Department of Computational Perception
JOHANNES KEPLER UNIVERSITY LINZ

# Part IV: Personalization, Context-awareness, and Hybrid Methods

**Markus Schedl**
**Peter Knees**

{markus.schedl, peter.knees}@jku.at

**Department of Computational Perception**
**Johannes Kepler University (JKU)**
**Linz, Austria**

# Overview

1. Personalization and Context-awareness

2. Hybrid Methods

# Computational Factors Influencing Music Perception and Similarity

(Schedl et al., JIIS 2013)

**music content**

**Examples:**
- rhythm
- timbre
- melody
- harmony
- loudness

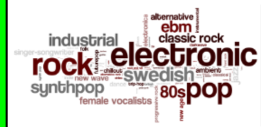
**music perception and similarity**

**Examples:**
- mood
- activities
- social context
- spatio-temporal context
- physiological aspects

**user context**

**Examples:**
- semantic labels
- song lyrics
- album cover artwork
- artist's background
- music video clips

**music context**




**Examples:**
- music preferences
- musical training
- musical experience
- demographics

**user properties**

Department of Computational Perception

# Computational Factors Influencing Music Perception and Similarity

(Schedl et al., JIIS 2013)

**music content**

**Examples:**
- rhythm
- timbre
- melody
- harmony
- loudness

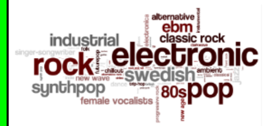
**Examples:**
- mood
- activities
- social context
- spatio-temporal context
- physiological aspects

**user context**

**personalized/context-aware methods:** typically extend music content or music context with a user-category

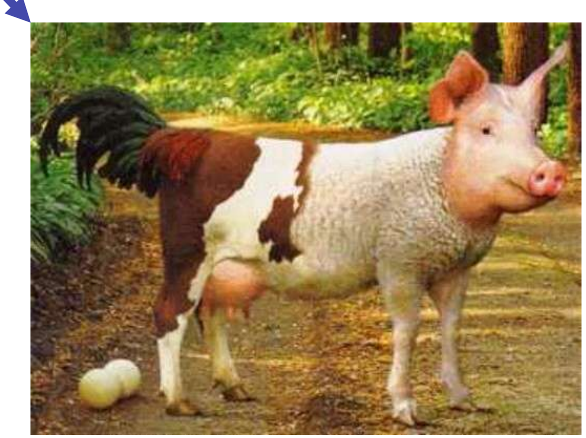
**Examples:**
- semantic labels
- song lyrics
- album cover artwork
- artist's background
- music video clips

**music context**




**Examples:**
- music preferences
- musical training
- musical experience
- demographics

**user properties**



# Computational Factors Influencing Music Perception and Similarity

(Schedl et al., JIIS 2013)

**music content**

**Examples:**
- rhythm
- timbre
- melody
- harmony
- loudness

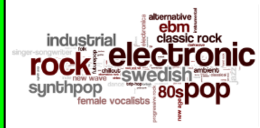
**Examples:**
- mood
- activities
- social context
- spatio-temporal context
- physiological aspects

**user context**

**hybrid methods:**
combine factors of at least two categories

**Examples:**
- semantic labels
- song lyrics
- album cover artwork
- artist's background
- music video clips

**music context**




**Examples:**
- music preferences
- musical training
- musical experience
- demographics

**user properties**

Department of Computational Perception

# Basic Categorization

- Personalized systems/methods
  - incorporate aspects of the *user properties*, i.e. static attributes
  - take into account music genre preference, music experience, age, etc.

- Context-aware systems/methods
  - incorporate aspects of the *user context*, i.e. dynamic aspects
    - **active user-awareness**: new user context is automatically incorporated into the system, adaptively changing its behavior
    - **passive user-awareness**: application presents the new context to the user for later retrieval/incorporation

Department of Computational Perception

# Typical Features used in CA

- Temporal and spatial features
  - temporal: weekday, time of day, season, month, etc.
  - spatial: position (coordinates), location (country, city, district; home, office)
- Physiological features
  - heart rate, pace, body temperature, skin conductance, etc.
  - application scenarios: music therapy [Liu, Rautenberg; 2009], sport trainer [Elliot, Tomlinson; 2006] [Moens et al.; 2010]
    - achieving and maintaining a healthy heart rate in music therapy
    - adapting music to pace of runner
    - selecting music suited to stimulate a particular running behavior, reach a performance level, or fit a training program

