

Wintersemester 2016/2017

Music Information Retrieval

Marko Harasic

Freie Universität Berlin

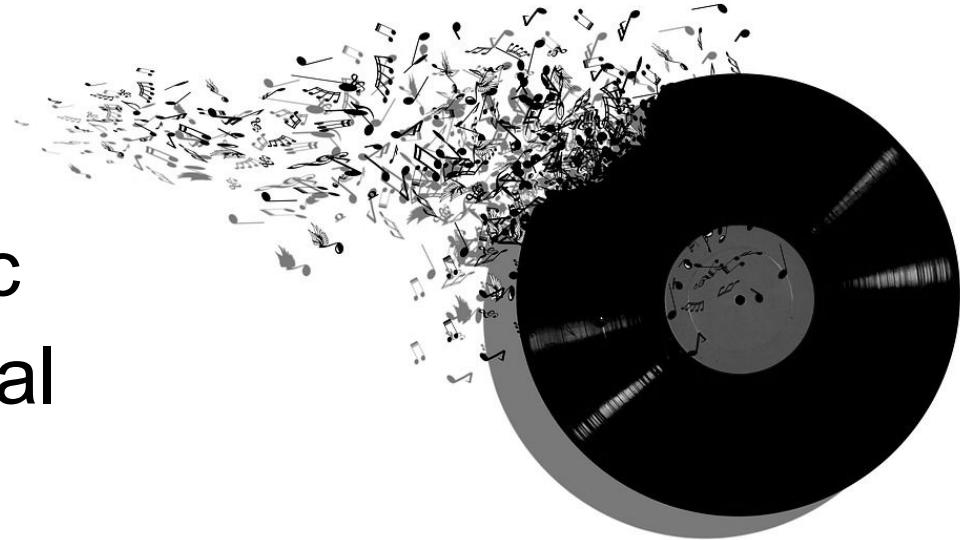
Department of Computer Science

Web Based Information Systems Workgroup

[mailto: harasic@inf.fu-berlin.de](mailto:harasic@inf.fu-berlin.de)



- Organisatorisches
- Einführung in Music Information Retrieval
- Themenvergabe





Organisatorisches

Veranstalter



- Veranstalter:
 - Marko Harasic
 - Mail: harasic@inf.fu-berlin.de
 - AG Netzbasierte Informationssysteme
 - Königin-Luise-Str. 24-26
 - **(NICHT Takustr. 9)**
 - Raum 114
- Sprechstunde:
 - Termine per Mail abstimmen





- Veranstaltungshomepage:
 - https://www.mi.fu-berlin.de/inf/groups/hcc/teaching/Wintersemester-16_17/Music-Information-Retrieval.html
- Hier finden sich:
 - Folien der Veranstaltung
 - Termine der Präsentationen
 - Hinweise/Links auf Literatur
 - Scheinkriterien

Mailingliste

- Eintrag in die Mailingliste:
- https://lists.spline.inf.fu-berlin.de/mailman/listinfo/nbi_s_mir
- **Wer:** Alle
- **Wann:** Jetzt
- **Warum:**
 - Raumänderungen
 - Terminänderungen
 - Ankündigungen
 - Material
 -





- Referatsbeitrag
 - Quellen: Vom Veranstalter bereitgestellte Paper + Eigenrecherche
- Ausarbeitung
 - Referatsinhalt als gegliederten wissenschaftlicher Text
- Note :
 - 50 % Ausarbeitung
 - 50 % Referatsbeitrag
- 4 bzw. 5 Leistungspunkte

