b_i + b_{album(i)} + b_{artist(i)} + \frac{1}{|genres(i)|} \sum_{g \in genres(i)} b_g + c_i^T f(t_{ui})$$

- Special treatment of implicit data (*preference* vs. *confidence*), re-recommendation, etc.

[Hu et al., 2008] *Collaborative Filtering for Implicit Feedback Datasets*, ICDM.

[Slaney, 2011] *Web-Scale Multimedia Analysis: Does Content Matter?*, IEEE MultiMedia 18(2).

[Koenigstein et al., 2011] *Yahoo! music recommendations: modeling music ratings with temporal dynamics and item taxonomy*, RecSys.

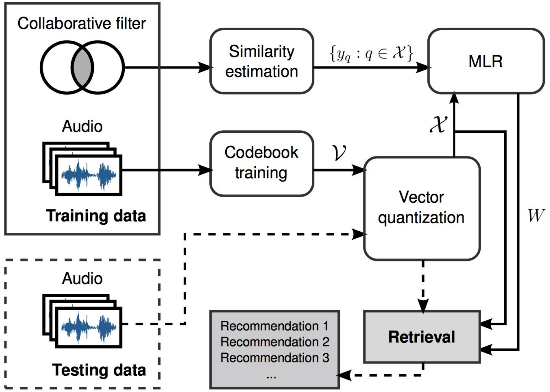
Feedback-Transformed Content

Interaction model as target for learning features from audio

- Dealing with cold-start
- Personalizing the mixture of content features

E.g.,

- Learning item-based CF similarity function from audio features using metric learning



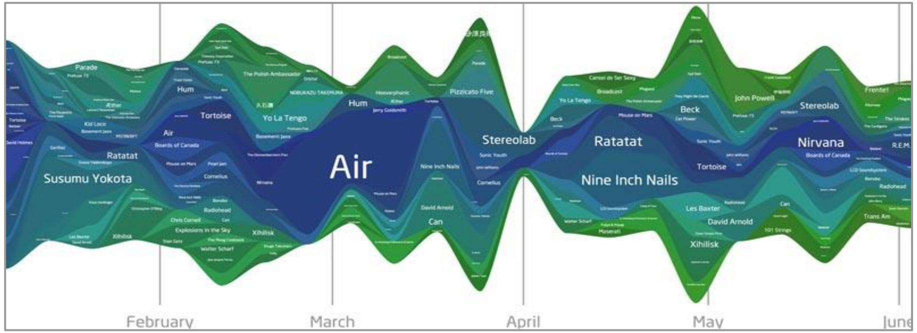
[McFee et al., 2012] *Learning Content Similarity for Music Recommendation.* IEEE TASLP 20(8).

- Learning latent item features using weighted matrix factorization
- Convolutional neural network with mel-spectrogram as input and latent item vectors as target

[van den Oord et al., 2013] *Deep Content-Based Music Recommendation*. NIPS workshop.

Sequence Mining

- Aims at modelling user preference + finer-grained session context (≈user context+user intent)
- User context should be reflected in <u>selected sequence of songs</u>



- Model (hyper-)graph, latent factors, or topic models (e.g. LDA) on tags over **listening histories** and **playlists**
  → “session model”, “playlist dialect”, etc.

[Zheleva et al., 2010] *Statistical models of music-listening sessions in social media*. WWW.

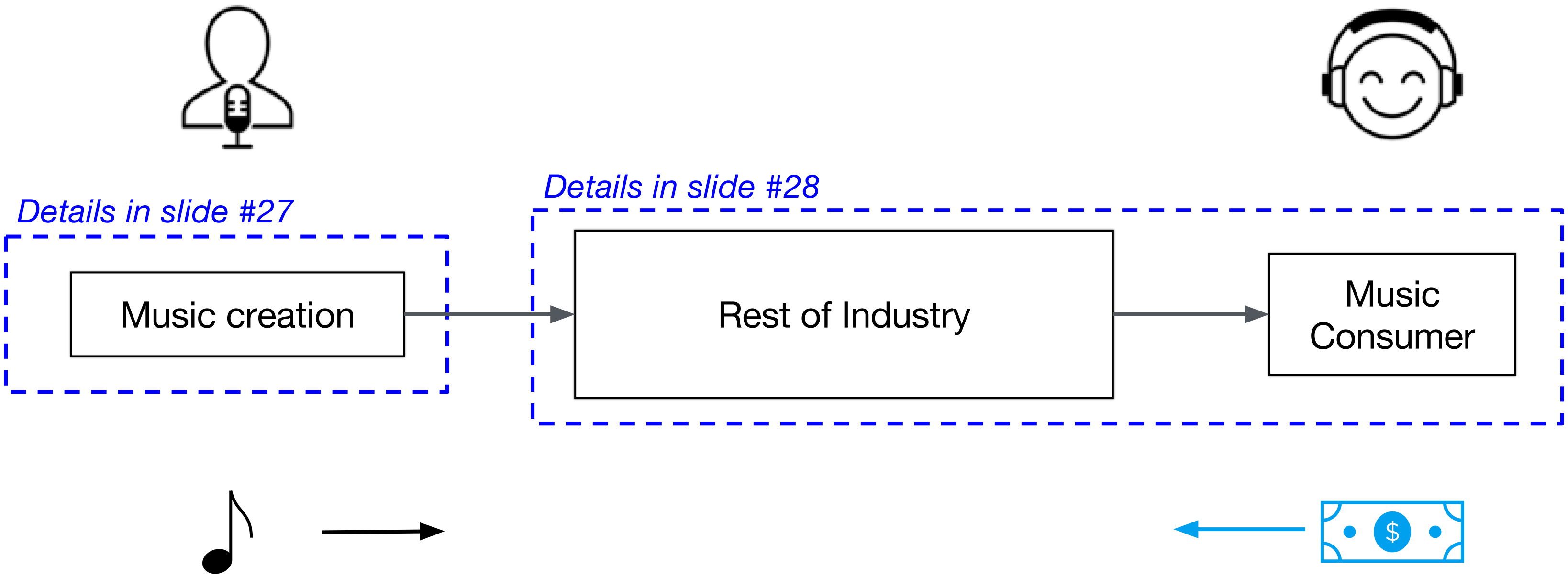
[Hariri et al., 2012] *Context-aware music recommendation based on latent topic sequential patterns*, RecSys.

[Aizenberg et al., 2012] *Build your own music recommender by modeling internet radio streams*. WWW.

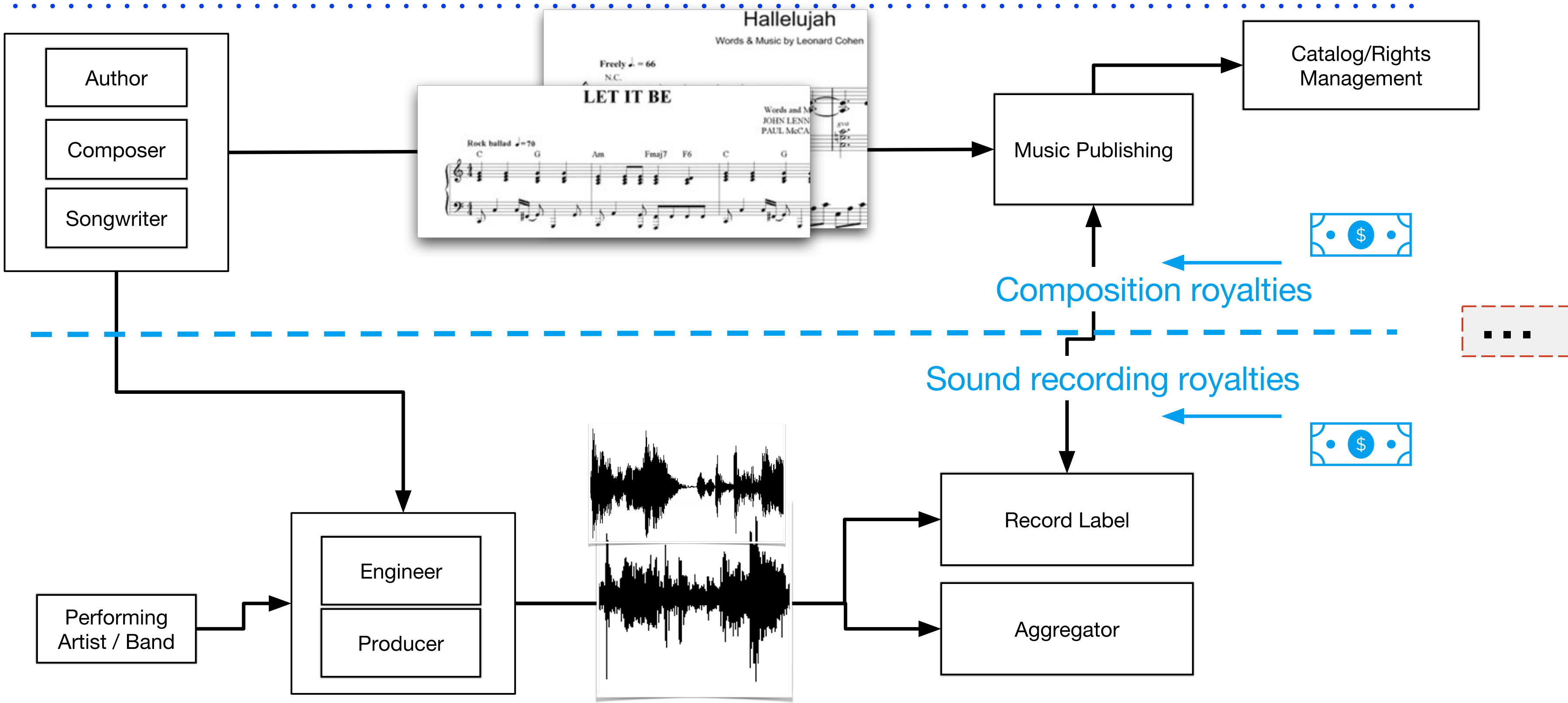
[McFee, Lanckriet, 2012] *Hypergraph Models of Playlist Dialects*, ISMIR.

It’s All About the Use Case

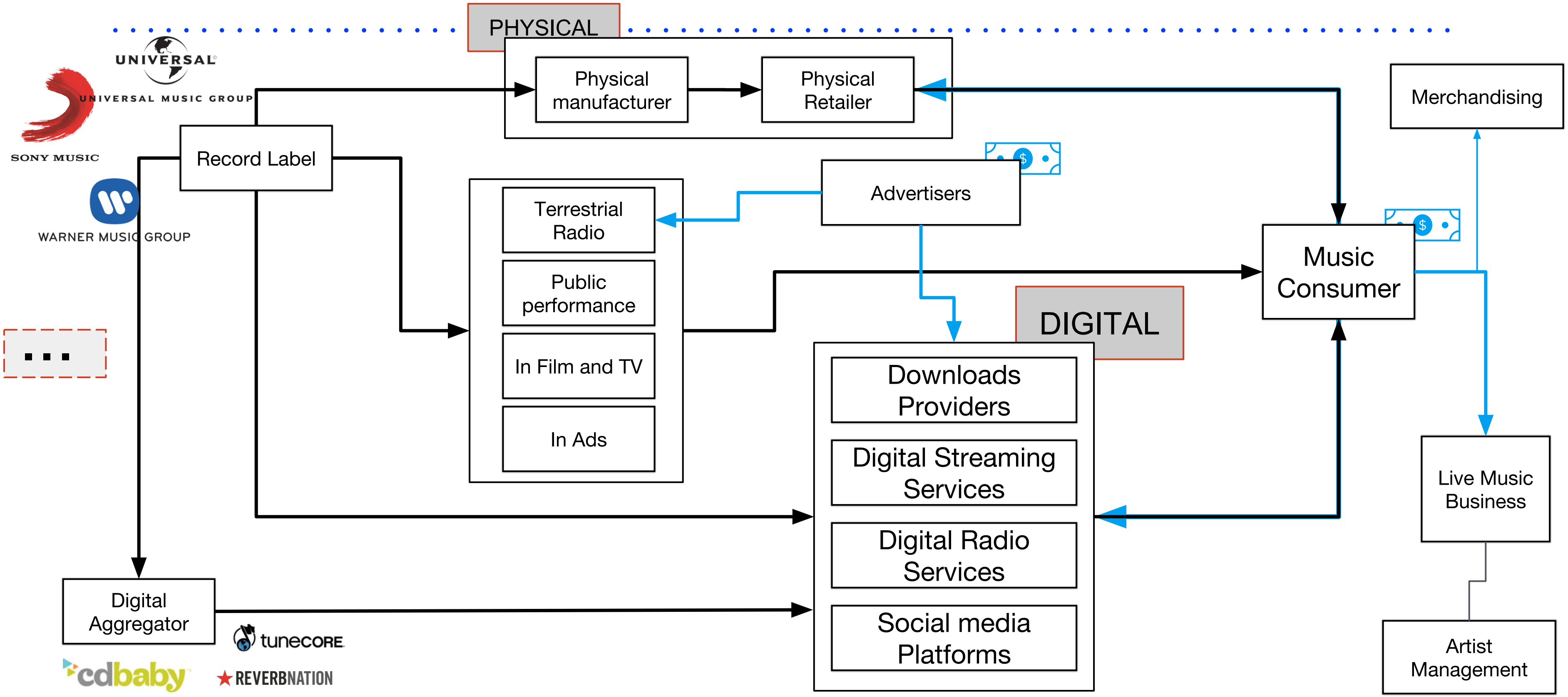
You said "Music Industry Landscape"?



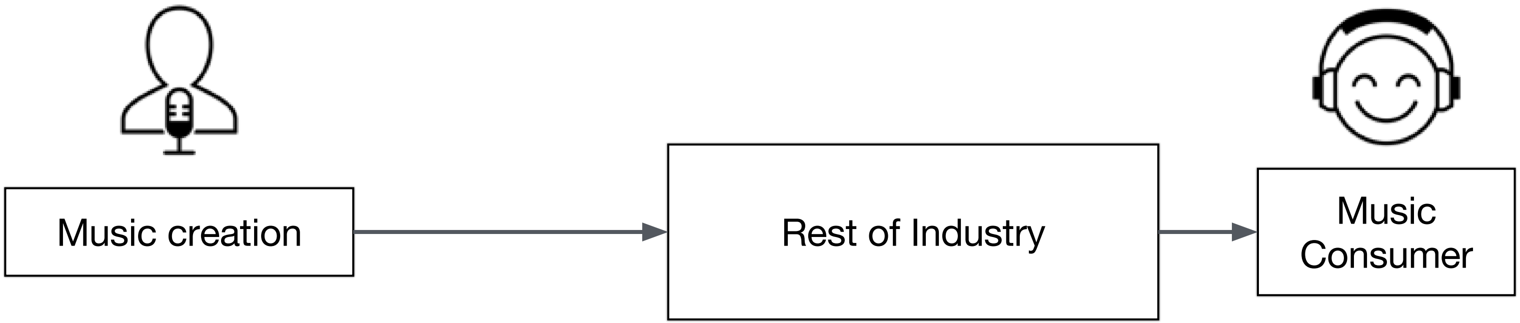
Music Industry Landscape

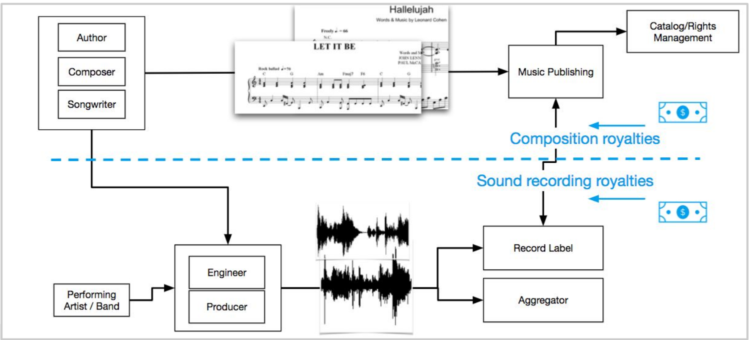


Music Industry Landscape

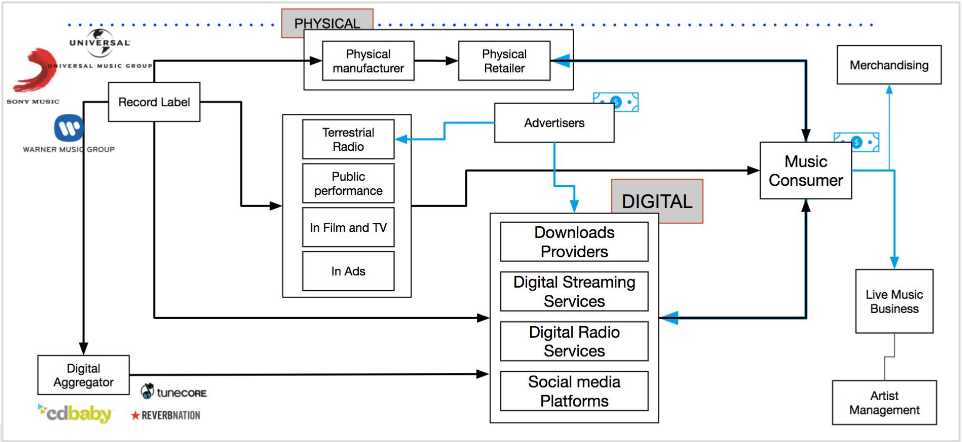


Music Industry Landscape (again)



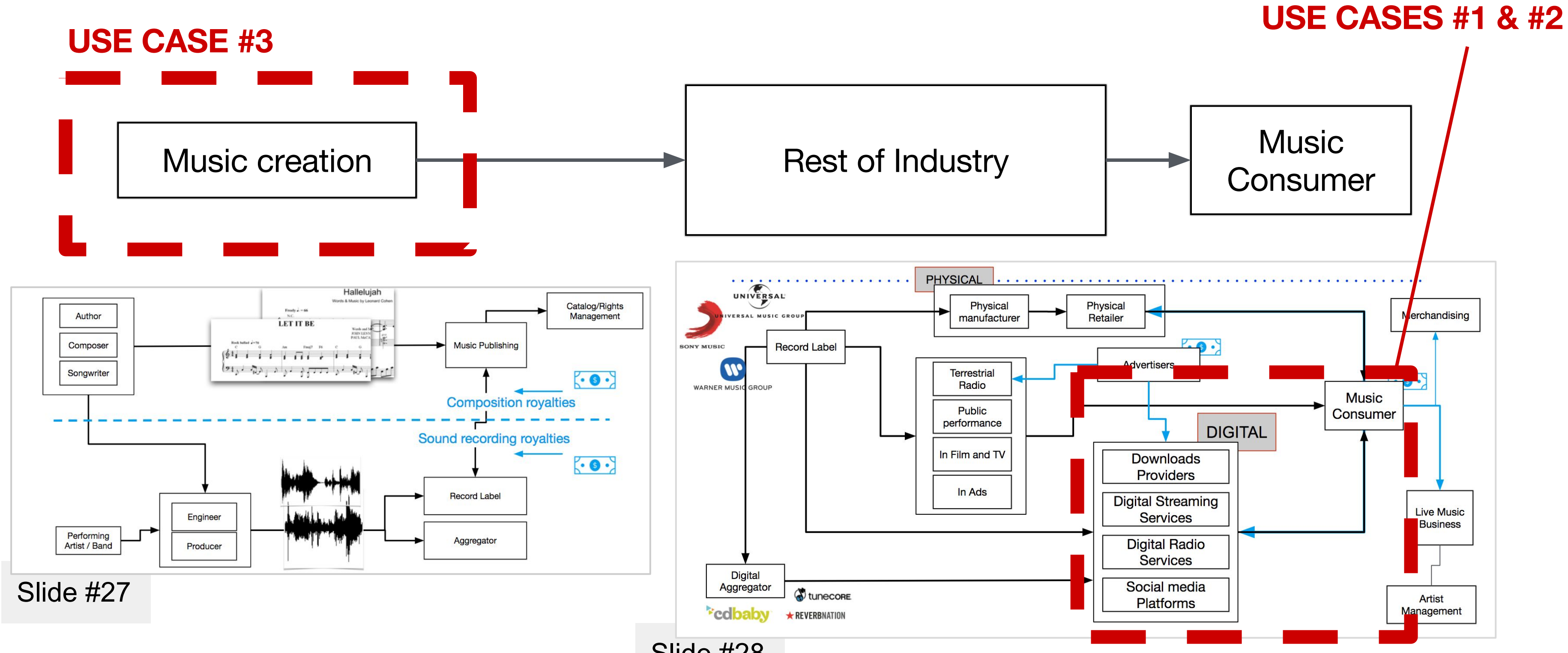


Slide #27



Slide #28

Music Industry Landscape (again)



Overview (again)

- Introduction to Music Recommendation
- It’s All About the Use Cases
- **Use Case 1: Station/Playlist Generation**
- **Use Case 2: Context-Aware Music Recommendation**
- **Use Case 3: Recommendation in the Creative Process of Music Making**
- What’s Next?