Department of Computational Perception

# Gathering the User Context

- Implicit
  - sensors: GPS, heart rate, accelerometer, pressure, light intensity, environmental noise level (now available in abundance through smart phones)
  - derived features: location + time → weather
  - learned features (via ML): accelerometer, speed → user activity

- Explicit
  - via user involvement/feedback
  - e.g., mood, activity, item ratings, skipping behavior [Pampalk et al.; 2005]

# Overview

1. Personalization and Context-awareness

2. Hybrid Methods
   – Music playlist generation using music content and music context
   – *#nowplaying* approaches: music taste analysis, browsing the world of music on the microblogosphere
   – Geospatial music recommendation
   – User-Aware music recommendation on smart phones
   – Matching places of interest and music

Department of Computational Perception

# Music playlist generation using music content and music context

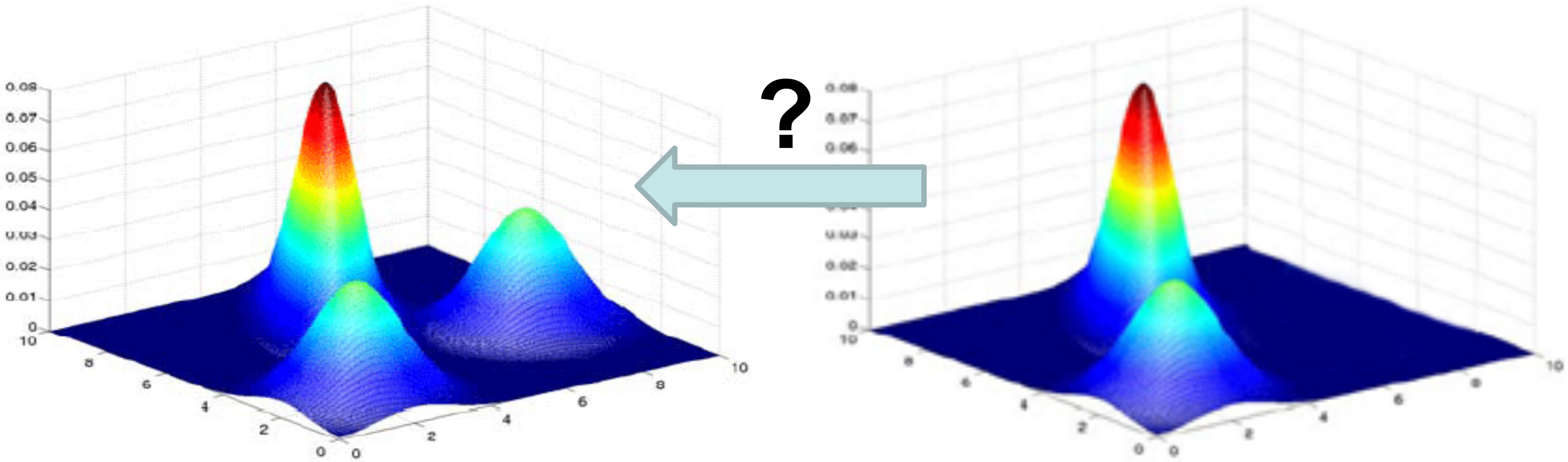
(Knees et al.; 2006)

- Idea: combine music content + music context features to improve and speed up playlist generation
- Application scenario: “The Wheel” – create a circular playlist containing all tracks in a user’s collection (consecutive tracks as similar as possible)
- Approach: use web features to confine search for similar songs (carried out on music content features)

# Music playlist generation using music content and music context

(Knees et al.; 2006)

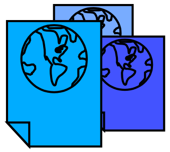

- Audio/content features: ♪
  - compute Mel-Frequency Cepstral Coefficients (MFCC)
  - model song's distribution of MFCCs via Gaussian Mixture Models (GMM)
  - estimate similarity between two songs *A* and *B* by sampling points from *A*'s GMM and computing probability that points "belong to" GMM of *B*



# Music playlist generation using music content and music context

(Knees et al.; 2006)


- Web/music context features:



  - query Google for `[artist "music"]`
  - fetch 50 top-ranked web pages
  - remove HTML, stop words, and infrequent terms
  - for each artist's virtual document, compute tf-idf vectors:

$$w_{ta} = \begin{cases} (1 + \log_2 tf_{ta}) \log_2 \frac{N}{df_t} & \text{if } tf_{ta} > 0 \\ 0 & \text{otherwise} \end{cases}$$

  - perform cosine normalization (different document length!)