Ablauf

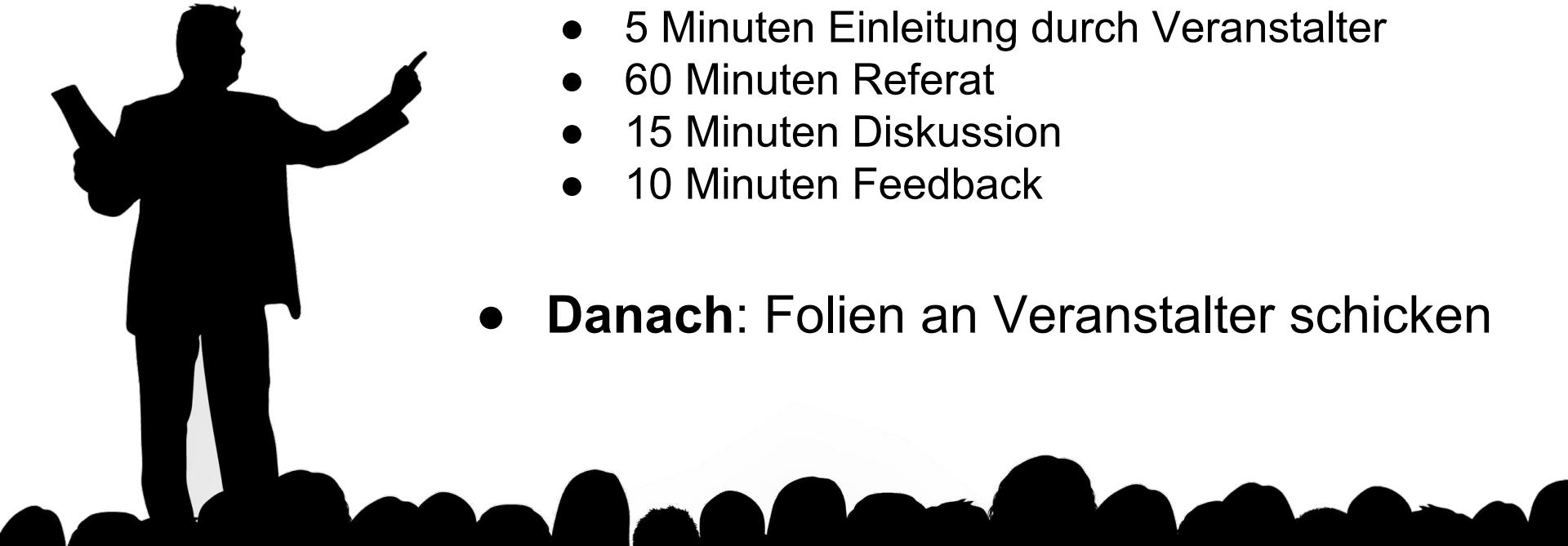


- Heute: Vergabe von Referatsthemen
- Jeweils zwei Wochen vor Referat:
 - Folienentwurf Veranstalter schicken
 - Vorbesprechungstermin vereinbaren
- Ohne Termin und Entwurf:
 **Keinen Schein**
- Referatstermin
- Abgabe der Ausarbeitung Ende der Vorlesungszeit (04.03.17)





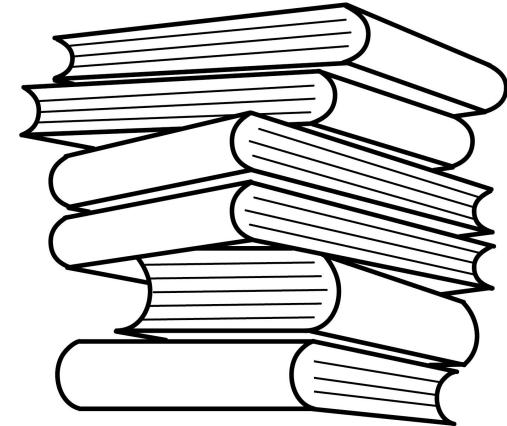
- **Davor:** Wird Notebook, Kabel... benötigt?
- Vorträge dauern 90 Minuten:
 - 5 Minuten Einleitung durch Veranstalter
 - 60 Minuten Referat
 - 15 Minuten Diskussion
 - 10 Minuten Feedback
- **Danach:** Folien an Veranstalter schicken



Ausarbeitung



- gibt Referat als gegliederten wissenschaftlichen Text wieder
- entspricht Ausarbeitungstechniken
- ermöglicht Erschließung des Themas für Abwesende
- hat **8 Seiten pro Person**
- verwendet das zu Verfügung gestellte LATEX-Template
- wird bis **04.03.17** dem Dozenten zugeschickt





Einführung in Music Information Retrieval

Was ist Music Information Retrieval?