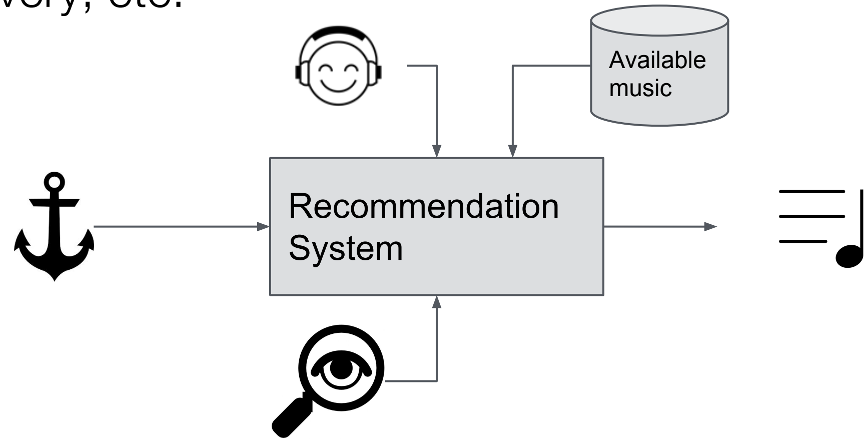
Use Case 1: Station/Playlist Generation

Station/Playlist generation problem

- A continuation problem
- Given a listener enjoying a particular musical experience, what track recommendations can we make to extend this experience as much as possible

A particular recommendation problem

- *The problem*: Given a listener, a set of available tracks to play, a musical "anchor", and a particular focus, **recommend best next tracks**
    - ***Musical anchor***: i.e. current music listening experience defined by e.g. a radio station, a set of tracks (e.g. a playlist, an album), a given artist, a genre, etc.
    - ***Focus / Listener intent***: lean-in vs. lean-back, new music, (re)discovery, etc.



Station/Playlist generation - Differences

Station:

- Anchored in a track, an artist, an album, a genre, etc.
- Recommendations: Sequential, 1 track after the other. Possibly hidden.
- Learning:
  - Learning data: Feedback (lots of), user-generated data (little)
  - System is the oracle, then adapts to feedback (must be real-time)

Playlist:

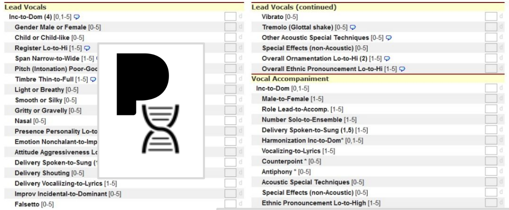
- Anchored in an arbitrary (finite) length set of tracks, either:
  - User-generated
  - Curated (e.g. by streaming service, 3rd-parties)
- Recommendations: In batch
- Learning:
  - Learning data: User-generated data, feedback

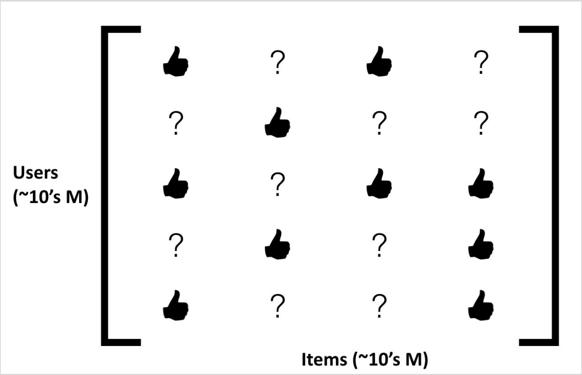
Data and Recommendation algorithms

Types of algorithms / approaches:

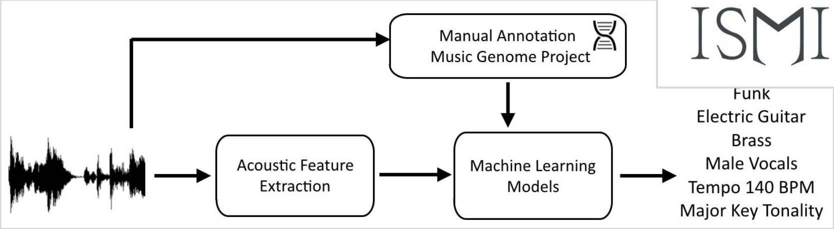
- Editorial
- Curatorial
- Collaborative filtering
- Content-based
  - Human labelling
  - Machine listening
- Personalized filtering

UNIVERSAL MUSIC GROUP · SONY MUSIC · WARNER MUSIC GROUP · DDEX · rovi · Discogs · WIKIPEDIA The Free Encyclopedia

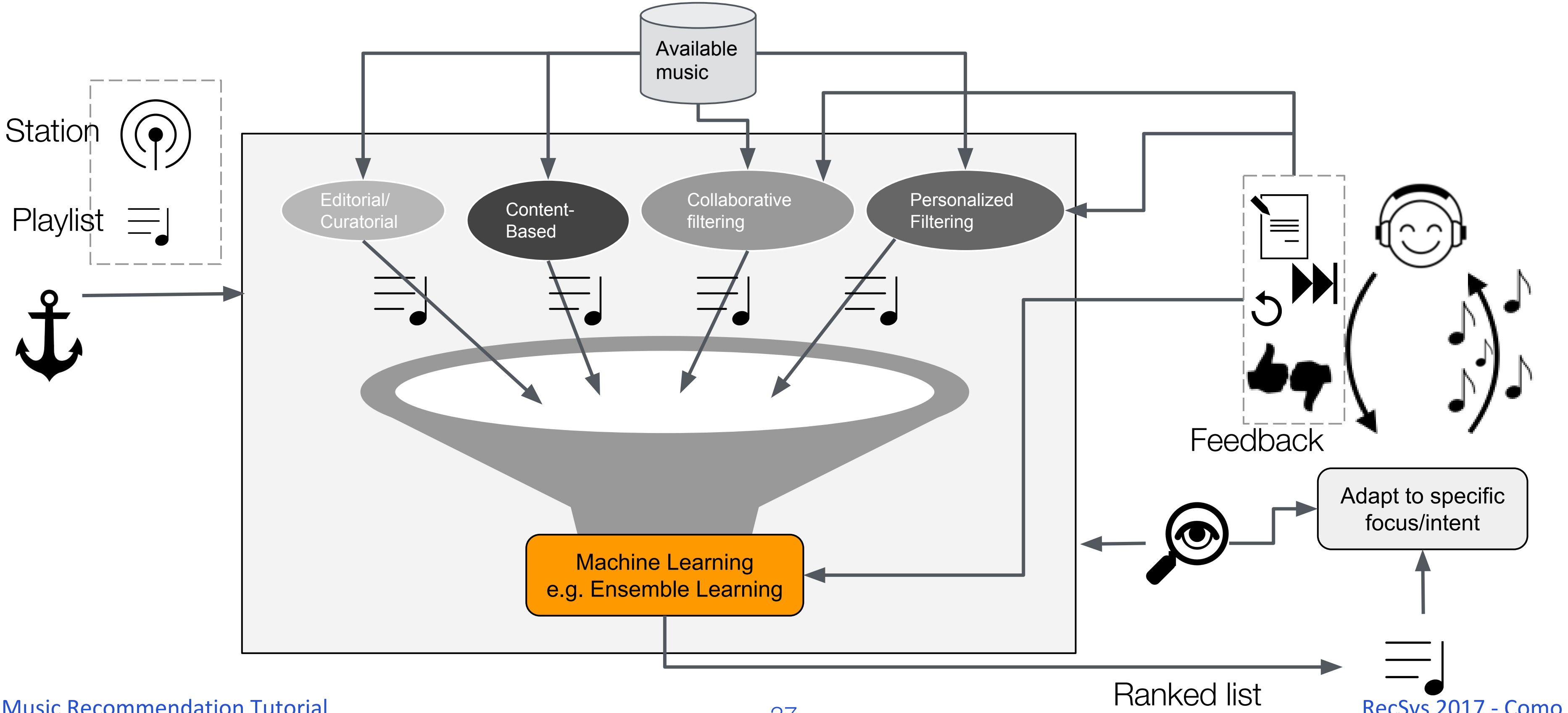




ISMIR



Recommendation pipeline



Temporal aspect

- *"Recommending next tracks"*... Temporal ordering matters
- Notion of "music rotation" from AM/FM radio programming, e.g.:
  - Popularity categories: "Current", "Recurrent", "Gold"
  - Musical attributes: tempo, male vs. female vocals, danceability, major vs. minor, etc.
  - Sound attributes: synth vs. acoustic, intensity, etc.
  - Artist separation

[Price, 2015] *After Zane Lowe: Five More Things Internet Radio Should Steal from Broadcast,* NewSlangMedia blog post

- Predict best time for next user interaction with an item

[Dai, Wang, Trivedi, Song, 2016] *Recurrent Coevolutionary Latent Feature Processes for Continuous-Time User-Item Interactions*, Workshop on Deep Learning for Recommender Systems @ RecSys

- Modelling transitions in listening habits (e.g. artist transitions)

[Figueiredo, Ribeiro, Almeida, Andrade, Faloutsos, 2016] *Mining Online Music Listening Trajectories*, ISMIR

[McFee, Lanckriet, 2012] *Hypergraph Models of Playlist Dialects*, ISMIR

[Bonnin, Jannach, 2014] *Automated Generation of Music Playlists: Survey and Experiments*, ACM Computing Surveys

A “good” recommendation?

What makes a good recommendation:

- Accuracy
- Good balance of:
  - Novelty vs. familiarity / popularity
  - Diversity vs. similarity
- Transparency / Interpretability
- Listener Context

Influential <u>factors</u>:

- Listener
- Musical anchor
- Focus / Intent

Remember: It’s about recommending an <u>experience</u>

[Celma, 2010] *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space*, Springer

[Celma, Lamere, 2011] *Music Recommendation and Discovery Revisited*, ACM Conference on Recommender Systems

[Jannach, Adomavicius, 2016] *Recommendations with a Purpose*, RecSys

[Amatriain, Basilico, 2016] *Past, Present, and Future of Recommender Systems: An Industry Perspective*, RecSys

Accuracy (is not enough)

- Typically, recommendations are based on predicting the relevance of unseen items to users. Or on item ranking.
- For recommendations to be accurate, optimize recommender to best predict <u>general</u> relevance
  - e.g. learning from historical data from <u>all</u> *users*

- Too much focus on accuracy → biases (i.e. **popularity** and **similarity** biases)
  - Tradeoff popularity vs. personalization (is pleasing both general user base *and* each individual even possible?...)
  - Particular risk of selection bias when recsys is the oracle (e.g. station)
  - Single-metric Netflix Prize (RMSE) → only one side of the coin

[Jannach, Kamehkhosh, Bonnin, 2016] *Biases in Automated Music Playlist Generation: A Comparison of Next-Track Recommending Techniques*, UMAP

Novelty

- Introducing novelty to balance against popularity (or familiarity) bias
- Both are key: Listeners want to hear what's hype (or what they already know). But they also need their dose of novelty... Once in a while.
  - How far novel? ("correct" dose?)
  - How often?
  - When?, etc...

	"Yep, novelty's fine"	*"No novelty, please!"*
Listener	Jazz musician	My mother
Musical anchor	Exploring a new friend's music library	Playlist for an official high-stake dinner
Focus	Discovery	Craving for my hyper-personalized stuff

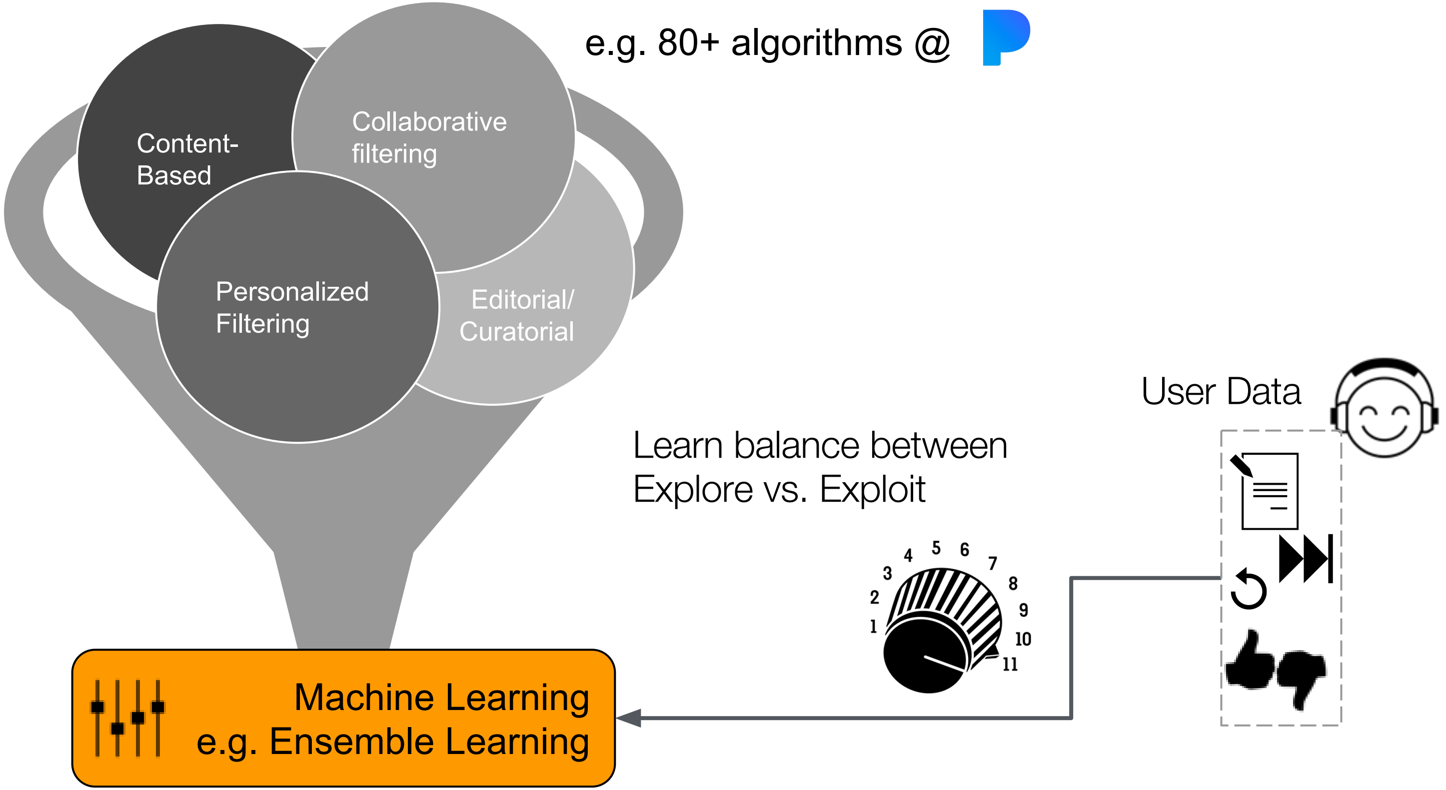
Diversity

- Introducing diversity to balance against similarity bias
- Similarity ≅ accuracy
  - Trade-off accuracy vs. diversity
  - As for Novelty, adding Diversity is a useful means for personalizing and contextualizing recommendations

	"Yep, bring on diversity"	*"No diversity, please!"*
Listener	A (good) DJ	Exclusive Metal-head
Musical anchor	Station anchored on "90's & 00's Hits"	Self-made playlist anchored on "Slayer"
Focus	Re-discovery, hyper-personalized	"Women in Post-Black Metal"

[Parambath, Usunier, Grandvalet, 2016] *A Coverage-Based Approach to Recommendation Diversity on Similarity Graph*, RecSys

Exploration vs. Exploitation



Exploration vs. Exploitation

- Exploit:
  - **Data** tells us what works best now, let's play exactly that
  - Play something **safe now**, don't worry about the future
  - Lean-back experience
  - "Don't play music I am not familiar with"