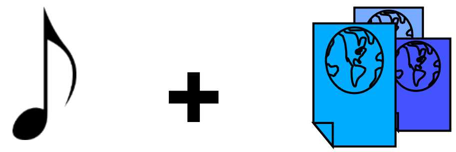
# Music playlist generation using music content and music context

(Knees et al.; 2006)


We computed so far…

- *similarities* based on music content (song level)
- *feature vectors* (tf-idf) from web content (artist level)

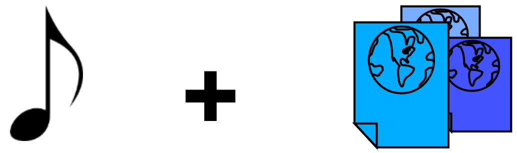
How to combine the two?



- adapt the content similarities according to web similarity
- penalize transitions (decrease similarity) between songs whose artists are dissimilar in terms of web features

# Music playlist generation using music content and music context

(Knees et al.; 2006)




To obtain the final, hybrid similarity measure:

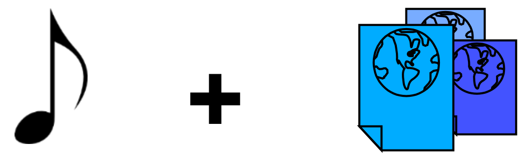
| | | | | |
|---|---|---|---|---|
| Folk−Rock(4)<br>Rap(4)<br>Jazz(1)<br>Punk−Rock(1) | Electronica(2) | Electronica(5) | Electronica(1) | Electronica(16)<br>Acid Jazz(1) |
| Folk−Rock(1)<br>Italian(1) | Electronica(1) | Acid Jazz(1) | | Acid Jazz(1)<br>Electronica(1) |
| Italian(3)<br>Electronica(1) | | Reggae(2)<br>Italian(1) | | Rap(2)<br>A Cappella(1)<br>Acid Jazz(1)<br>Electronica(1) |
| Punk−Rock(4)<br>Electronica(1) | Rap(4) | | Blues(1) | Jazz(3) |
| Electronica(12)<br>Punk−Rock(1) | Rap(1)<br>Electronica(1) | Celtic(2)<br>Reggae(1) | Celtic(3)<br>A Cappella(1) | Jazz(5)<br>Bossa Nova(4)<br>Blues(3)<br>A Cappella(2)<br>Rap(1) |

train Self-Organizing Map (SOM) on artist web features

Department of Computational Perception

# Music playlist generation using music content and music context

(Knees et al.; 2006)




To obtain the final, hybrid similarity measure:

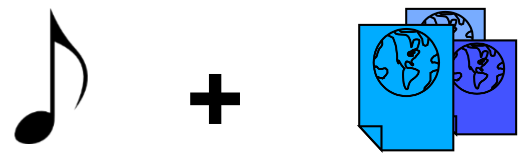
| | | | | |
|---|---|---|---|---|
| Folk–Rock(4)<br>Rap(4)<br>Jazz(1)<br>Punk–Rock(1) | Electronica(2) | Electronica(5) | Electronica(1) | Electronica(16)<br>Acid Jazz(1) |
| Folk–Rock(1)<br>Italian(1) | Electronica(1) | Acid Jazz(1) | | Acid Jazz(1)<br>Electronica(1) |
| Italian(3)<br>Electronica(1) | | Reggae(2)<br>Italian(1) | | Rap(2)<br>A Cappella(1)<br>Acid Jazz(1)<br>Electronica(1) |
| Punk–Rock(4)<br>Electronica(1) | Rap(4) | | Blues(1) | Jazz(3) |
| Electronica(12)<br>Punk–Rock(1) | Rap(1)<br>Electronica(1) | Celtic(2)<br>Reggae(1) | Celtic(3)<br>A Cappella(1) | Jazz(5)<br>Bossa Nova(4)<br>Blues(3)<br>A Cappella(2)<br>Rap(1) |

– set to zero content-based similarity of songs by dissimilar artists (according to position in SOM)
– i.e., when creating playlists, consider as potential next track only songs by artists close together on SOM

Department of Computational Perception

# Music playlist generation using music content and music context

(Knees et al.; 2006)