- Interdisziplinäres Forschungsgebiet
 - Musikwissenschaften
 - Informatik
 - Signalanalyse



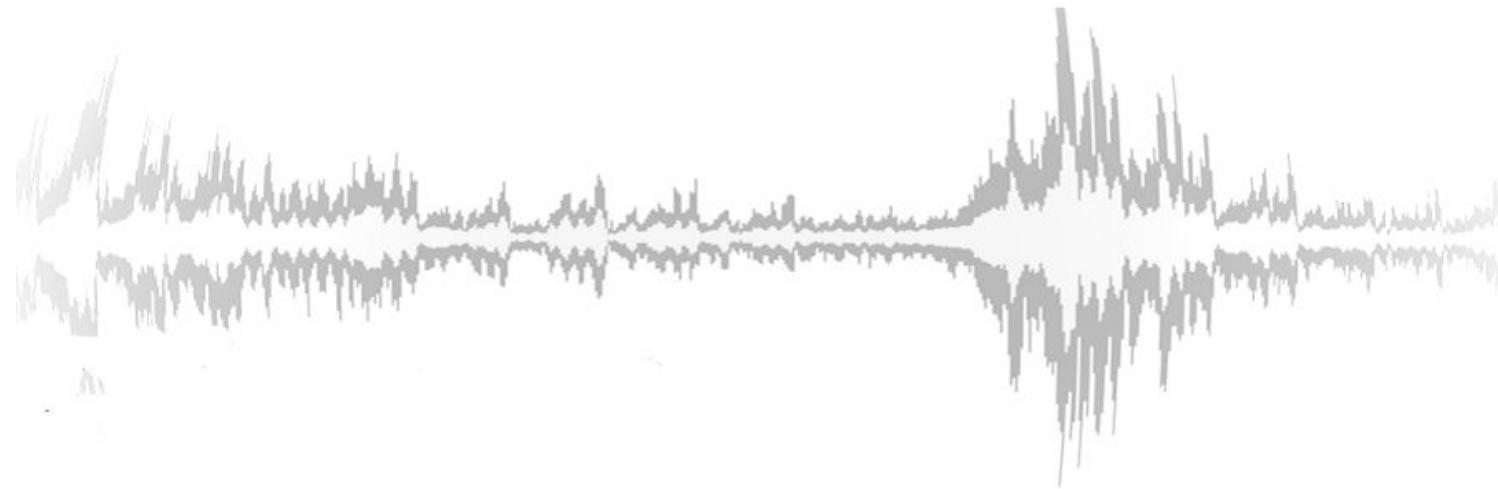
Stellt Methoden zur Extraktion, Analyse und Verarbeitung beliebiger musikalischer Entitäten auf jeder Repräsentationsebene bereit.

- Similarity
- Classification / Clustering
- Search
- Recommendation
- Visualisation



Similarity by Audio-Signals

- Content-based Retrieval von Liedern
- Analyse des Audio-Signals nach:
 - Loudness
 - Pitch
 - Brightness
 - Bandwidth
 - Harmonicity
 - Dominant Frequency



Similarity by Lyrics



- Liedtexte repräsentieren einen semantischen Aspekt von Musik
- Musikrichtungen haben meist spezifisches Vokabular
- Ansatz:
 - Analyse der Liedtexte
 - Extraktion von N Topics z.B. mit PLSA
 - Topiczuweisung zu den Liedern
- Ähnliche Lieder haben ähnliche Topics

*Someday you gotta find another way
You better right your mind and live by what you say
Today is just another day
Unless you set your sights and try to find a way*

*I say fuck authority
Silent majority
Raised by the system
Now it's time to rise against them
We're sick of your treason
Sick of your lies
Fuck no we won't listen
We're gonna open your eyes*

*Frustration, domination
Feel the rage of a new generation
We're living we're dying and we're
Never ever gonna stop, stop trying*

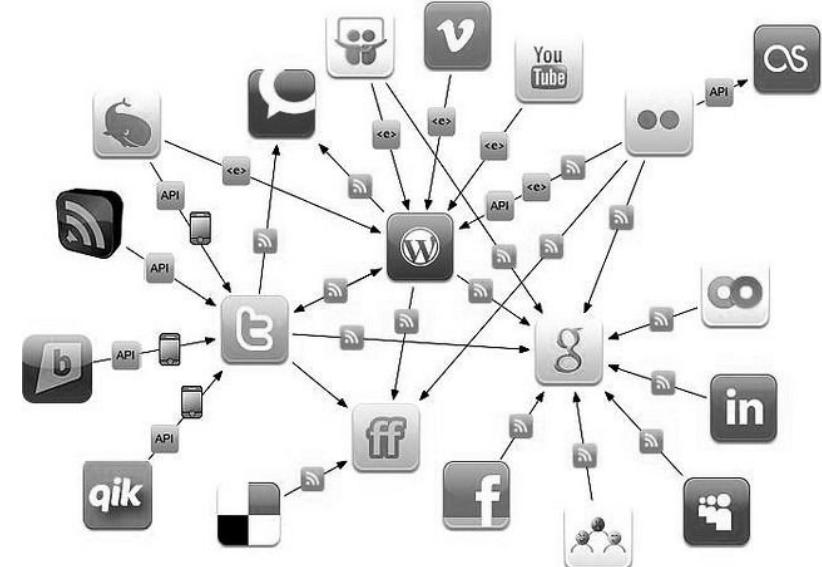
*You know the time is right to take control
We got to take offense against the status quo
No way not gonna stand for it today
Fight for your rights it's time we had our say!*