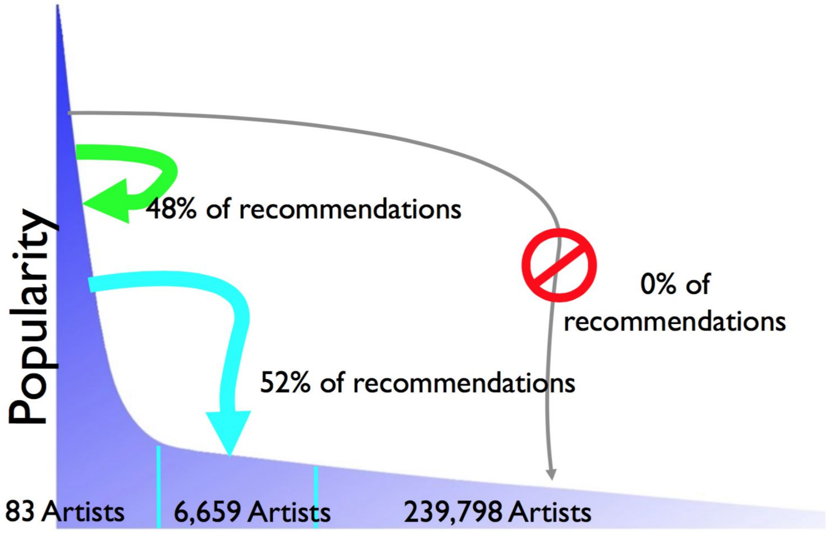
Short-term reward

Long-term reward

- Explore:
  - Let's **learn** (i.e. gather some more data points on) what **might** work
  - Play something **risky now**, preparing for tomorrow
  - Lean-in experience
  - "I'm ready to open up. Just don't play random stuff"

[Xing, Wang, Wang, 2014] *Enhancing Collaborative Filtering Music Recommendation by Balancing Exploration and Exploitation*, ISMIR

Exploration vs. Exploitation



(CF-based recommendations, Last.fm data)

Helps alleviate limited reach of some recsys:

- Coldplay, Drake, etc. vs. "Working-class" musicians (long-tail)
- Radio typically plays 10's artists per week
- Streaming has the potential to play 100k's artists per week
- Caveat of collaborative filtering-based algorithms

[Celma, 2010] *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space*, Springer

Transparency / Interpretability

- *“Why am I recommended this?”*

If you like Bernard Herrmann

You might like “Gimme some more” by Busta Rhymes

Transparency / Interpretability

- *“Why am I recommended this?”*

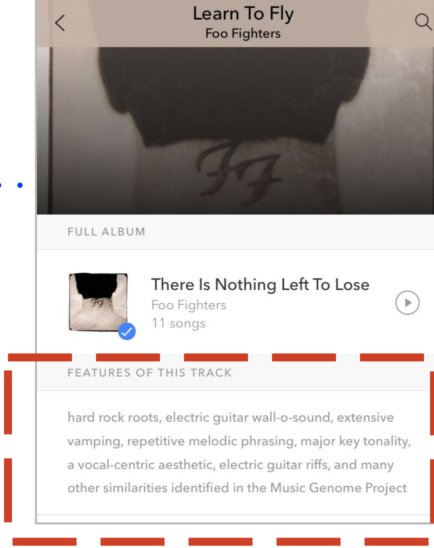
If you like Bernard Herrmann

You might like “Gimme some more” by Busta Rhymes

Because:
He sampled Herrmann’s work

Transparency / Interpretability

- Explain how the system works: Transparency
- Increases users’ confidence in the system: Trust
- Facilitates persuasion
- Fun factor → increases time spent listening
- Increases personalization (e.g. *“because you like guitar”*)
- Better experience overall
- Caveat: Users will then want to correct potentially erroneous assumptions
  → Extra level of interactivity needed



[Tintarev, Masthoff, 2015] *Explaining Recommendations: Design and Evaluation*, Recommender Systems Handbook (2nd ed.), Kantor, Ricci, Rokach, Shapira (eds), Springer

[Musto, Narducci, Lops, de Gemmis, Semeraro, 2016] *ExpLOD: A Framework for Explaining Recommendations based on the Linked Open Data Cloud*, RecSys

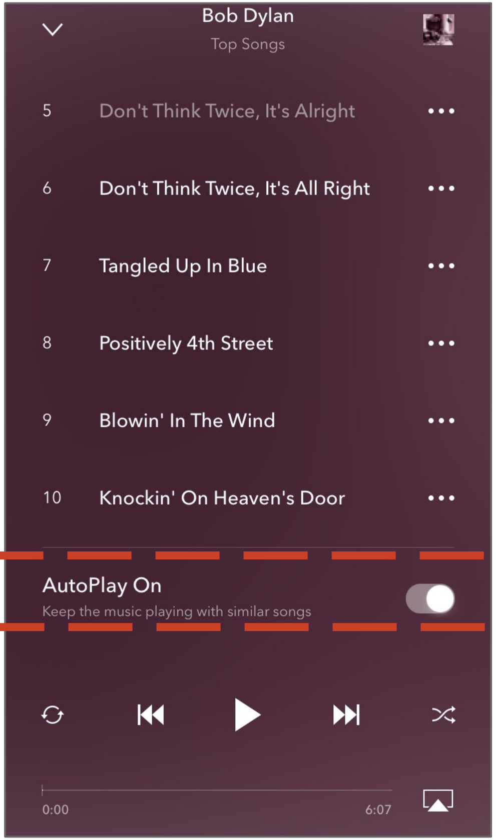
[Chang, Harper, Terveen, 2016] *Crowd-based Personalized Natural Language Explanations for Recommendations*, RecSys

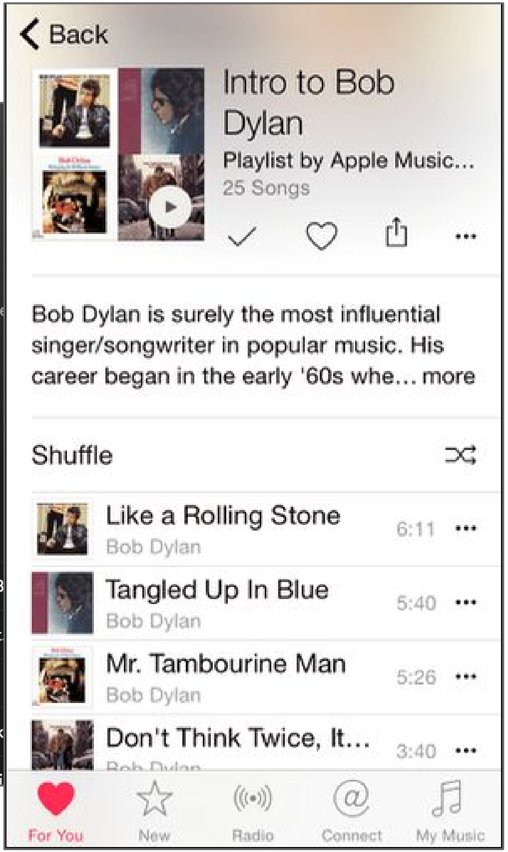
Listener Context

- Big picture: → *Context-Aware Music Recommendation (next Use Case)*

- **Explicit focus / listener intent**:
  - Focus on newly released music (new stuff)
  - Focus on discovery (new *for me*)
  - Re-discovery (throwback songs)
  - Focus on a particular listening experience (lean-in vs. lean-back)
  - Hyper-personalized (extreme lean-back, *my* best-of)
  - etc.
- Specific focus defines:
  - Which recommendations are best
  - Which **vehicle** for recommendations is best (**how** to recommend)