To obtain the final, hy

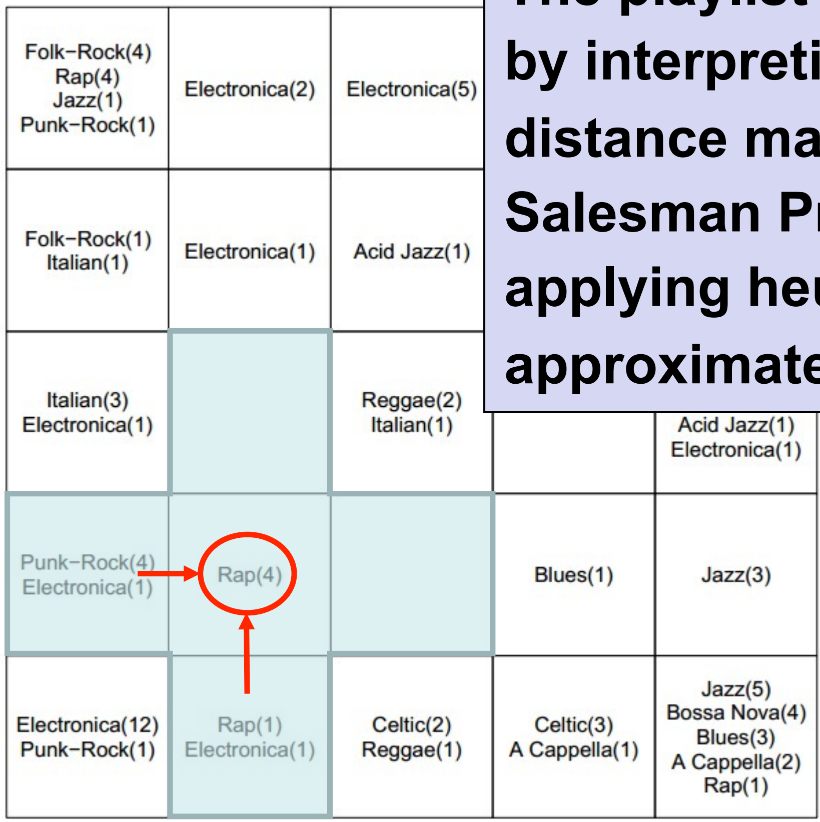
ased

y

cording to

aylists,

consider as potential next track only songs by artists close together on SOM




**The playlist is eventually created by interpreting the adapted distance matrix as Traveling Salesman Problem (TSP) and applying heuristics to approximate a solution.**

Department of Computational Perception

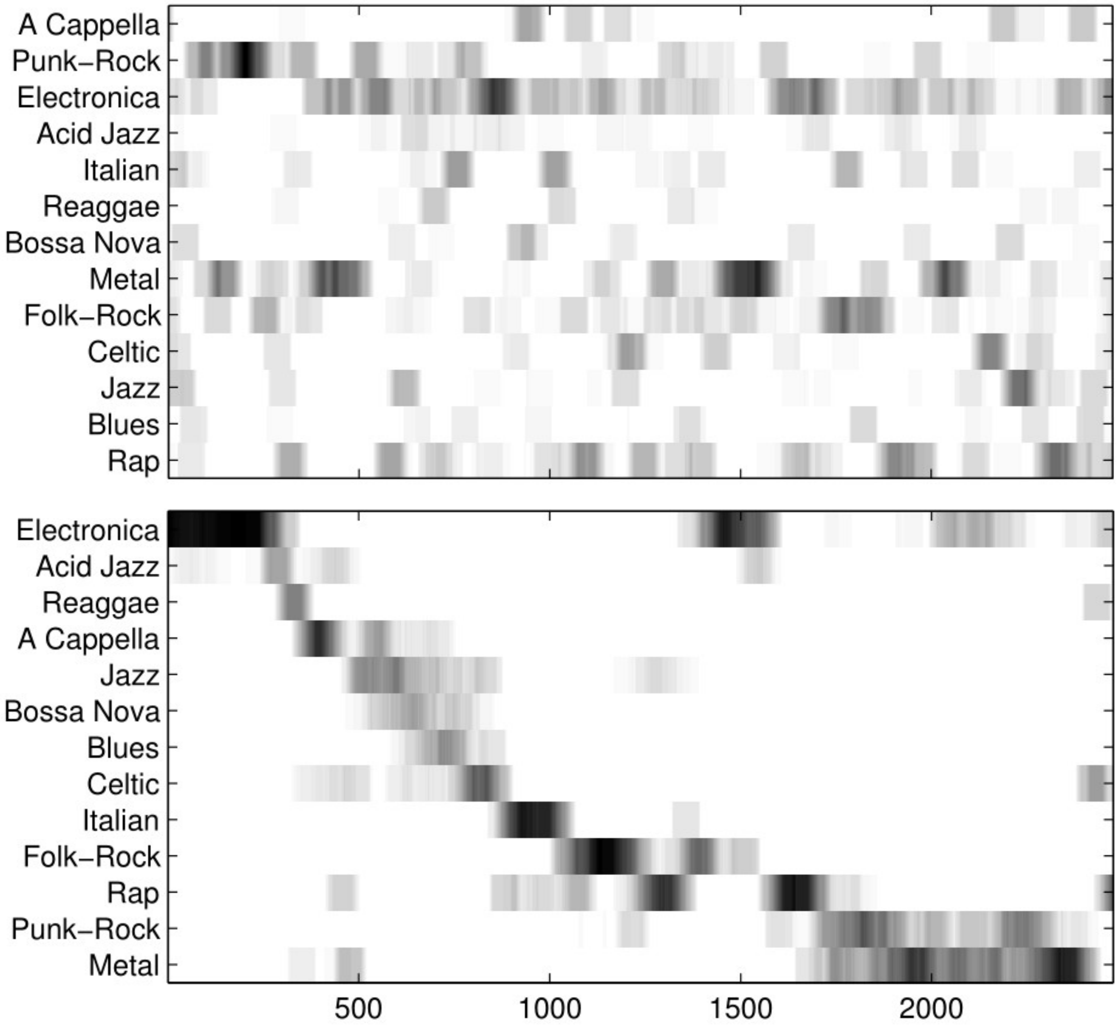
# Music playlist generation using music content and music context