*I say fuck authority
Silent majority
Raised by the system
Now it's time to rise against them
we're sick of your treason
Sick of your lies
Fuck no we won't listen
We're gonna open your eyes!*

*Frustration, domination
Feel the rage of a new generation
We're living, we're dying
We're sick and tired of relentless lying
Destroy, enjoy, your fuckin world is our new toy
Dominate, eliminate, You're gonna feel the wrath, wrath of hate*

Similarity by Meta-Data

- Meta-Daten liefern zusätzliche Informationen über Musik
- Ähnliche Künstler haben ähnliche Meta-Informationen
- Ansatz:
 - Crawling von externen Quellen
 - RSS-Feeds
 - Blogs
 - Musik-spezifische Webseiten
 - Wikipedia
 - Twitter
 - Social - Tags (Last.fm...)
 -
 - Verwendung von klassischen Verfahren aus dem Document-Retrieval



Mood Detection



- Musik hat Einfluss auf menschliche Gefühle
- Klassifikation von Musik in Kategorien von Emotionen
- Ansatz:
 - Machine-Learning supervised Learning
 - Emotionsklassen sind vorgegeben
 - Vergleich anhand Audiosignale
 - Game with a purpose
 - Collaborative Tagging
 - Extraktion von Emotions-Tags aus Webseiten



Genre determination

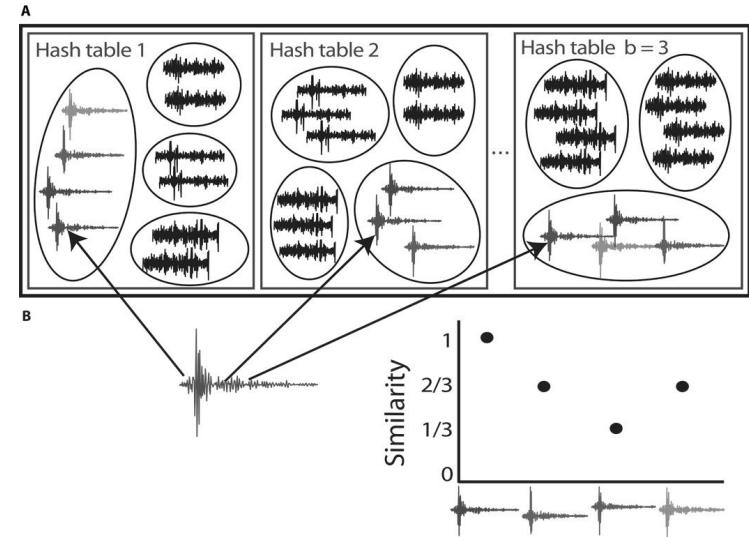


- Genre Bestimmung entspricht Klassifikation
 - Auto-Tagging hilfreich für Sammlungen
 - Keine einheitliche Taxonomie existiert
 - Ansatz:
 - User-Generated Tags (last.fm, Allmusic)
 - Low-level Analyse der Audiosignale



Recognition of Coversongs

- Coversongs sind alternative Wiedergaben eines Liedes
- Wahren meist die melodische Struktur
- Berücksichtigung unterschiedlicher Aspekte
 - Instrumentalisierung
 - Tempo
 - Tonhöhe
- Skalierbarkeitsproblem
- Ansatz:
 - Locality Sensitive Hashing
 - Ähnliche Lieder in gleiche Buckets



Acoustic Fingerprinting / Shazam

- Ermittlung eines konkreten Songs
 - Repräsentation von Lied-Fragmenten per Hashwert
 - Robust gegen Störungen
 - Vergleich der gehashten Aufnahme mit gespeicherten Hashes