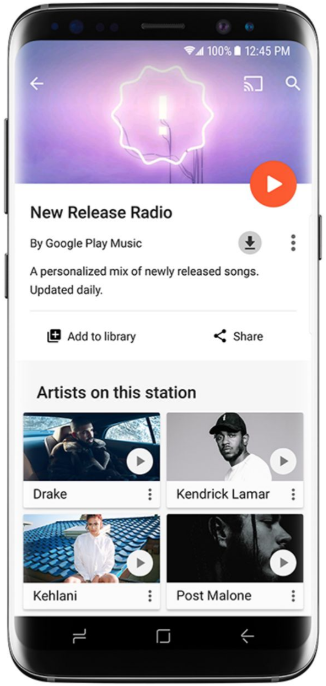
Focus on: Discovering an artist

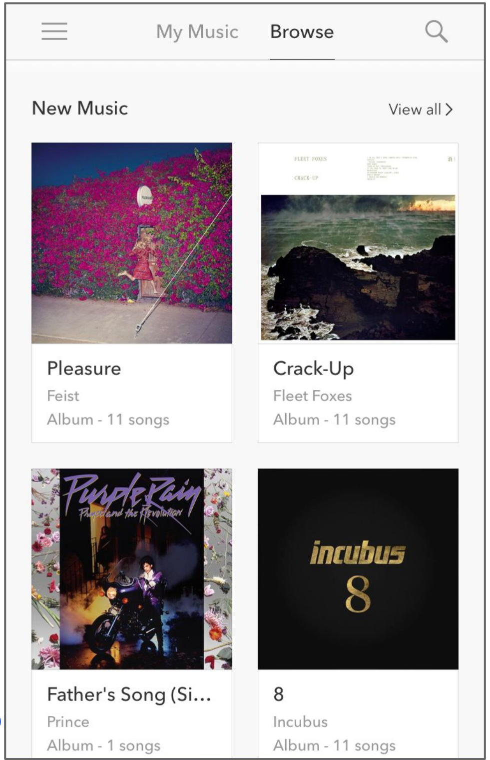


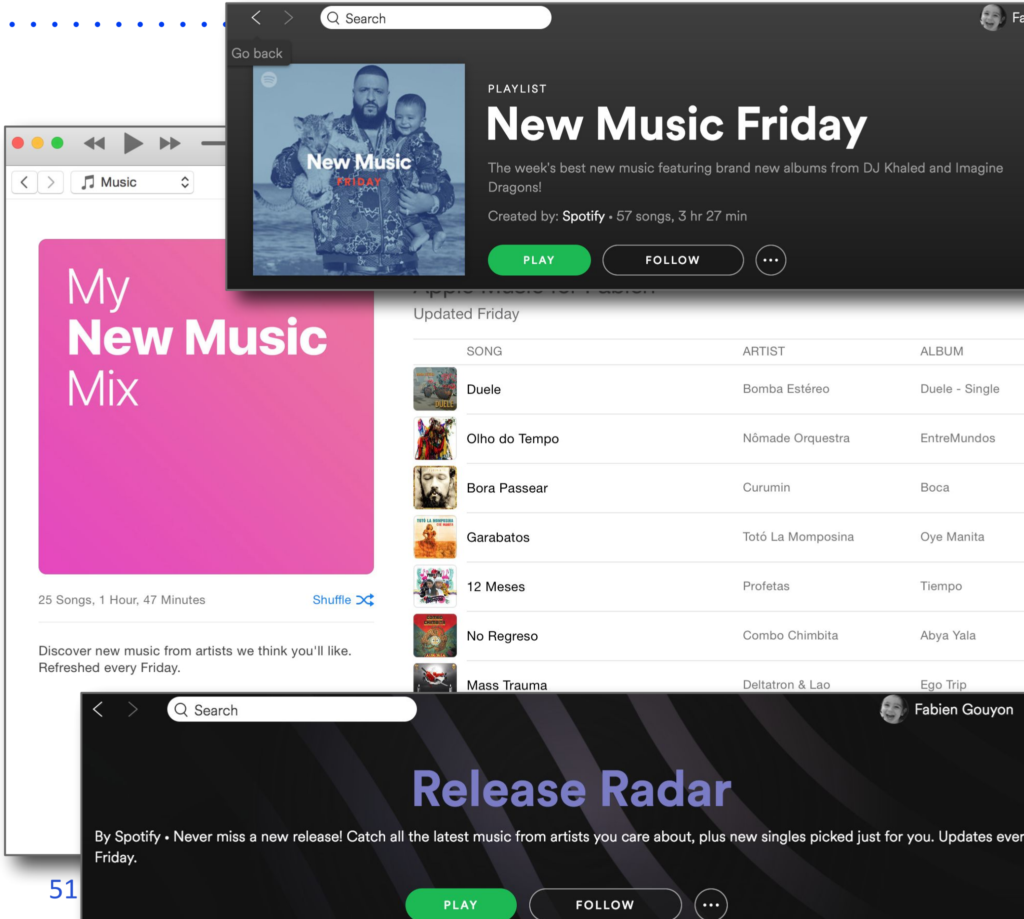


Focus on: New music

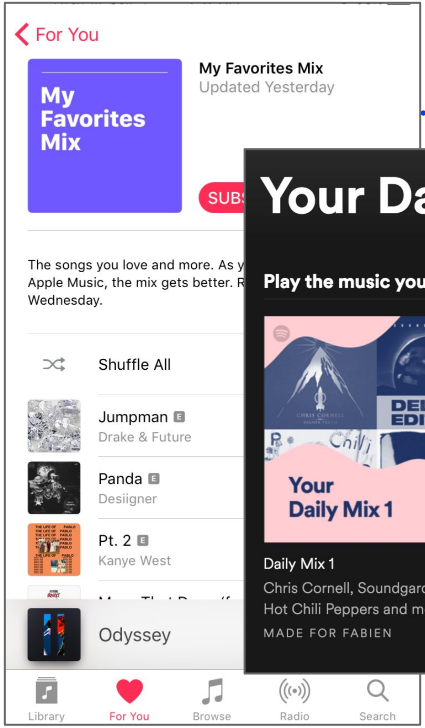
Personalized vs. non-personalized

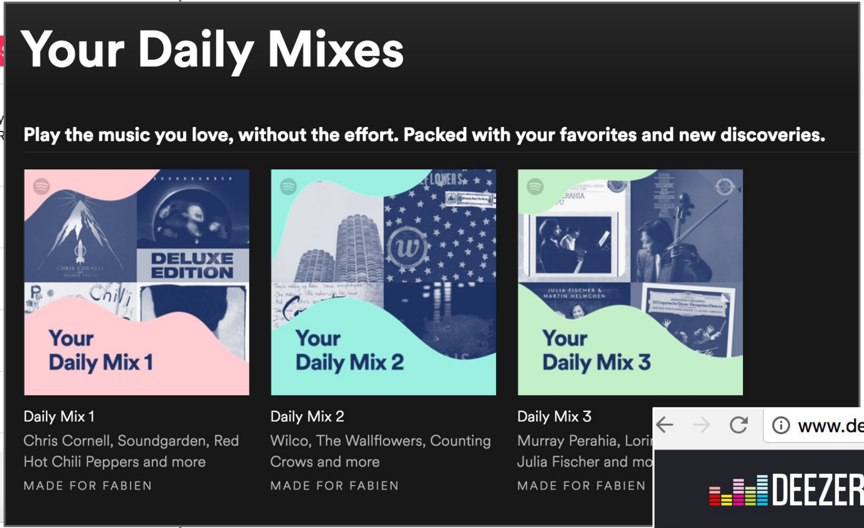


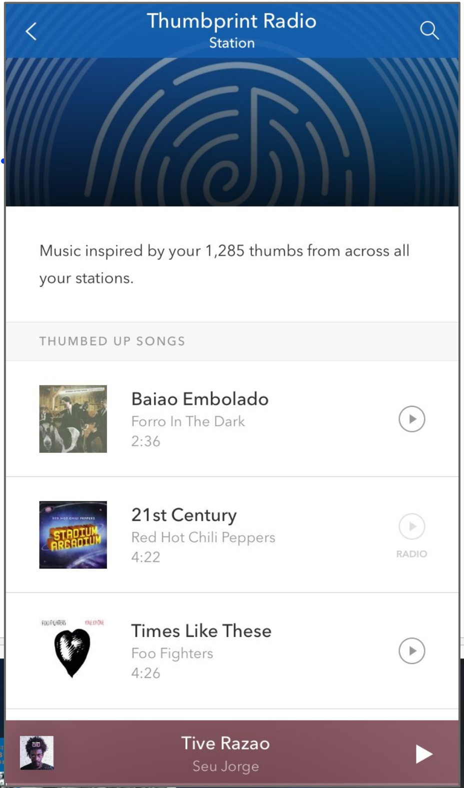


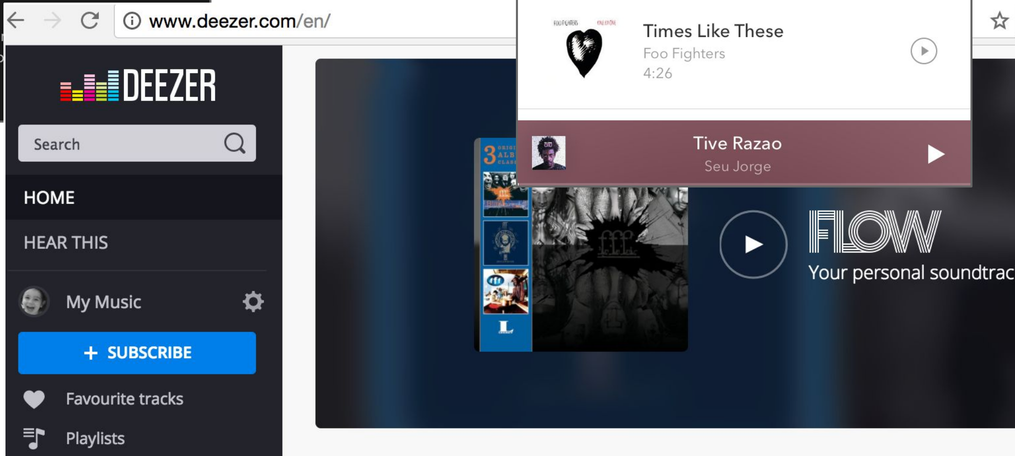


Focus on: Re-discovery



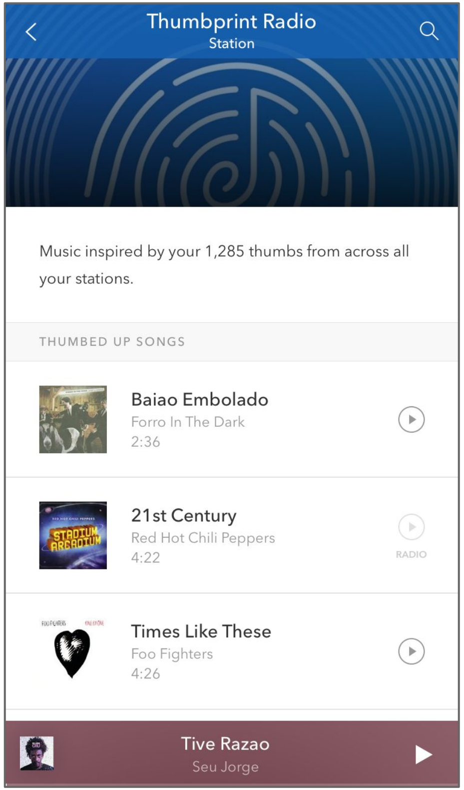


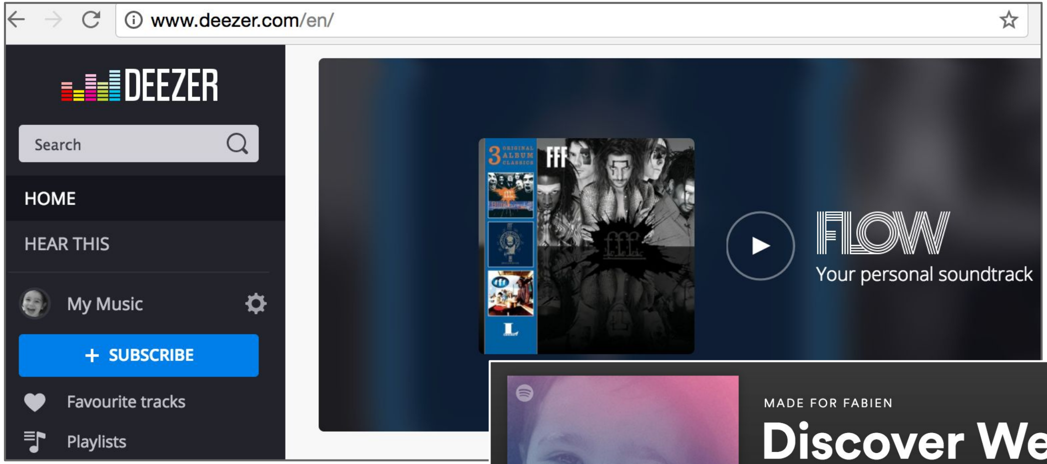


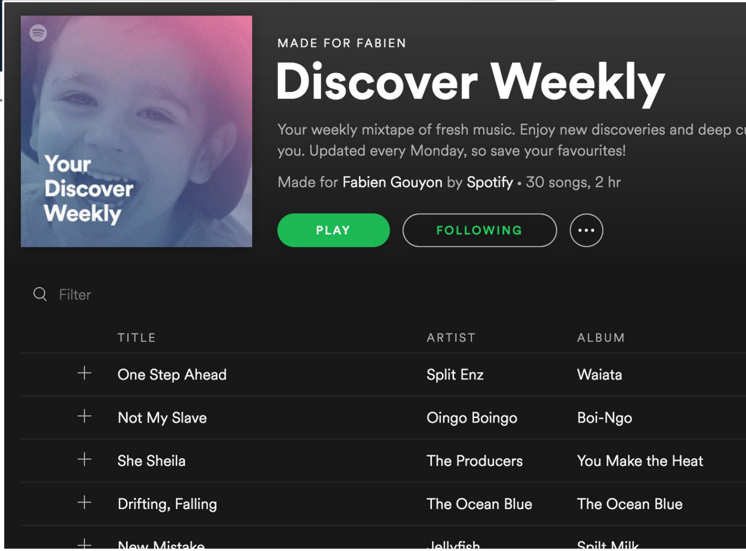


Focus on stuff you know you like
Personalized, leaning towards exploit

Focus on: Hyper-personalized Discovery





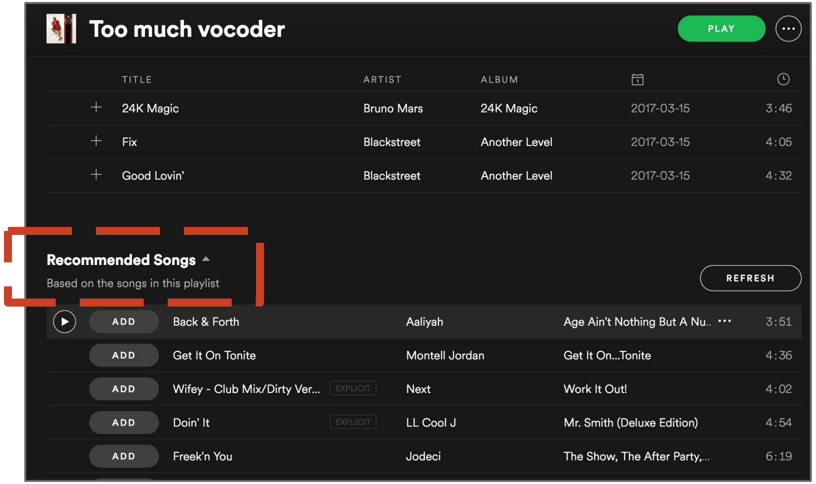


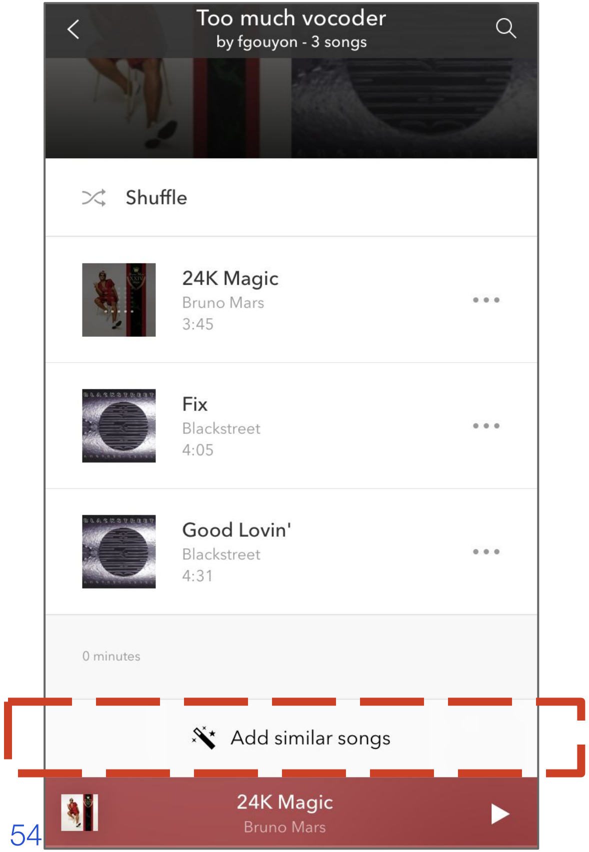
About discovering new stuff.
Intended to feel like it's curated. Just. For. Me.

Leaning towards explore

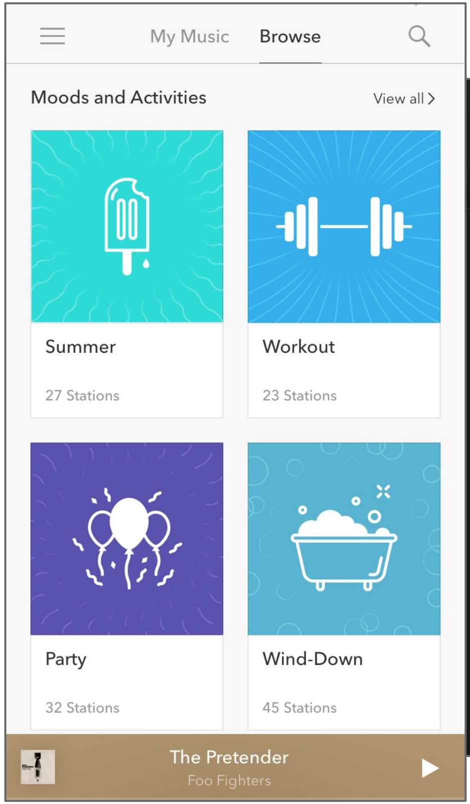
Focus on: Lean-in experience

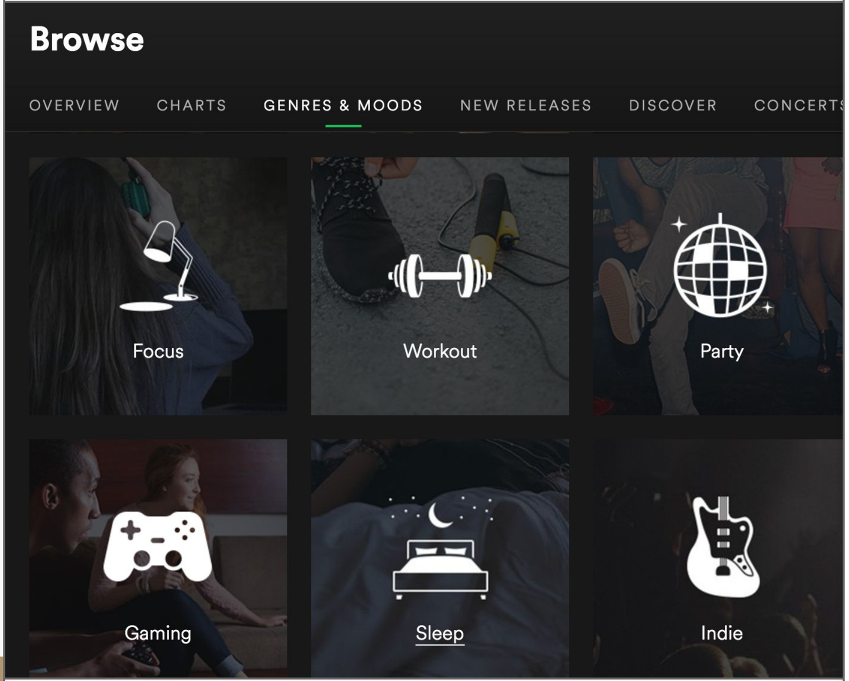
Lean in:
Building Playlists

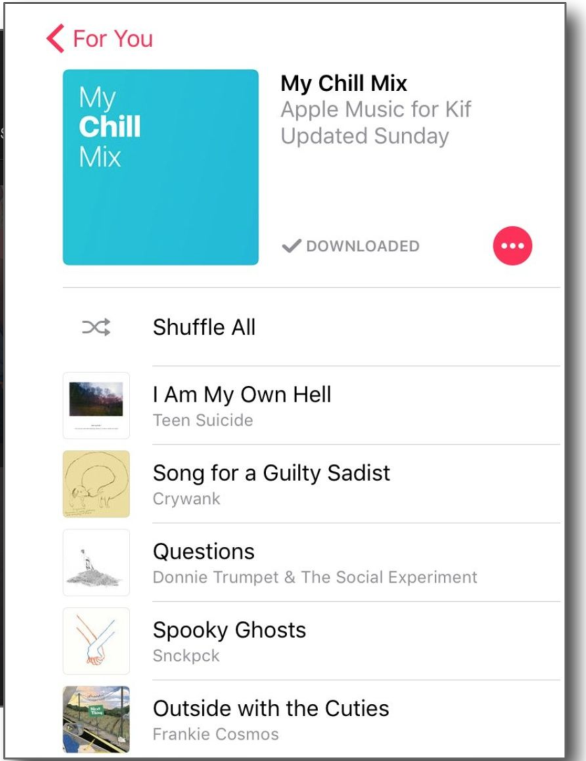




Focus on: Mood /Activity







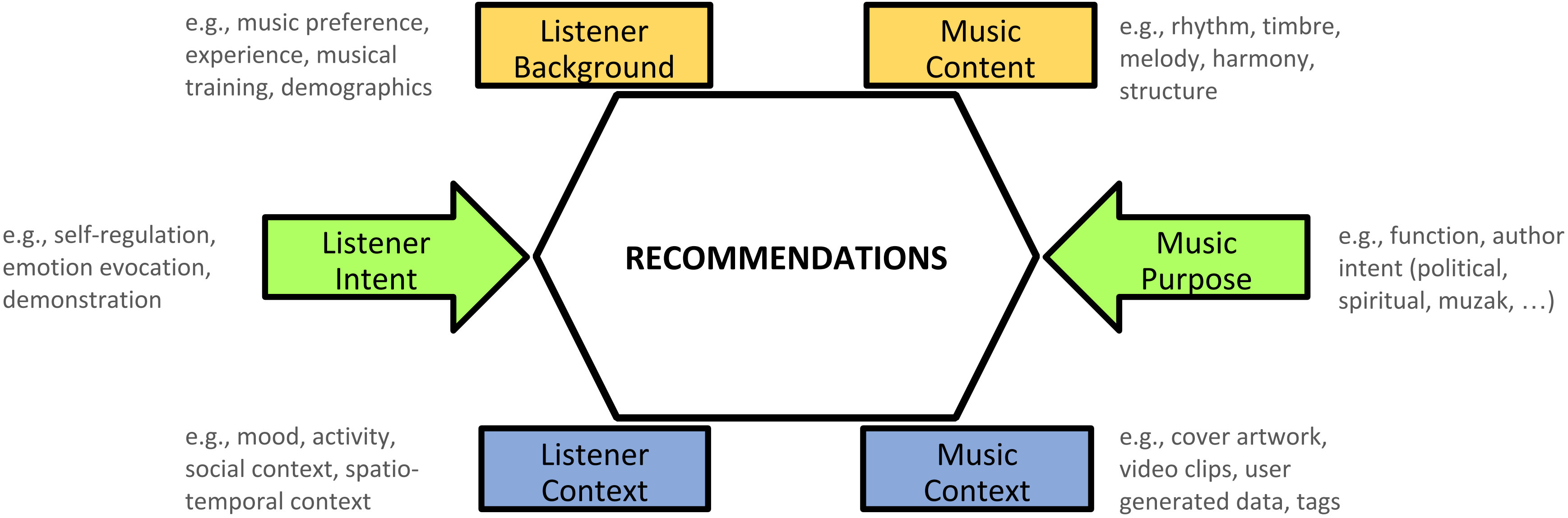
Personalized vs. non-personalized

Use Case 2:
Context-Aware Music Recommendation

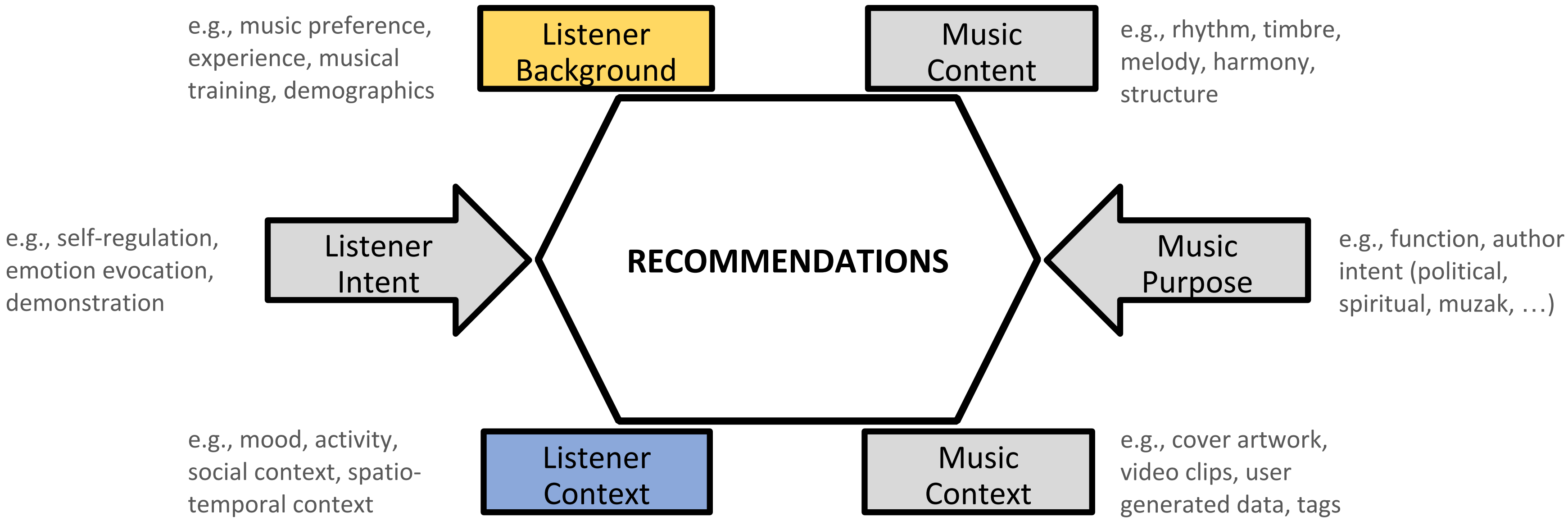
Overview

- **Context categories and acquisition**: We categorize various dimensions of the user context, e.g., time, location, activity, weather, social context, personality, etc.
- **Methods/examples**: We outline the most frequently adopted approaches in context-aware MRS.
- **Cultural/regional specificities**: We summarize findings about country-specific differences in music preferences.
- **Evaluation**: We highlight particular challenges in evaluating context-aware MRS.