(Knees et al.; 2006)


- Evaluation:
  - dataset: 2,545 tracks from 13 genres, 103 artists
  - performance measure: consistency of playlists (for each track, how many of its 75 consecutive tracks belong to a certain genre)

# Music playlist generation using music content and music context

(Knees et al.; 2006)




music content similarity only

hybrid approach


Department of Computational Perception

# #nowplaying approaches: Basics

(Schedl, ECIR 2013)

- Extract listening events from microblogs

(a) Filter Twitter stream (#nowplaying, #itunes, #np, …)
(b) Multi-level, rule-based analysis (artists/songs) to find relevant tweets (MusicBrainz)
(c) Last.fm, Freebase, Allmusic, Yahoo! PlaceFinder to annotate tweets

„Alice Cooper“
„BB King“
„Prince“
„Metallica“
…

{"id_str":"142338125895696385","place":null,"text":"**#NowPlaying Christmas Tree-Lady Gaga**","in_reply_to_user_id":null,"favorited":false,"geo":null,"retweet_count":0,"in_reply_to_screen_name":null,"in_reply_to_status_id_str":null,"source":"web","retweeted":false,"in_reply_to_user_id_str":null,"coordinates":null,"created_at":"Thu Dec 01 20:23:48 +0000 2011","in_reply_to_status_id":null,"contributors":null,"user":{"id_str":"20209983","profile_link_color":"2caba5","screen_name":"tamse77","follow_request_sent":null,"geo_enabled":false,"favourites_count":26,"location":"Maryland ","following":null,"verified":false,"profile_background_color":"e80e0e","show_all_inline_media":true,"profile_background_tile":true,"followers_count":309,"profile_image_url":"http:\/\/a1.twimg.com\/profile_images\/1647613274\/392960_10150559294659517_793614516_11700077_1689597400_n_normal.jpg","description":"being awesome since 1990. ","is_translator":false,"profile_background_image_url_https":"https:\/\/si0.twimg.com\/profile_background_images\/359728130\/frames.gif","friends_count":148,"profile_sidebar_fill_color":"ffffff","default_profile":false,"listed_count":3,"time_zone":"Central Time (US & Canada)","contributors_enabled":false,"created_at":"Fri Feb 06 01:51:10 +0000 2009","profile_sidebar_border_color":"f5f8ff","protected":false,"notifications":null,"profile_use_background_image":true,"name":"Katie","default_profile_image":false,"statuses_count":22172,"profile_text_color":"615d61","url":null,"profile_image_url_https":"https:\/\/si0.twimg.com\/profile_images\/1647613274\/392960_10150559294659517_793614516_11700077_1689597400_n_normal.jpg","id":20209983,"lang":"en","profile_background_image_url":"http:\/\/a2.twimg.com\/profile_background_images\/359728130\/frames.gif","utc_offset":-21600},"truncated":false,"id":142338125895696385,"entities":{"hashtags":[{"text":"NowPlaying","indices":[0,11]}],"urls":[],"user_mentions":[]}}