Query by Humming



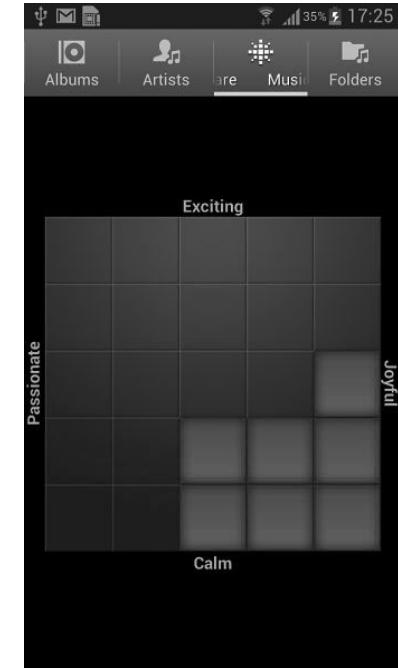
- Suche nach Liedern anhand einer “gesummtten” Melodie
- Extraktion signifikanter Melodien des Lieds
- Erstellung von Melodie-Datenbanken durch Domänen-Experten
- Ansatz:
 - Fingerprinting ist nicht möglich
 - Vergleich von Pitches im Signal
 - n-gram Überlappung
 - Hidden Markov Modelle
 - Spracherkennung



Matrix Factorisation / Spotify

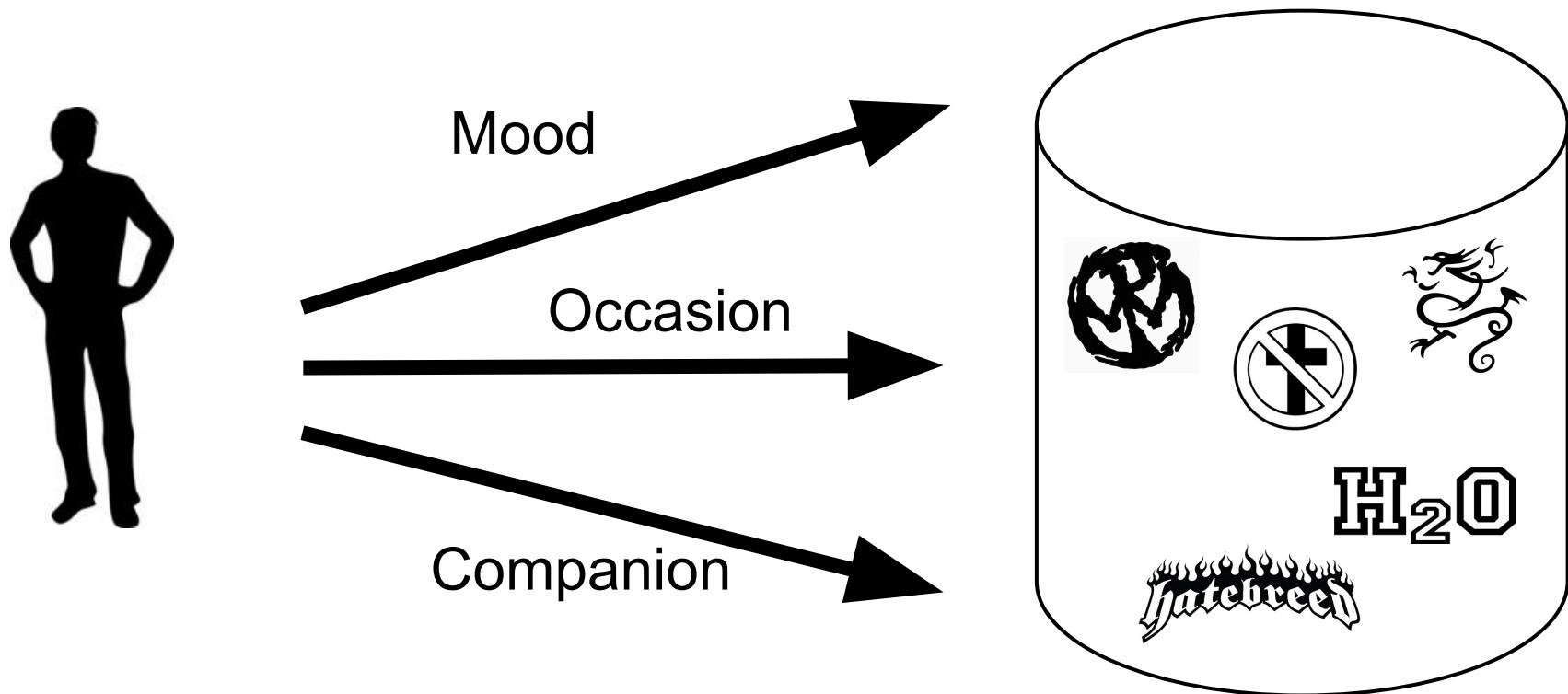


- Collaborative Filtering Ansatz
- Benutzer und Lieder bilden $N \times M$ Matrix
 - Listening Events als implizite Bewertungen
 - Nullstellen entsprechen ungehörten Liedern
- Zerlegung in zwei Matrizen $N \times k$ und $k \times M$
 - Lieder haben k latente Eigenschaften
 - Benutzer haben k latente Präferenzen
- Dabei Füllung der Nullstellen
- Höchste Werte dann die Empfehlungen