Listening Hexagon



Listening Hexagon



Context categories

Environment-related context

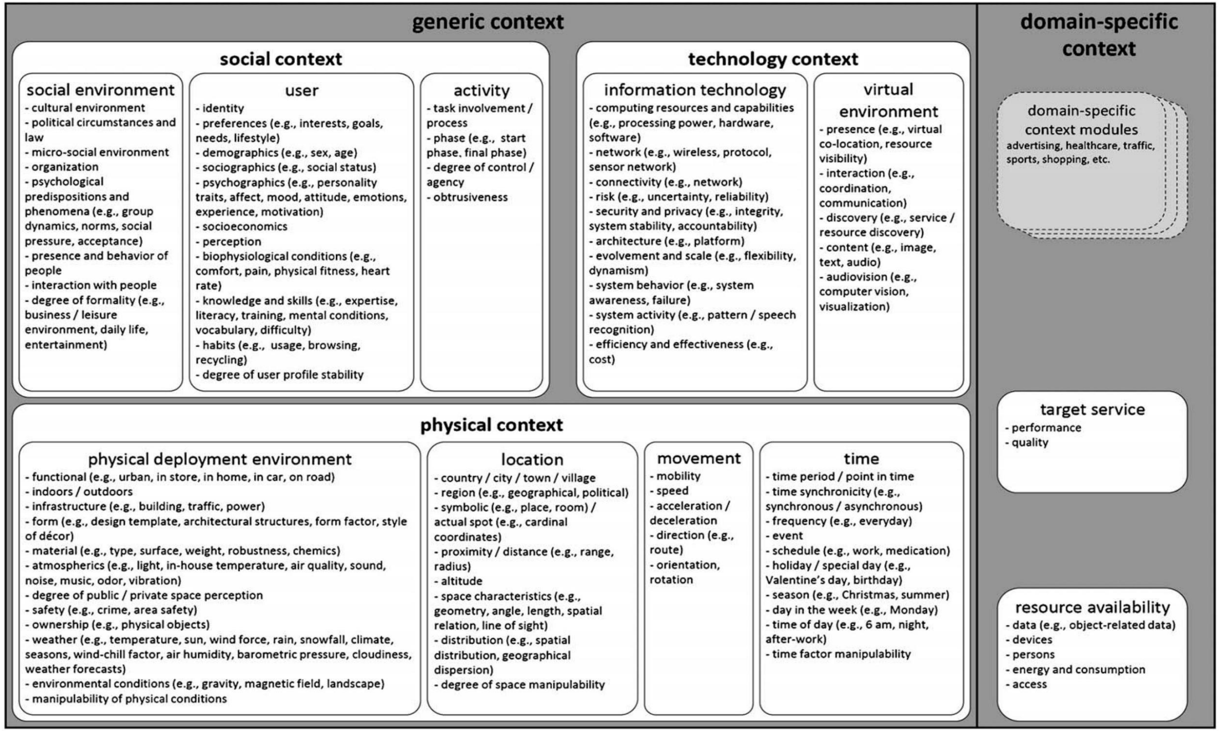
- Exists irrespective of a particular user
- Ex.: time, location, weather, traffic conditions, noise, light

User-related context/background

- Is connected to an individual user
- Ex.: activity, emotion, personality, social and cultural context

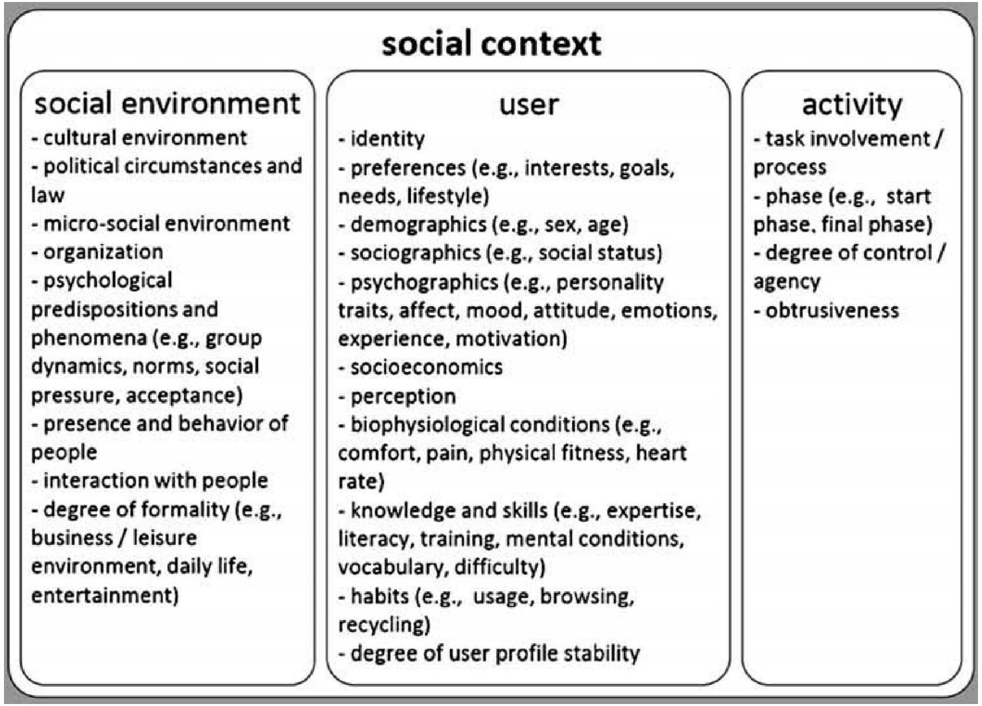
[Schedl et al., 2015] chapter *Music Recommender Systems*, Recommender Systems Handbook, Ricci et al. (eds.), 2nd ed., pp. 453-492.

Many more context categories



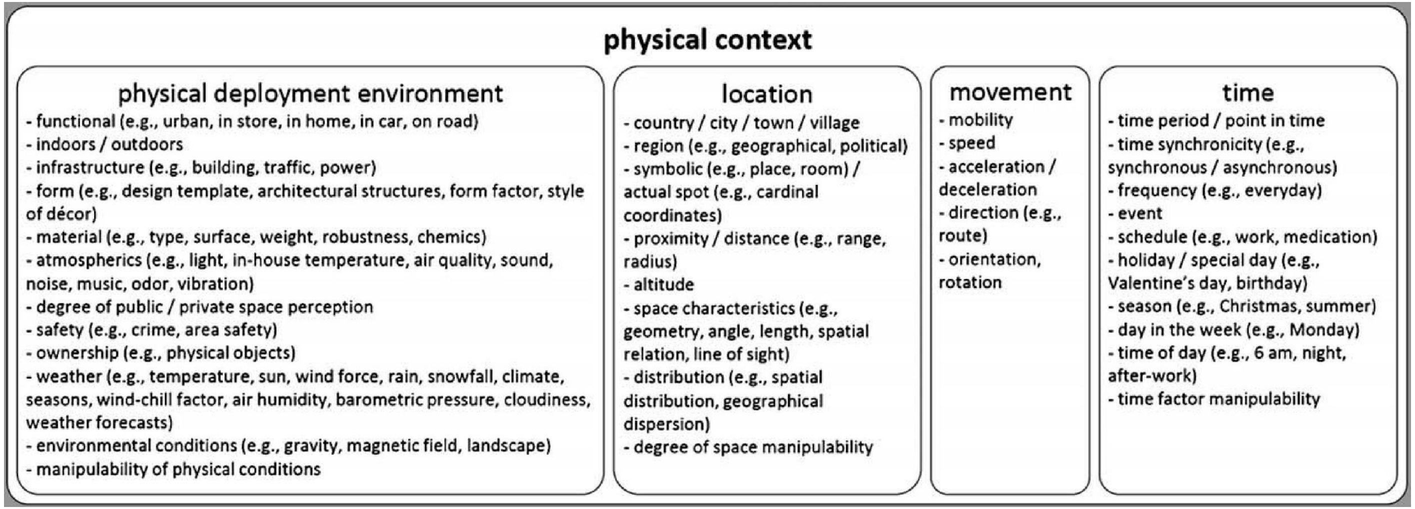
[Bauer & Novotny, 2017] *A consolidated view of context for intelligent systems*. Journal of Ambient Intelligence and Smart Environments, 9(4), 377-393. doi:10.3233/ais-170445

Many more context categories



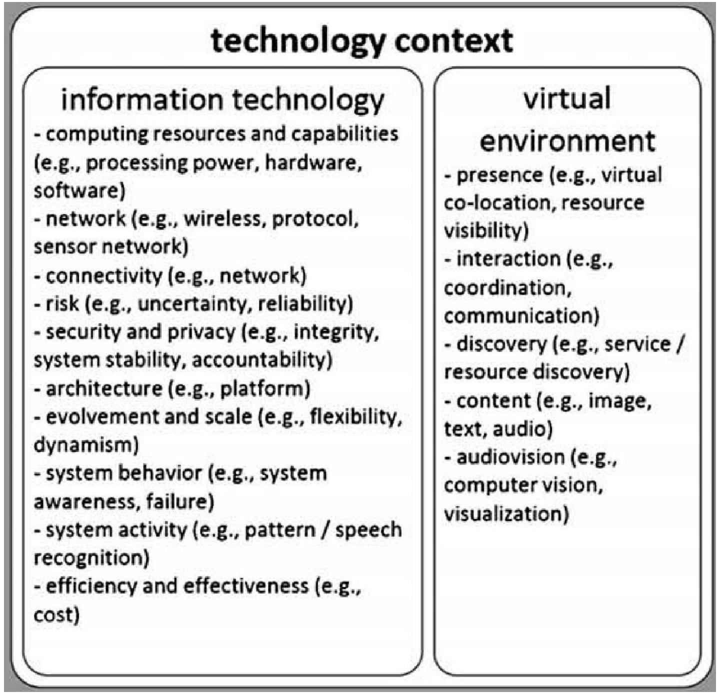
[Bauer & Novotny, 2017] *A consolidated view of context for intelligent systems*. Journal of Ambient Intelligence and Smart Environments, 9(4), 377-393. doi:10.3233/ais-170445

Many more context categories



[Bauer & Novotny, 2017] *A consolidated view of context for intelligent systems*. Journal of Ambient Intelligence and Smart Environments, 9(4), 377-393. doi:10.3233/ais-170445

Many more context categories



[Bauer & Novotny, 2017] *A consolidated view of context for intelligent systems*. Journal of Ambient Intelligence and Smart Environments, 9(4), 377-393. doi:10.3233/ais-170445

Obtaining context data

- **Explicitly**: elicited by direct user interaction (questions, ratings in context)
  Ex.: asking for user's mood or music preference (Likert-style ratings)

- **Implicitly**: no user interaction necessary
  Ex.: various sensor data in today's smart devices (heart rate, accelerometer, air pressure, light intensity, environmental noise level, etc.)

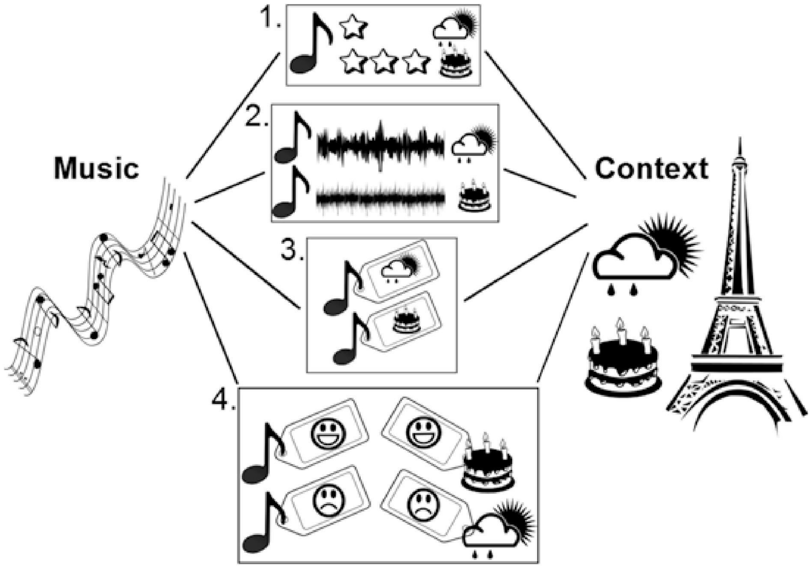
- **Inferring** (using rules or ML techniques):
  Ex.: time, position → *weather*; device acceleration (x, y, z axes), change in position/movement speed → *activity*; skipping behavior → music preferences

[Adomavicius & Tuzhilin, 2015] chapter *Context-Aware Recommender Systems*, Recommender Systems Handbook, Ricci et al. (eds.), 2nd ed., pp. 191-226.

Obtaining context data

Methods to establish **relationship: music ↔ context**

1. Rating music in context
2. Mapping audio/content features to context attributes
3. Direct labeling of music with context attributes
4. Predicting an intermediate context



[Schedl et al., 2015] chapter *Music Recommender Systems*, Recommender Systems Handbook, Ricci et al. (eds.), 2nd ed., pp. 453-492.

Methods/examples for context-aware MRS

- Mobile Music Genius

  [Schedl et al., 2014] *Mobile Music Genius: Reggae at the Beach, Metal on a Friday Night?*, Proceedings of the 2014 ACM International Conference on Multimedia Retrieval (ICMR).

- Just-for-me

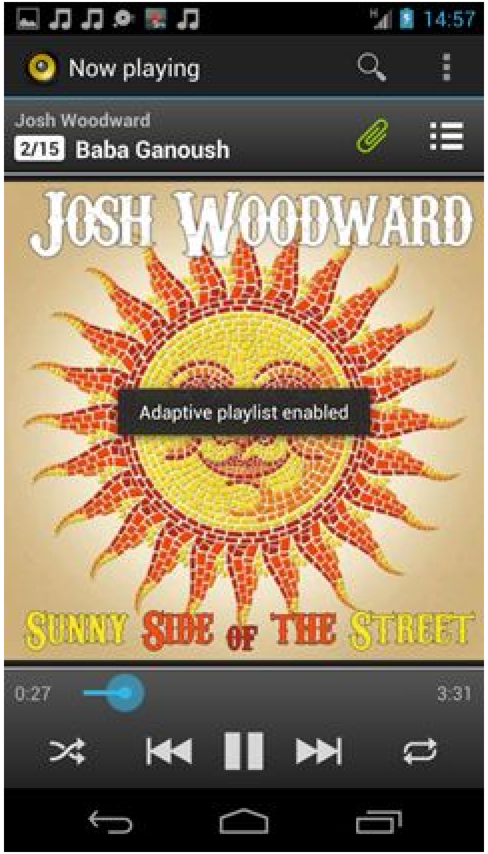
  [Cheng & Shen, 2014] *Just-for-Me: An Adaptive Personalization System for Location-Aware Social Music Recommendation*, Proceedings of the 2014 ACM International Conference on Multimedia Retrieval (ICMR).

- Music Recommendation for POIs

  [Kaminskas et al., 2013] *Location-aware Music Recommendation Using Auto-Tagging and Hybrid Matching*, Proceedings of the 7th ACM Conference on Recommender Systems (RecSys).

Mobile Music Genius

- Context-aware recommendation of next track in playlist
- Variety of context/sensors used, e.g., time, location, place, weather, device, activity, ambient (light, noise, etc.)
- Decision tree classifier continuously learns relationships: genre, artist, track → context attributes from user interactions (e.g., play, skip, stop events)

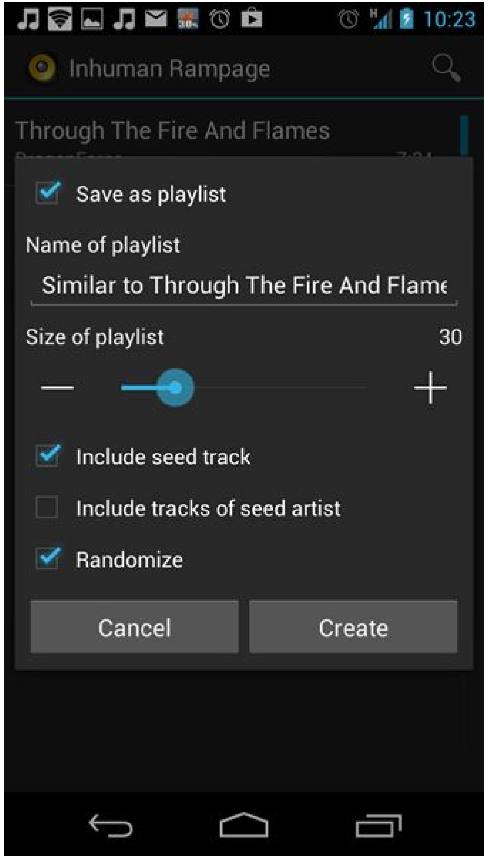


Mapping music/content features to context attributes

[Schedl et al., 2014] *Mobile Music Genius: Reggae at the Beach, Metal on a Friday Night?*, Proceedings of the 2014 ACM International Conference on Multimedia Retrieval (ICMR).

Recommendation approach

- Playlists created by track similarity, computed from Last.fm tags (cosine similarity on weighted artist and song tags)
- During playback: if change in context attributes exceeds sensitivity parameter, classifier is used to predict new track, which is played next



Evaluation

Classification accuracy of different classifiers and prediction targets:

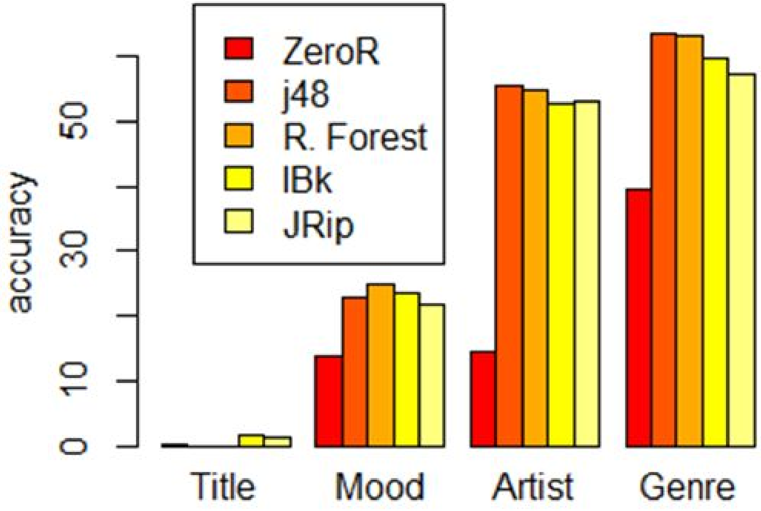
<u>Best results</u>

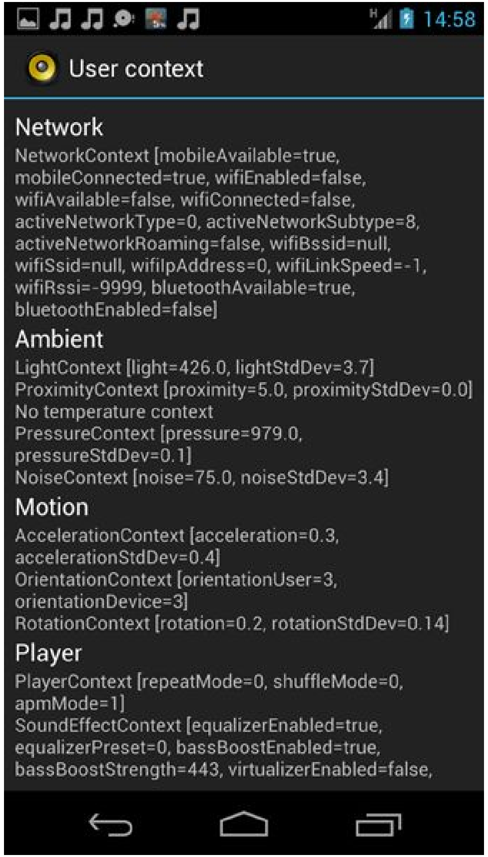
Title: 0.015

Mood: 0.230

Artist: 0.550

Genre: 0.610

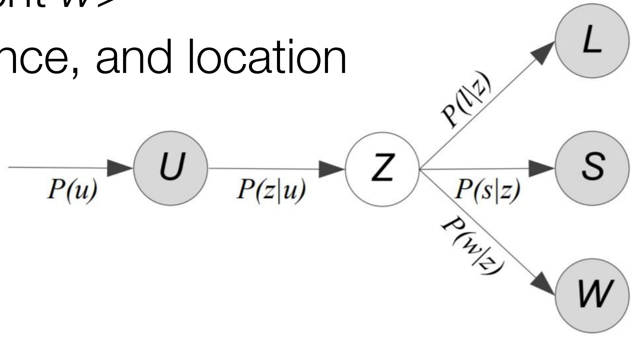




[Schedl et al., 2014] *Mobile Music Genius: Reggae at the Beach, Metal on a Friday Night?*, Proceedings of the 2014 ACM International Conference on Multimedia Retrieval (ICMR).