Department of
Computational
Perception

# #nowplaying approaches: Basics

(Schedl, ECIR 2013)

- Annotate identified listening events and create a database

{"id_str":"142338125895696385","place":null,"text":"**#NowPlaying Christmas Tree- Lady Gaga**","in_reply_to_user_id":null,"favorited":false,"geo":null,"retweet_count":0,"in_reply_to_screen_name":null,"in_reply_to_status_id_str":null,"source":"web","retweeted":false,"in_reply_to_user_id_str":null,"coordinates":null,"created_at":"Thu Dec 01 20:23:48 +0000 2011","in_reply_to_status_id":null,"contributors":null,"user":{"id_str":"20209983","profile_link_color":"2caba5","screen_name":"tamse77","follow_request_sent":null,"geo_enabled":false,"favourites_count":26,"location":"Maryland ","following":null,"verified":false,"profile_background_color":"e80e0e","show_all_inline_media":true,"profile_background_tile":true,"followers_count":309,"profile_image_url":"http:\/\/a1.twimg.com\/profile_images\/1647613274\/392960_10150559294659517_793614516_11700077_1689597400_n_normal.jpg","description":"being awesome since 1990. ","is_translator":false,"profile_background_image_url_https":"https:\/\/si0.twimg.com\/profile_background_images\/359728130\/frames.gif","friends_count":148,"profile_sidebar_fill_color":"ffffff","default_profile":false,"listed_count":3,"time_zone":"Central Time (US & Canada)","contributors_enabled":false,"created_at":"Fri Feb 06 01:51:10 +0000 2009","profile_sidebar_border_color":"f5f8ff","protected":false,"notifications":null,"profile_use_background_image":true,"name":"Katie","default_profile_image":false,"statuses_count":22172,"profile_text_color":"615d61","url":null,"profile_image_url_https":"https:\/\/si0.twimg.com\/profile_images\/1647613274\/392960_10150559294659517_793614516_11700077_1689597400_n_normal.jpg","id":20209983,"lang":"en","profile_background_image_url":"http:\/\/a2.twimg.com\/profile_background_images\/359728130\/frames.gif","utc_offset":-21600},"truncated":false,"id":142338125895696385,"entities":{"hashtags":[{"text":"NowPlaying","indices":[0,11]}],"urls":[],"user_mentions":[]}}

```
134243700380401664  127821914  11  2  106.83  -6.23  1  1  202085  3529910  0  1 ...
134243869201154048  174194590  11  2  -0.142  51.52  2  2  330061  5762915  1  0 ...
```

**twitter-id user-id month weekday longitude latitude country-id city-id artist-id track-id <tag-ids>**

"MusicMicro" dataset available:
http://www.cp.jku.at/datasets/musicmicro

Department of Computational Perception

# Some statistics on spatial distribution

| #nowplaying | | #itunes | |
|---|---|---|---|
| country | tweets | country | tweets |
| Brazil | 725,389 | USA | 78,460 |
| USA | 673,839 | Japan | 30,932 |
| Japan | 458,558 | Mexico | 23,047 |
| Mexico | 419,584 | Brazil | 16,390 |
| Indonesia | 284,082 | UK | 15,134 |
| South Korea | 251,132 | Canada | 11,266 |
| China | 183,178 | South Korea | 8,652 |
| UK | 128,744 | Australia | 5,119 |
| Netherlands | 121,134 | China | 4,492 |
| Venezuela | 110,336 | Germany | 3,157 |