Context Aware Recommender Systems

- Traditionelle RS: Benutzer X Lieder → Bewertung
- Kontextuelle RS: Benutzer X Lieder X Kontext → Bewertung



Automatic Playlist Generation

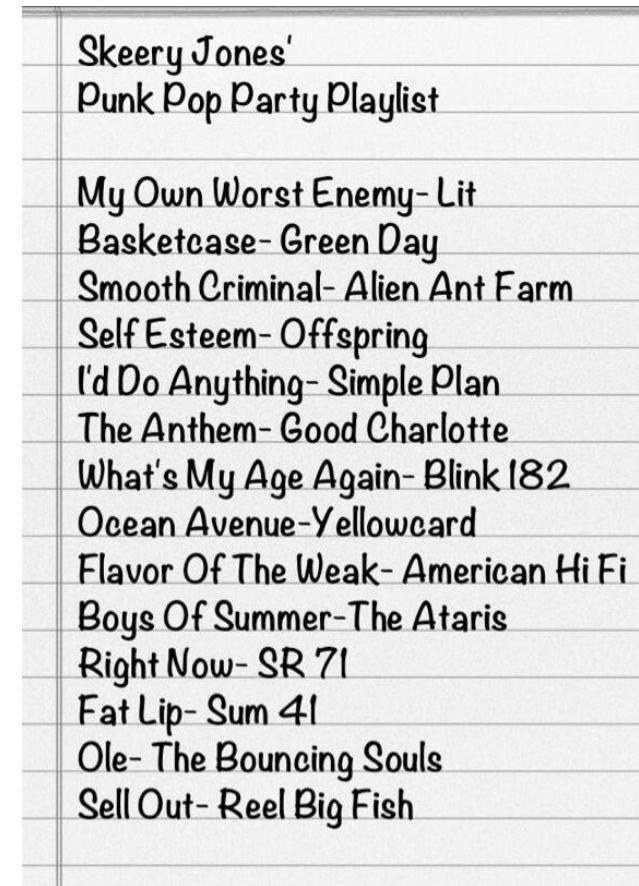


- Musik wird sequenziell konsumiert
- Ermöglicht Erzeugung eines personalisierten Streams
- Ansätze:
 - Playlisten basieren auf Listen anderer Benutzer
 - Benutzer liefern einen Seed-Song
 - Benutzer geben Constraints an
 - Lieder berücksichtigen Popularität
 - Machine-Learning trainierte Modelle



Visualisation of Music Libraries

- Benutzer benötigen Unterstützung beim Stöbern in ihren Sammlungen
- Klassische Listen sind unübersichtlich
- Ansätze:
 - Visualisierung durch 3D Modelle
 - Darstellung von Features
 - Gruppierung von ähnlichen Liedern





Themenvorgabe

Fragen?

Marko Harasic
harasic@inf.fu-berlin.de

Themen und Termine (1)

20.10.2016	Vorstellung und Verteilung der Themen	Harasic
27.10.2016	Hinweise zur Gestaltung von Referaten	Harasic
03.11.2016	Similarity by Audio-Signals (MFCC)	
10.11.2016	Similarity by Lyrics (PLSA)	
17.11.2016	Similarity by Meta-Data (Webmining)	
01.12.2016	Mood Detection	
08.12.2016	Genre determination	
15.12.2016	Recognition of Coversongs (LSH)	
22.12.2016	entfällt (akad. Weihnachstferien)	

Themen und Termine (2)

29.12.2016	entfällt (akad. Weihnachstferien)	
05.01.2017	Acoustic Fingerprinting / Shazam	
12.01.2017	Query by Humming	
19.01.2017	Matrix Factorisation / Spotify	
26.01.2017	Context-Aware Recommender Systems	
02.02.2016	Automatic Playlist Generation	
09.02.2016	Visualisation of Music Libraries	
16.02.2016	Abschluss, Rückblick, Feedback	Harasic