Just-for-me

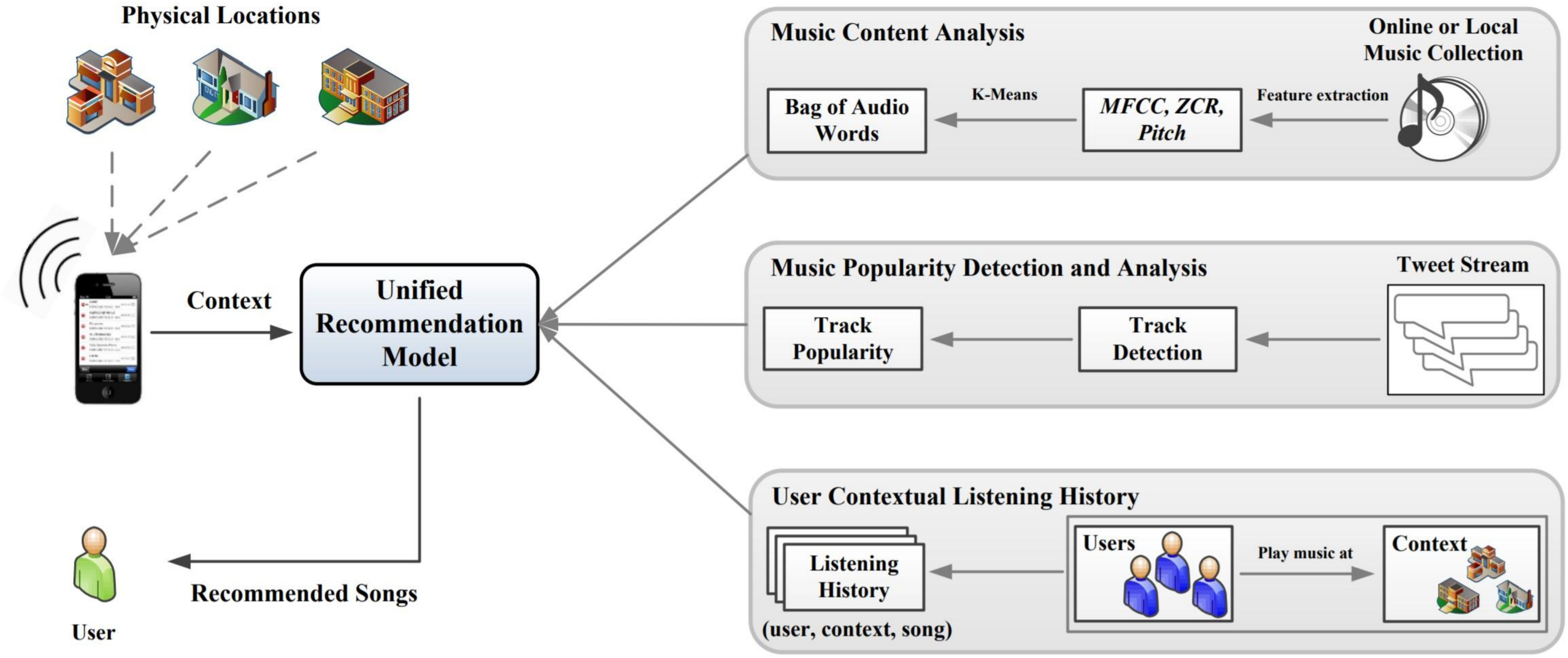
- Location-aware mobile music recommender
- Representation of play events:
  <user *u*, location *l*, track preference *s*, audio content *w*>
- Latent topic model used to relate content, preference, and location
- Trained via EM on existing user data
- Trained model used to predict *Pr(s|u,l)*
- Popularity estimation from tweets and integrated into track preference score (updated weekly)



Mapping music/content features to context attributes

[Cheng & Shen, 2014] *Just-for-Me: An Adaptive Personalization System for Location-Aware Social Music Recommendation*, Proceedings of the 2014 ACM International Conference on Multimedia Retrieval (ICMR).

Recommendation approach

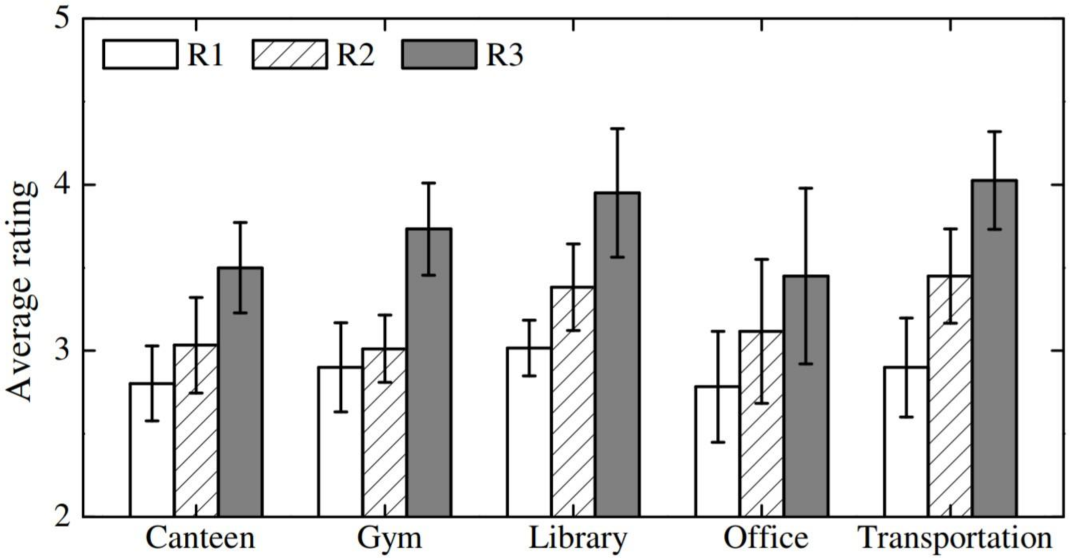


[Cheng & Shen, 2014] *Just-for-Me: An Adaptive Personalization System for Location-Aware Social Music Recommendation*, Proceedings of the 2014 ACM International Conference on Multimedia Retrieval (ICMR).

Evaluation

- 10 subjects (Asian, 6m/4f) rated up to 250 tracks in 5 contexts (canteen, gym, library, office, transportation), which are used for training
- 750 tracks used to create recommendations for user *u* at location *l*
- Subjects rated recommended tracks on Likert scale (1-5)

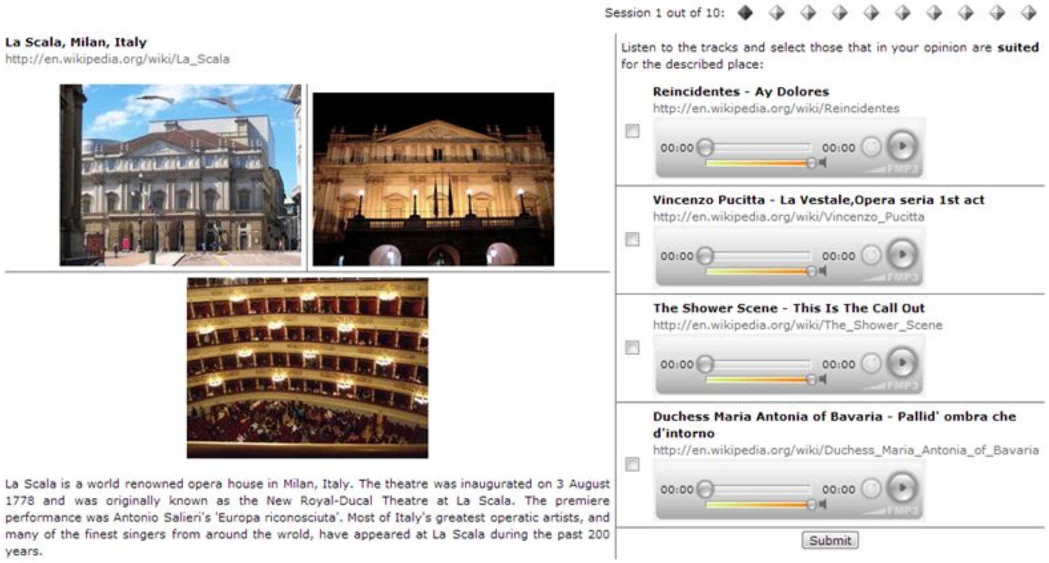
- <u>Baselines</u>
  R1: random track selection
  R2: location-based filtering w/o user preferences



Music recommendation for places of interest

- Combines: direct labeling, mapping audio/content features to context attributes, and predicting intermediate context

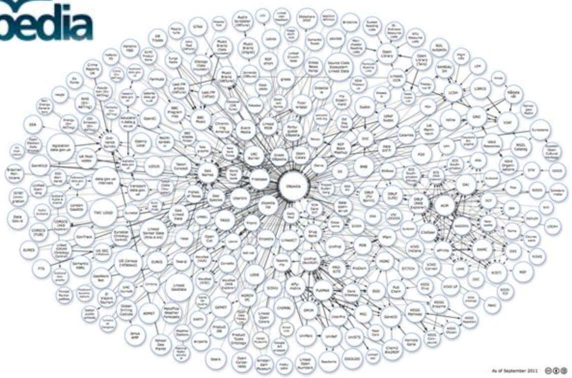
Direct labeling to create ground truth



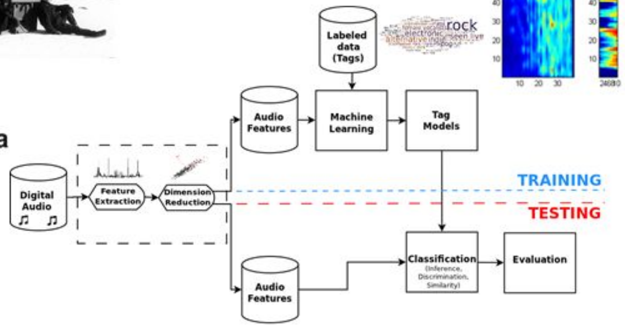
[Kaminskas et al., 2013] *Location-aware Music Recommendation Using Auto-Tagging and Hybrid Matching*, Proceedings of the 7th ACM Conference on Recommender Systems (RecSys).

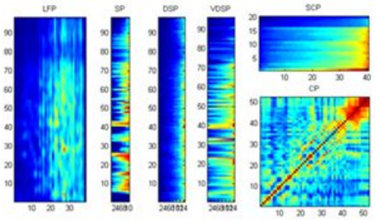
Recommendation approach

- Hybrid MRS fusing **knowledge-based** recommendations and **audio content-based** recommendations obtained via auto-tagging (rank fusion)



+





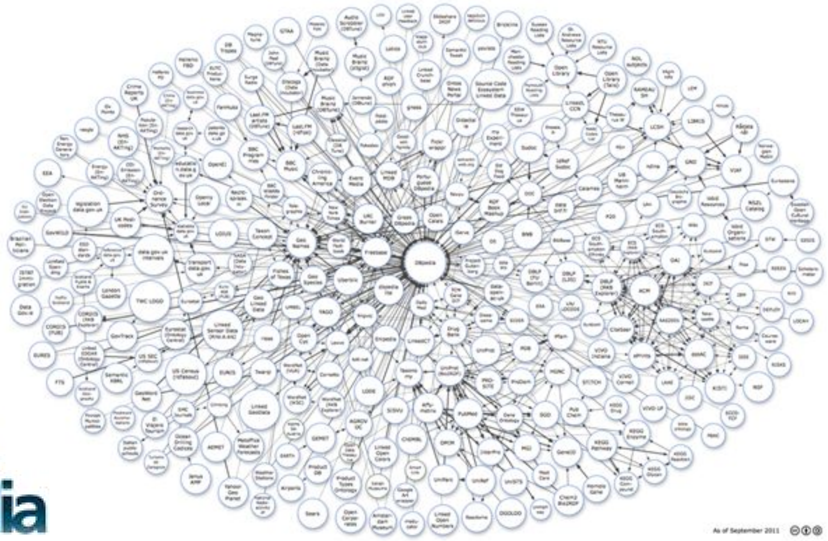
[Kaminskas et al., 2013] *Location-aware Music Recommendation Using Auto-Tagging and Hybrid Matching*, Proceedings of the 7th ACM Conference on Recommender Systems (RecSys).

Knowledge-based recommendation

- DBpedia knowledge graph
- Identify relations between musician and POIs (e.g., POI *located in* city, city *birthplace of* musician)
- Assign relevance weights to nodes and edges
- Estimate similarity/relatedness between POI and musicians via weight spreading

Predicting intermediate context

DBpedia



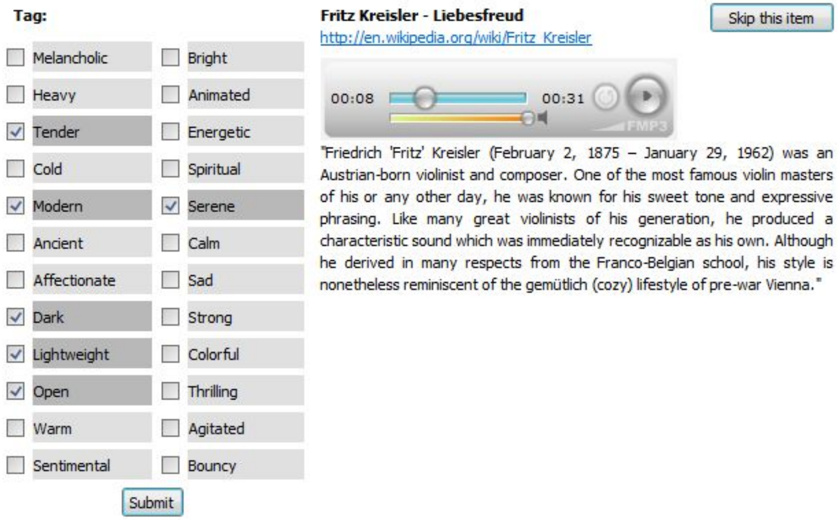
[Kaminskas et al., 2013] *Location-aware Music Recommendation Using Auto-Tagging and Hybrid Matching*, Proceedings of the 7th ACM Conference on Recommender Systems (RecSys).

Audio content-based recommendation

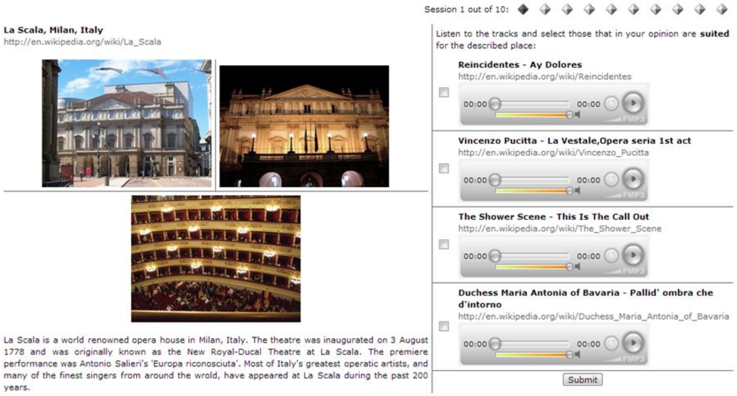
- Establish ground truths: track ← emotions, POI ← emotions (web survey)
- Train a music auto-tagger from ground truth data (track ← emotions)
- Use auto-tagger to predict emotions for unseen tracks (track → emotions)
- Establish similarity between POI and track via Jaccard index on “bag-of-tags” representations

Predicting intermediate context (emotions)

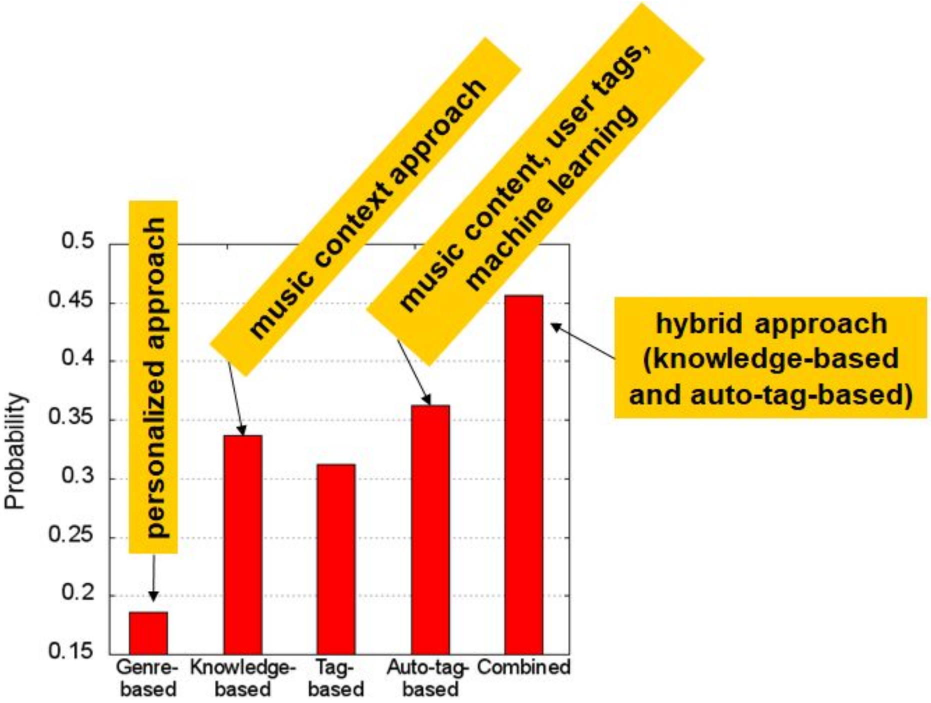
Mapping audio/content features to context attributes



Evaluation



Share of tracks marked as well-suited for POI, among all tracks recommended by given approach:



[Kaminskas et al., 2013] *Location-aware Music Recommendation Using Auto-Tagging and Hybrid Matching*, Proceedings of the 7th ACM Conference on Recommender Systems (RecSys).

More examples for context-aware MRS

- Music recommendation in **social context**,
  based on social graph via friendship relationships on Last.fm and KKBOX

[Chen et al., 2015] *Exploiting Latent Social Listening Representations for Music Recommendations*, Proceedings of the 9th ACM Conference on Recommender Systems (RecSys).

- Music recommendation **in a car** (InCarMusic),
  ratings in context (genre ↔ situation, e.g., driving style, sleepiness, weather)

[Baltrunas et al., 2011] *InCarMusic: Context-Aware Music Recommendations in a Car*, Proceesings of the International Conference on Electronic Commerce and Web Technologies (EC-Web).

- Music recommendation based on listener **emotion**,
  content-based approach based on direct labeling and emotion classification

[Bodarwé et al., 2011] *Emotion-based music recommendation using supervised learning*, Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia (MUM).

More examples for context-aware MRS

- Music recommendation based on **activity** and **mood**,
  based on real-life user annotations of activity and mood on a smartphone,
  plus sensor data, using factorization machines as RS

  [Teng et al., 2013] *A large in-situ dataset for context-aware music recommendation on smartphones*, Proceedings of the IEEE International Conference on Multimedia and Expo Workshops (ICME).

- Music recommendation for **daily activities**,
  based on automatic activity recognition from smartphone sensor data,
  matching with audio content features via probabilistic Bayes classifier

  [Chen et al., 2015] *Exploiting Latent Social Listening Representations for Music Recommendations*, Proceedings of the 9th ACM Conference on Recommender Systems (RecSys).