most active countries

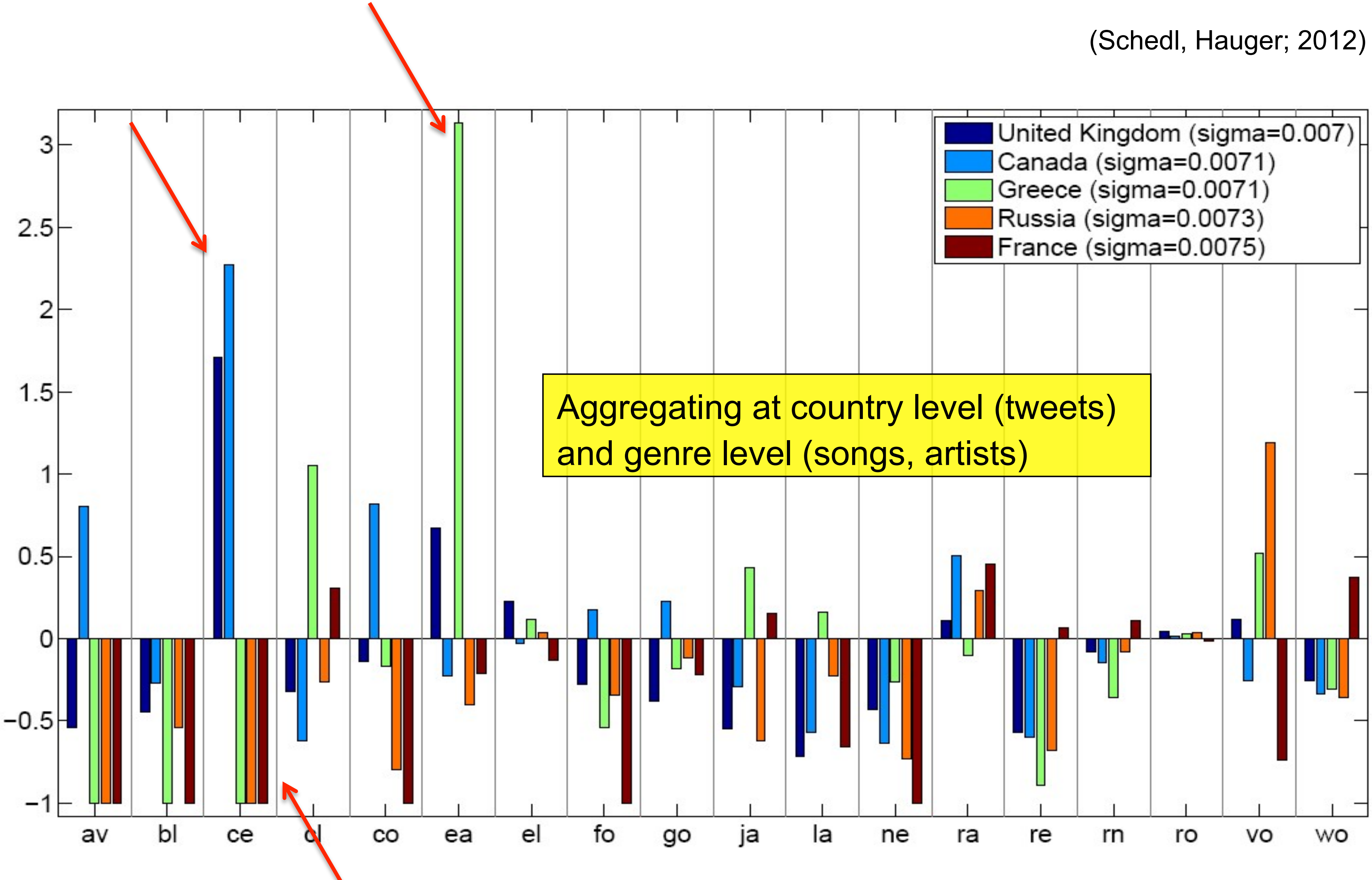
# Some statistics on artist distribution

most frequently listened artists

| #nowplaying | | #itunes | |
|---|---|---|---|
| artist | tweets | artist | tweets |
| Paramore | 9,066 | The Beatles | 939 |
| Drake | 7,697 | Daft Punk | 683 |
| Katy Perry | 6,998 | Britney Spears | 567 |
| Bruno Mars | 6,932 | Adele | 462 |
| Lady Gaga | 6,919 | Coldplay | 428 |
| Coldplay | 6,434 | Bruno Mars | 416 |
| Eminem | 6,352 | Katy Perry | 374 |
| Rihanna | 6,038 | The Black Eyes Peas | 373 |
| Taylor Swift | 5,844 | Kanye West | 367 |
| Usher | 5,445 | Lady Gaga | 358 |
| Muse | 5,383 | Avril Lavigne | 308 |
| Justin Bieber | 5,028 | Arcade Fire | 299 |
| The Beatles | 4,579 | Radiohead | 266 |
| Michael Jackson | 4,476 | Kings of Leon | 240 |
| Linkin Park | 4,285 | Duran Duran | 238 |
| Oasis | 4,190 | Michael Jackson | 229 |
| Kanye West | 4,013 | Linkin Park | 228 |
| Chris Brown | 3,943 | Eminem | 211 |
| Avril Lavigne | 3,780 | Muse | 209 |
| Radiohead | 3,756 | The Black Keys | 203 |