Cultural/regional specificities

- Example to analyze and integrate **listener background**

- Music preferences vary strongly between countries
  → recommendations should be tailored to cultural background
  → country information can be used to alleviate cold start ("single sign-on")

- Ex.: music preferences analyzed using LFM-1b dataset (>1b listening events of 120k Last.fm users, 585k artists, user demographics)

[Schedl, 2017] *Investigating Country-specific Music Preferences and Music Recommendation Algorithms with the LFM-1b Dataset*, International Journal of Multimedia Information Retrieval 6(1):71-84.

Populations' preferences

country	age	gender	users	rnb	rap	elect.	rock	blues	folk	jazz	punk	altern.	pop	metal	α
-	-	-	120175	3.34	3.41	11.18	18.27	3.28	5.61	3.97	6.19	16.75	13.64	3.98	0.493
US	-	-	10255	**3.00**	3.22	11.17	18.82	3.07	6.06	3.79	7.53	17.69	13.56	**3.29**	0.554
RU	-	-	5024	**1.55**	3.10	**14.30**	20.60	2.28	**4.58**	3.03	7.76	18.14	10.58	6.10	0.564
DE	-	-	4578	1.96	3.15	11.90	19.80	2.59	5.67	3.10	**7.93**	17.26	12.02	6.00	0.510
UK	-	-	4534	2.88	2.76	12.08	18.47	3.10	5.49	4.02	7.32	18.10	13.55	3.35	0.582
PL	-	-	4408	2.18	**3.81**	11.14	19.45	2.72	4.85	3.49	7.28	**19.08**	10.96	7.19	0.503
BR	-	-	3886	2.88	**1.90**	**8.29**	19.91	3.26	6.05	3.47	7.49	18.72	13.92	5.92	0.586
FI	-	-	1409	1.88	3.40	11.55	**21.45**	**2.20**	4.95	**2.92**	6.56	**16.41**	11.48	**9.85**	0.520
NL	-	-	1375	2.64	2.70	11.81	18.18	**3.65**	6.17	4.20	5.64	17.18	13.37	4.32	0.532
ES	-	-	1243	2.41	2.09	9.86	19.64	3.25	6.07	3.71	6.60	16.95	**14.22**	5.12	0.560
SE	-	-	1231	2.29	2.60	12.01	19.03	3.07	6.12	3.53	6.15	17.44	14.11	4.82	0.584
UA	-	-	1143	1.69	2.82	13.42	20.86	2.46	4.92	3.13	7.25	18.16	**10.56**	6.64	0.565
CA	-	-	1077	2.20	2.89	11.76	19.16	2.78	**6.37**	3.53	7.48	18.26	13.02	4.35	0.575
FR	-	-	1055	2.87	3.44	12.77	**17.58**	3.25	5.68	**4.71**	**5.55**	16.89	12.99	3.73	0.535

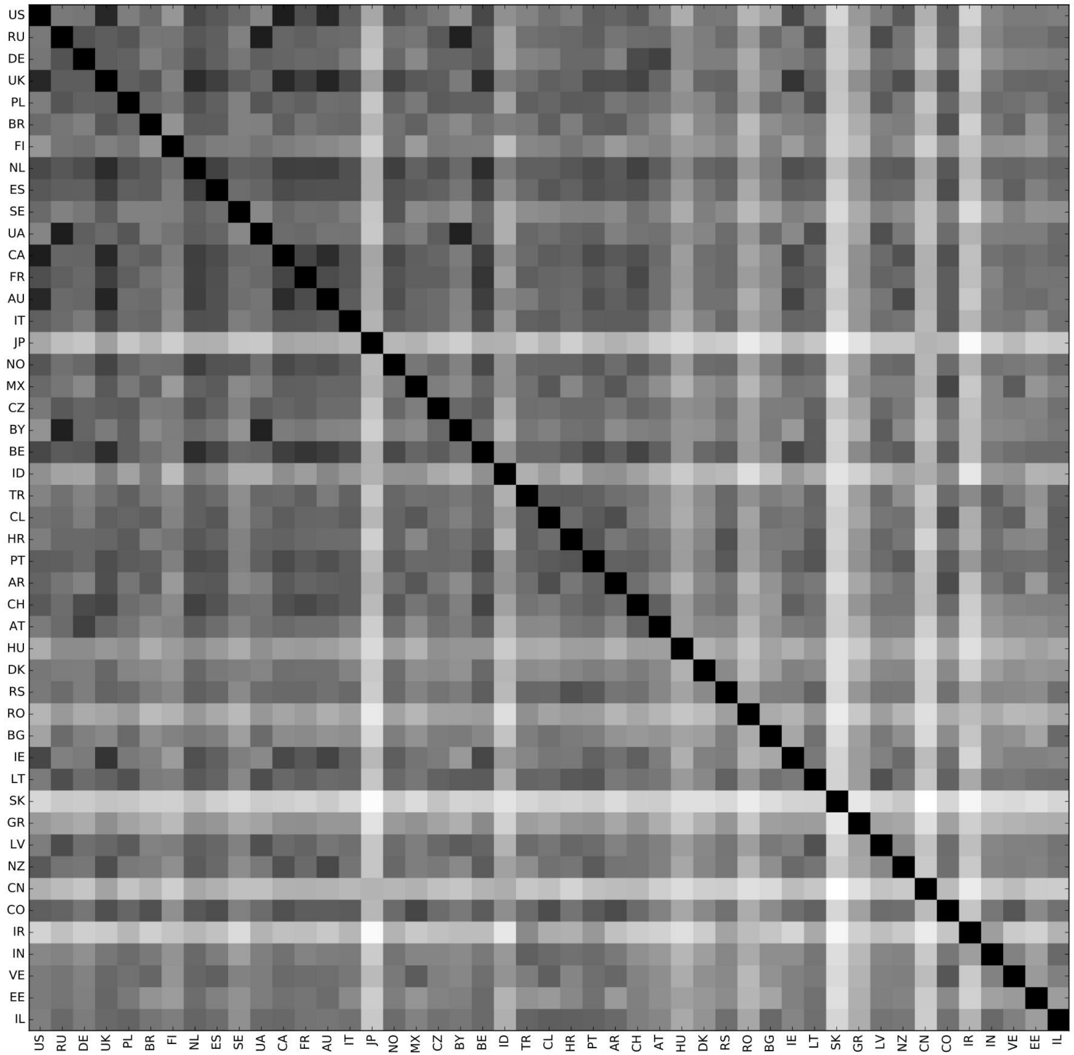
More fine-grained

U.S.A.		Japan		Finland	
Genre tag	PC	Genre tag	PC	Genre tag	PC
Rock	12.51	Rock	16.01	Rock	11.31
Alternative	9.63	Alternative	8.37	Metal	11.15
Alternative rock	5.86	J-pop	5.77	Alternative	7.30
Metal	4.77	Pop	4.56	Alternative rock	4.56
Pop	3.62	Metal	4.55	Hard rock	4.28
Indie	3.59	Alternative rock	4.26	Heavy metal	3.44
Hard rock	3.12	Indie	3.63	Death metal	2.74
Indie rock	3.09	Electronic	2.29	Classic rock	2.61
Classic rock	2.92	Hard rock	2.24	Pop	2.21
Electronic	2.33	Classic rock	2.23	Indie	2.13
Dance	2.21	Visual Kei	2.03	Electronic	2.00
Psychedelic	1.84	Indie rock	2.02	Indie rock	1.75
Blues	1.77	Heavy metal	1.68	Dance	1.71
Hip-Hop	1.72	Dance	1.66	Progressive rock	1.67
Punk	1.61	Punk	1.53	Nu metal	1.57
Heavy metal	1.49	Psychedelic	1.45	Progressive	1.50
Singer-songwriter	1.34	Anime	1.43	Power metal	1.46
Progressive	1.25	Electronica	1.43	Punk	1.45
Electronica	1.24	Blues	1.18	Alternative metal	1.32
Progressive rock	1.16	Japanese rock	1.17	Psychedelic	1.18
New Wave	1.08	Progressive rock	1.06	Hip-Hop	1.10
Punk rock	1.03	Pop punk	0.91	Electronica	0.90
Nu metal	0.99	Nu metal	0.86	Speed metal	0.89
Alternative metal	0.85	Progressive	0.86	Blues	0.84

Likeminded populations

Observations:

- Clusters of countries with same language: e.g., US, UK, Ireland, Australia, New Zealand
- Clusters of countries with same historical/cultural background: e.g., Russia, Ukraine, Belarus, (Lithuania, Latvia)
- Several outliers: e.g., Japan, China, Iran



Predicting music taste from culture

- Improve MRS in cold start situations

[Skowron et al.; 2017] *Predicting Genre Preferences from Cultural and Socio-economic Factors for Music Retrieval*, Proceedings of the 39th European Conference on Information Retrieval (ECIR).

Predicting music taste from culture

- Improve MRS in cold start situations
- **Culture** model: *Hofstede*

[Skowron et al.; 2017] *Predicting Genre Preferences from Cultural and* the
39th European Conference on Information Retrieval (ECIR).

Predicting music taste from culture

- Improve MRS in cold start situations
- **Culture** model: *Hofstede*
- **Socio-economic** model: *Quality of Government*
  (e.g., GDP, life expectancy, press freedom, ethnic fractionalization)
- Predicting music preferences of country as shares of genres (Gradient Boosting and Random Forest, 16% reduction of RMSE compared to global genre shares)

QOG DATA

[Skowron et al.; 2017] *Predicting Genre Preferences from Cultural and Socio-economic Factors for Music Retrieval*, Proceedings of the 39th European Conference on Information Retrieval (ECIR).

Evaluation summary

- Listener context and background are highly individual
- Need for context-sensitive evaluation strategies
- Automatic approaches typically fail:
  Which preferences are due to context and which due to other factors?
- Typically via user questionnaires/web surveys
- Careful selection of participants necessary (balance w.r.t. gender, age, profession, musical knowledge and background, etc.)

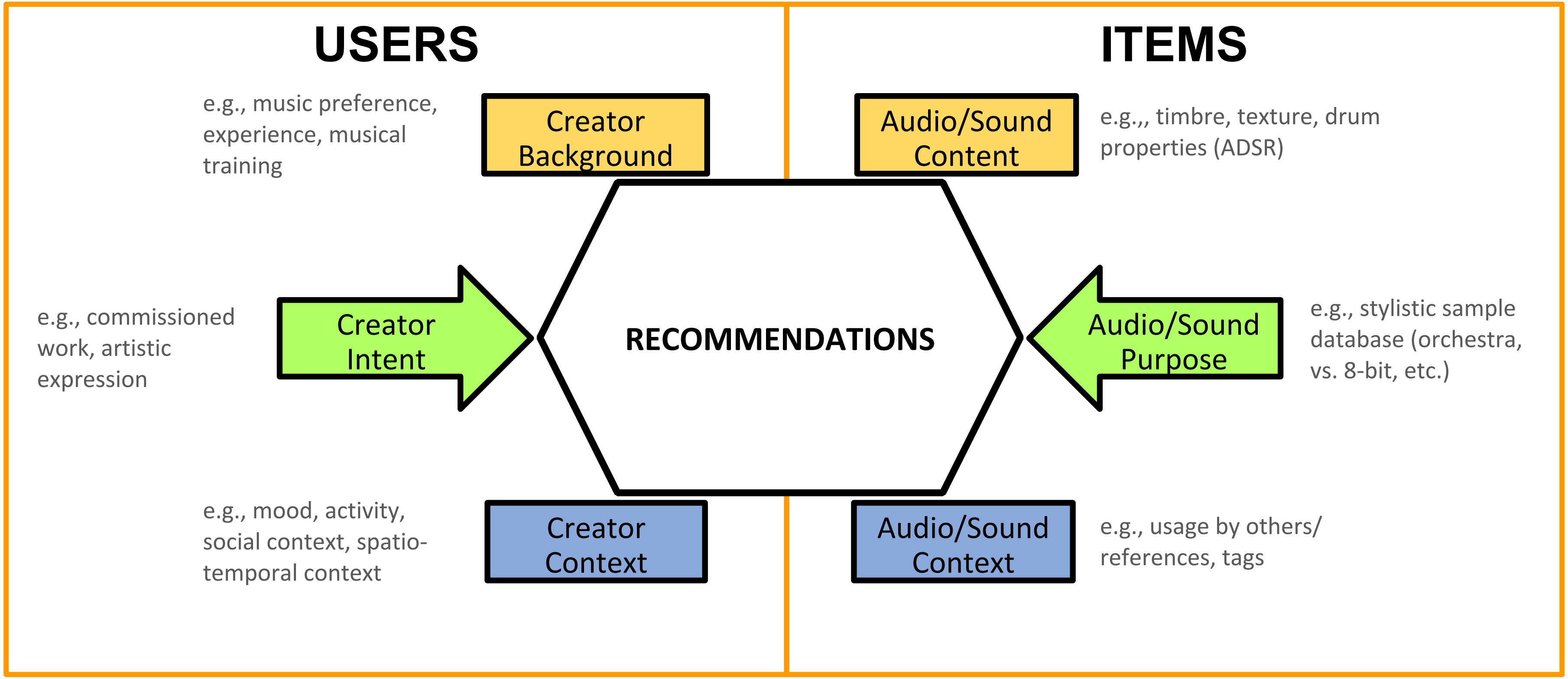
Trending topics

- Integrating intermediate representations (e.g., infer activity on smart phone)
- Culture-aware MRS
- Emotion-aware MRS
- Personality-aware MRS
- Exploit multimodal signals in context-aware MRS
- Automatic feature learning / deep learning

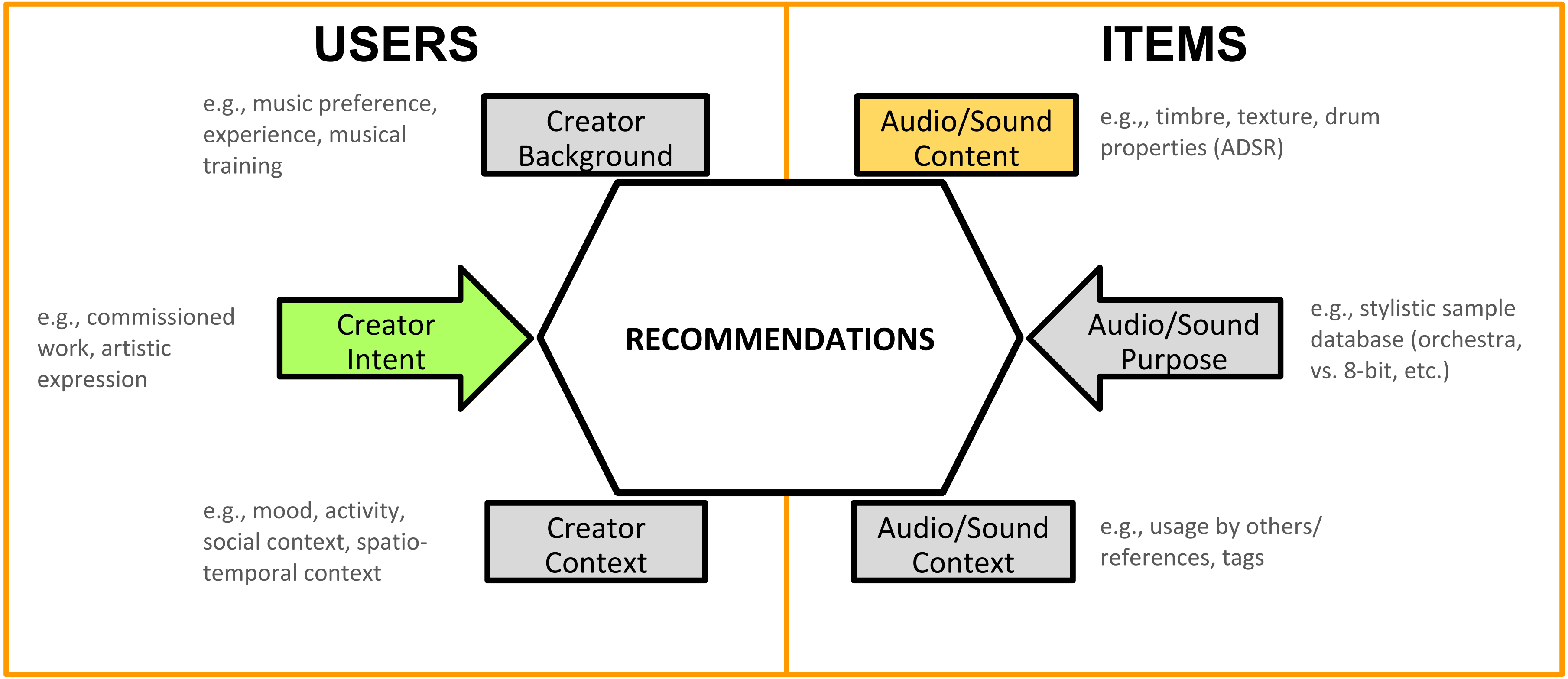
Use Case 3: Recommendation in the Creative Process

Music Industry Landscape (again)

Creator Hexagon



Creator Hexagon



RecSys for Music Producers

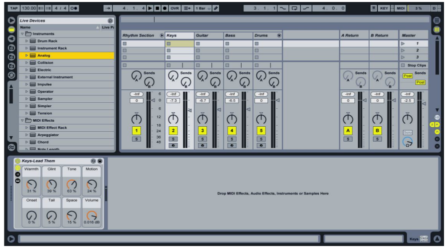
- Today, basically all music and audio production becomes digital at one point
- Used tools reflect current practice of music making
  - Sound synthesis, virtual instruments, samples, pre-recorded material, loops, effects
  - Mixing, mastering, control for live performances
- Finding the right sound remains a central challenge:

  "Because we usually have to browse really huge libraries [...] that most of the time are not really well organized." (TOK003)

  "Like, two hundred gigabytes of [samples]. I try to keep some kind of organization." (TOK006)

- Actually the ideal target group for music retrieval and recommendation

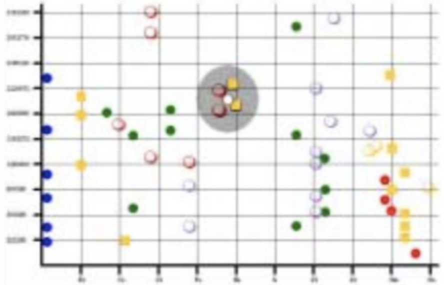
Digital Audio Workstations (DAWs)



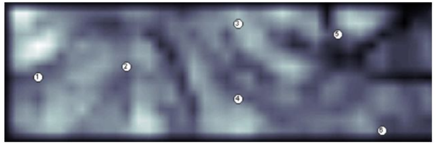
- Commercial products come with very large databases of sounds
- Screen optimized for arrangement/mixing
- UI for finding material marginalized or external window
- Incorporated strategies:
  - Name string matching
  - Tag search/filtering
  - Browsing (=scrolling lists)
- No one tags their library!

Facilitating Sound Retrieval

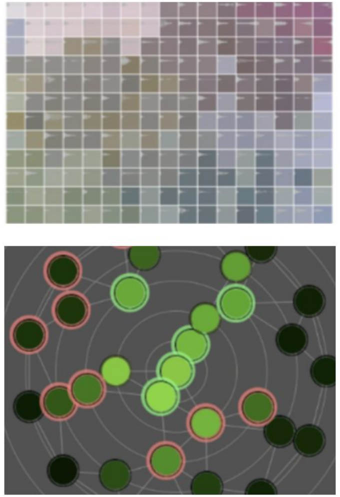
- New (academic) interfaces for sample browsing



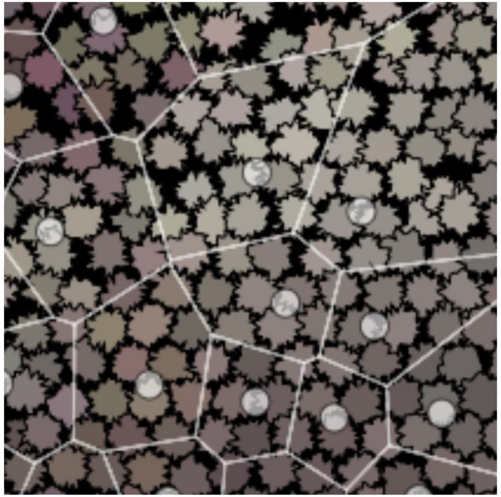
Sonic browser
(Fernström and Brazil, ICAD 2001)



Drum sample browser
(Pampalk et al., DAFx 2004)



Audio Quilt: snare, synth
(Fried et al., NIME 2014)



Texture browser
(Grill and Flexer, ICMC 2012)

- Not so much recommendation. Why?

Let’s Ask the Users!

- Interviews, tests, and feedback sessions
  - Participatory workshops
  - Music Hack Days
  - Red Bull Music Academy

- Unique opportunity for research to get access to up-and-coming musicians from around the world

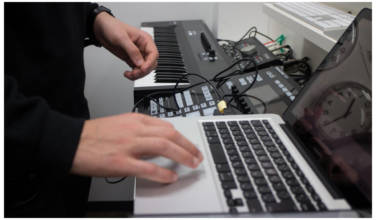
- Peer-conversations through semi-structured interviews

- Potentially using non-functional prototypes as conversation objects

[Andersen, Knees; 2016] *Conversations with Expert Users in Music Retrieval and Research Challenges for Creative MIR*. ISMIR.

[Ekstrand, Willemsen; 2016] *Behaviorism is Not Enough: Better Recommendations through Listening to Users*. RecSys.

The Role of Recommendation



- Recommenders are seen critical in creative work

"I am happy for it to make suggestions, especially if I can ignore them" (TOK007)

- Who is in charge?

"as long as it is not saying do this and do that." (TOK009)

- Artistic originality in jeopardy

"as soon as I feel, this is something you would suggest to this other guy as well, and then he might come up with the same melody, that feels not good to me. But if this engine kind of looked what I did so far in this track [...] as someone sitting next to me" (NIB4)

"then it's really like, you know, who is the composer of this?" (NIB3)

[Andersen, Grote; 2015] *GiantSteps: Semi-structured conversations with musicians*. CHI EA.

The Role of Recommendation (2)

- Users open to **personalization**, would accept cold-start

"You could imagine that your computer gets used to you, it learns what you mean by grainy, because it could be different from what that guy means by grainy" (PA008)

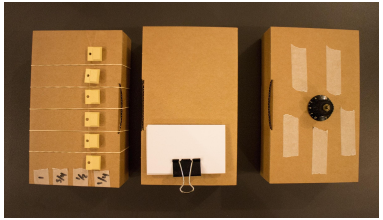
- Imitation is not the goal: **opposition** is the challenge

"I'd like it to do the opposite actually, because the point is to get a possibility, I mean I can already make it sound like me, it's easy." (TOK001)

"Make it complex in a way that I appreciate, like I would be more interested in something that made me sound like the opposite of me, but within the boundaries of what I like, because that's useful. Cause I can't do that on my own, it's like having a bandmate basically." (TOK007)

[Knees et al.; 2015] *"I'd like it to do the opposite": Music-Making Between Recommendation and Obstruction*. DMRS workshop.

The Role of Recommendation (3)



Two recurring themes wrt. recommendation:

1. Virtual band mate (controlled “collaborator”)

“I like to be completely in charge myself. I don’t like other humans sitting the chair, but I would like the machine to sit in the chair, as long as I get to decide when it gets out.” (TOK014)

2. Exploring non-similarity (“the other”, “the strange”)

“So if I set it to 100% precise I want it to find exactly what I am searching for and probably I will not find anything, but maybe if I instruct him for 15% and I input a beat or a musical phrase and it searches my samples for that. That could be interesting.” (TOK003)

cf. *defamiliarization*: art technique to find inspiration by making things different

“The Other” in RecSys and Creative Work

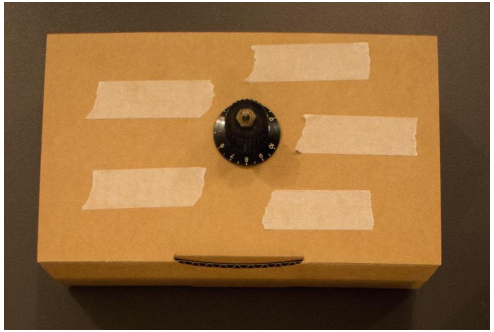
- **“Filter bubble” effects** in recommender systems:
  obvious, predictable, redundant, uninspiring, disengaging results
- Responses: optimizing for diversity, novelty, serendipity, unexpectedness
- In particular in creative work
  - no interest in imitating existing ideas and “more of the same” recommendations
  - challenging and questioning expectations and past behavior
- For **collaboration with an intelligent system** for creativity, opposite goals matter:
  - **change of context** instead of *contextual preservation*
  - **defamiliarization** instead of *predictability, explainability*
  - **opposition** instead of *imitation*
  - **obstruction** instead of *automation*

[Adamopoulos, Tuzhilin; 2015] *On Unexpectedness in Recommender Systems: Or How to Better Expect the Unexpected*. ACM TIST 5(4)

[Zhao, Lee; 2016] *How Much Novelty is Relevant?: It Depends on Your Curiosity.* SIGIR.

Testing the Idea of Controlled "Strangeness"

- Instead of retrieving "more of the same" through top-N results
- As a response, we **propose the idea of the Strangeness Dial**
- Device to **control the degree of otherness**
  → turn to left: standard similarity-based recommendations,
  → turn to right: "the other"
- Built as a non-functional prototype (cardboard box) to enable conversations
- Also tested as a software prototype for strangeness in rhythm variation



[Knees, Andersen; 2017] *Building Physical Props for Imagining Future Recommender Systems.* IUI HUMANIZE.

Responses to the Strangeness Dial (Idea)

- Idea and concept are received well (via non-functional prototype)

  "For search it would be amazing." (STRB006)

  "In synth sounds, it's very useful [...] Then the melody can also be still the same, but you can also just change the parameters within the synthesizer. That would be very cool." (STRB003)

  "That would be crazy and most importantly, it's not the same strange every time you turn it on." (TOK016)

- … but everybody understands it differently

  "Strangeness of genre maybe, how different genre you want. [...] It depends how we chart the parameter of your strangeness, if it's timbre or rhythm or speed or loudness, whatever." (STRB001)

  "No, it should be strange in that way, and then continue on in a different direction. That's the thing about strange, that there's so many variations of strange. There's the small, there's the big, there's the left, there's the right, up and down." (STRB006)

Responses to the Strangeness Dial (Prototype)

- The software prototype tried to present "otherness" in terms of <u>rhythm</u>
- This was perceived by some but didn't meet expectations of the majority

"I have no idea! It's just weird for me!" (UI03)

"It can be either super good or super bad." (UI09)

- Concept is highly subjective, semantics differ
- Demands for personalization (i.e., "which kind of strange are you talking about?")

"Then you have a lot of possibility of strange to chose from, actually. Like for me, I would be super interested to see it in 'your' strange, for example." (STRB006)

Some Takeaways

- User intent is a major factor
- Experts need recommenders mostly for inspiration: serendipity is key
- Control over recommendation desired (...transparency could help)
- Not much collaborative interaction data in this domain

  → Strong focus on content-based recommenders

  → To find what is unexpected, new sources of (collaborative) usage data need to be tapped
- Making music is mostly a collaborative task and a useful recommender needs to be a collaborator

Trending Topics

Intelligent machines in music creation: **AI for automatic composition**

Flow Machines (ERC project; François Pachet, now at Spotify)
e.g., assisted composition, automatic continuation/accompaniment, composition in style of X (“Daddy’s Car” … in the style of Beatles)

Magenta (Google project building on top of TensorFlow)
deep neural networks for, e.g., expressive renderings, sound generation, interactive note sequence generation

Jukedeck
automatic creation of royalty-free soundtracks, personalized music

IBM Watson Music

→ "Virtual Collaborator"

Working with Watson

Grammy award-winning music producer Alex Da Kid paired up with Watson to see if they could create a song together. Watson's ability to turn millions of unstructured data points into emotional insights would help create a new kind of music that for the first time ever, listened to the audience.

Not Easy
Alex Da Kid, X A...
Not Easy

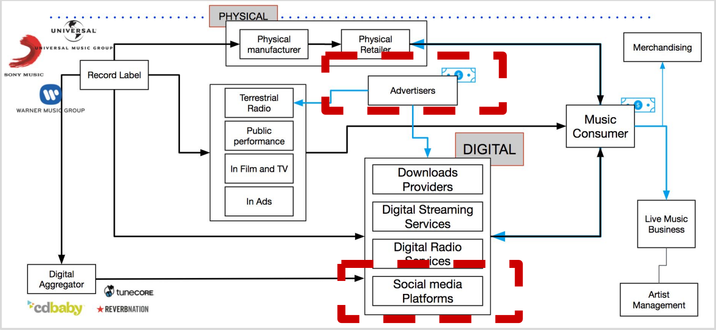
Available on iTunes

Cognitive creation

Alex Da Kid used Watson's emotional insights to develop 'heartbreak' as the concept for his first song, 'Not Easy,' and explored musical expressions of heartbreak by working with **Watson Beat**. Alex then collaborated with X Ambassadors to write the song's foundation, and lastly added genre-crossing artists Elle King and Wiz Khalifa to bring their own personal touches to the track. The result was an audience-driven song launching us all into the future of music.

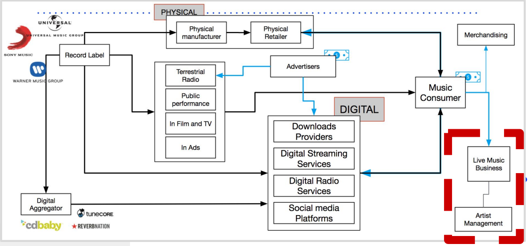
RecSys just an intermediary step to personalized content creation?

Further use cases



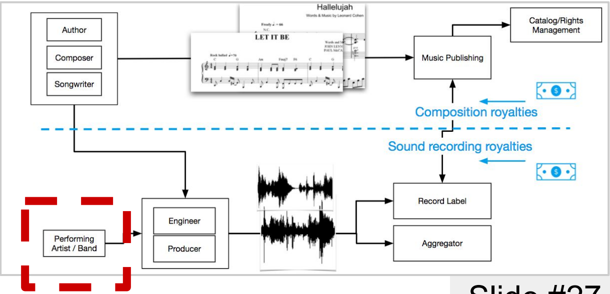
Slide #28

- Alternative audio content to music, e.g.
  - Ads (where a lot of $$$ is)
  - News, Podcasts
  - Artist messages
- Central battle-place of competition with AM/FM radio
  - Streaming in a better place for ads-targetting
  - Radio in a better place for alternative content
- Open problems:
  - How to sequence different types of content? (i.e. what content when?)
  - How to personalize?
  - How to present it to the listener?
  - How to blend music and audio in social media platform experiences?



Slide #28

Further use cases



Slide #27

- Live Music Business, e.g.
  - Recommending upcoming concerts to listeners
  - Recommending artists to e.g. music festivals

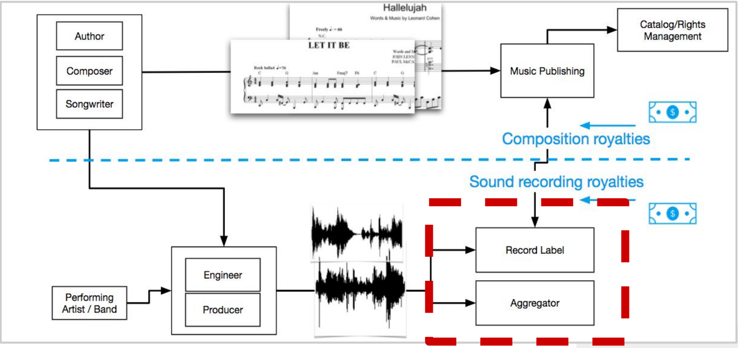
- Recommendations for artist management, e.g.
  - Help agents find best opportunities for artists

- Recommendations to artists
  - Recommending artists where to play
  - Help artists grow their careers, with insights based on data
  - Help artists communication with their fanbase

Further use cases

- RecSys (and data science) for record labels, e.g.
  - Assist A&R in finding new talents
  - An artist is launching an album, which track(s) to promote?
  - Make the best use / better monetization of back-catalogue
  - General assistance in business decisions
  - Marketing (where, to whom, how)
  - etc.

NB: Interesting explore/exploit trade-off



Slide #27

- NB: Some of these use cases addressed in upcoming H2020 project *FuturePulse*
  (Multimodal Predictive Analytics and Recommendation Services for the Music Industry)

Ethics

- Business-related recommendations (e.g. promotional content) vs. what the user actually wants/needs
- Impact on popular culture (shaping what makes popular culture)
    - Responsibility to counteract algorithmic biases and business-only metrics
    - “Filter bubble”
- Impact on “how” people listen to music (e.g. influence on curiosity)
- Impact on artists, on what’s successful, on the type of music composed

- Privacy (couldn’t attend tutorial next door right now ;-)

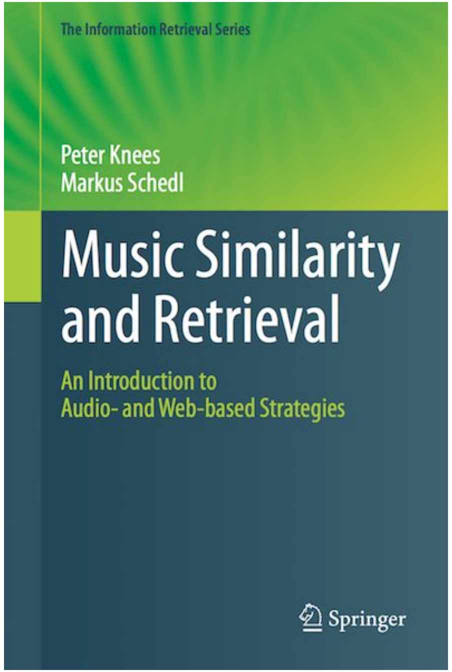
[Knijnenburg, Berkovsky, 2017] *Privacy for Recommender Systems*, Tutorial RecSys 2017

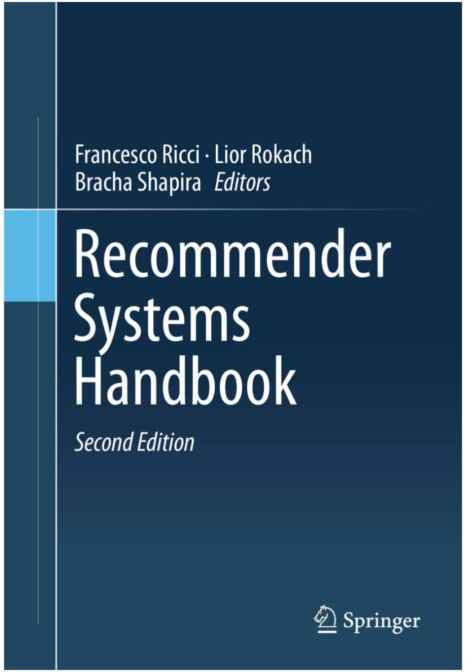
Challenges

- Recommending diverse types of content
- Understanding listening behavior in context
- Blending social interactions in music streaming
- Blending human-curated recommendations with algorithmic ones
- Transparency and trust
- Managing a listener's plurality of tastes without being disruptive
- Metrics for approximating long-term user satisfaction
- Voice-driven music interactions (in car, at home)

[Motajcsek et al. 2016] *Algorithms Aside: Recommendations as the Lens of Life*, RecSys 2016

More on This…





Music Similarity and Retrieval

by P. Knees and M. Schedl

Recommender Systems Handbook (2nd ed.)

Chapter 13: Music Recommender Systems

by M. Schedl, P. Knees, B. McFee, D. Bogdanov, and M. Kaminskas

Take-Home Messages

- Music is not “just another item”
- Dramatic changes in music consumption (growth, ownership → access) imply great challenges and impact/benefit for RecSys community
- RecSys technology has potential to be disruptive in many parts of the music industry (not just end-user consumption)
- Creating truly personalized music RecSys and evaluating user satisfaction is still challenging

Practical: Datasets

- Million Song Dataset: https://labrosa.ee.columbia.edu/millionsong
- Million Musical Tweets Dataset: http://www.cp.jku.at/datasets/mmtd
- #nowplaying Spotify playlists dataset: http://dbis-nowplaying.uibk.ac.at
- LFM-1b: http://www.cp.jku.at/datasets/LFM-1b
- Celma's Last.fm datasets: http://www.dtic.upf.edu/~ocelma/MusicRecommendationDataset/index.html
- Yahoo! Music: http://proceedings.mlr.press/v18/dror12a.html
- Art of the Mix (AotM-2011) playlists: https://bmcfee.github.io/data/aotm2011.html

Acknowledgments

- Faculty of Informatics, TU Wien
- The good people of the GiantSteps project
- Austrian Science Fund (FWF)
- Oscar Celma
- Andreas Ehmann
- Gordon Rios
- Tao Ye
- Sid Patil
- Cristina Sá
- Massimo Quadrana
- Kristina Andersen

The End

Thank you

Q&A

in

https://www.linkedin.com/in/markus-schedl-b16884/
https://www.linkedin.com/in/peterknees/
https://www.linkedin.com/in/fgouyon/

@m_schedl
@peter_knees
@fgouyon