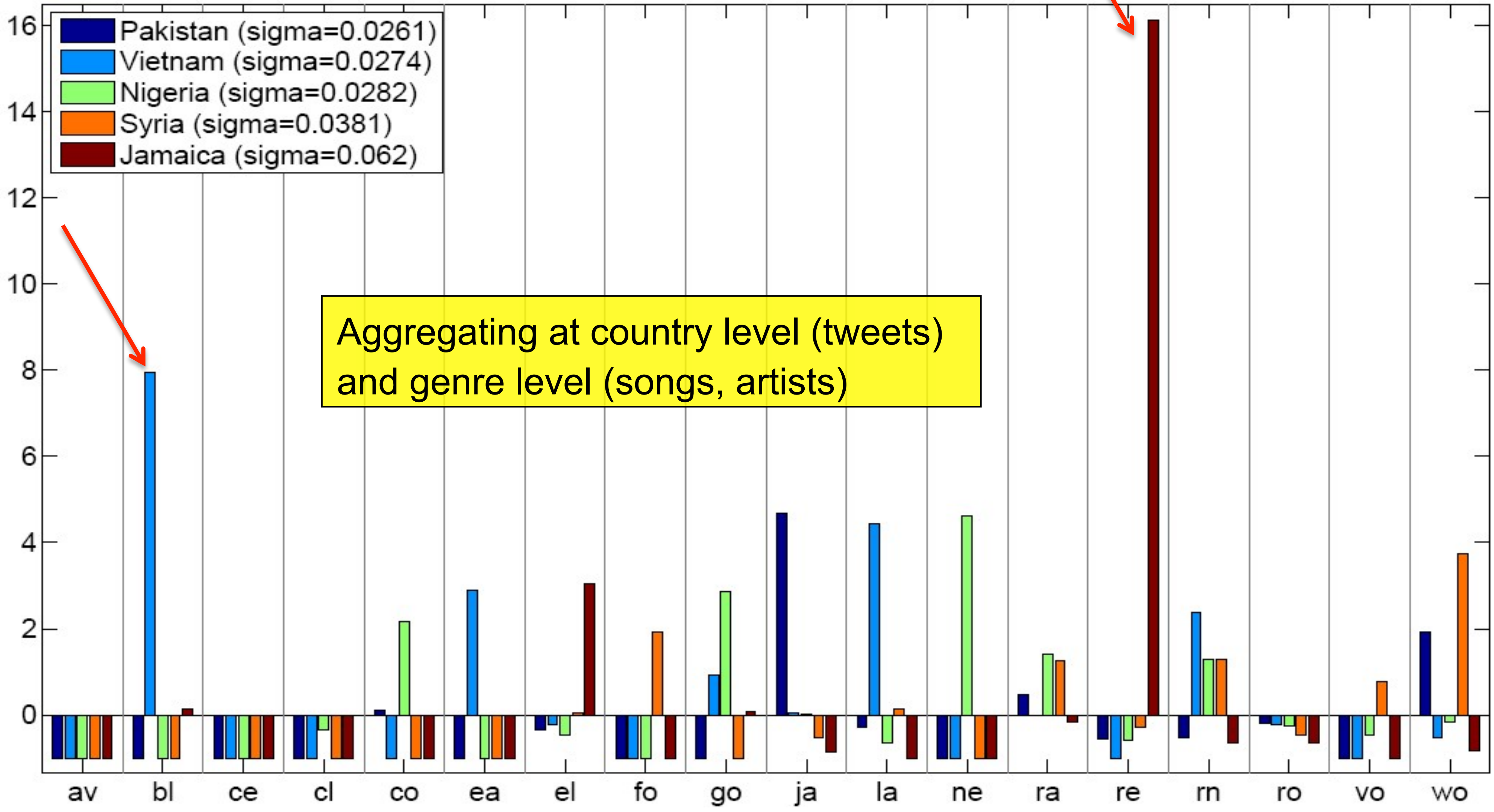
# #nowplaying approaches: Music taste analysis Most mainstreamy countries

(Schedl, Hauger; 2012)



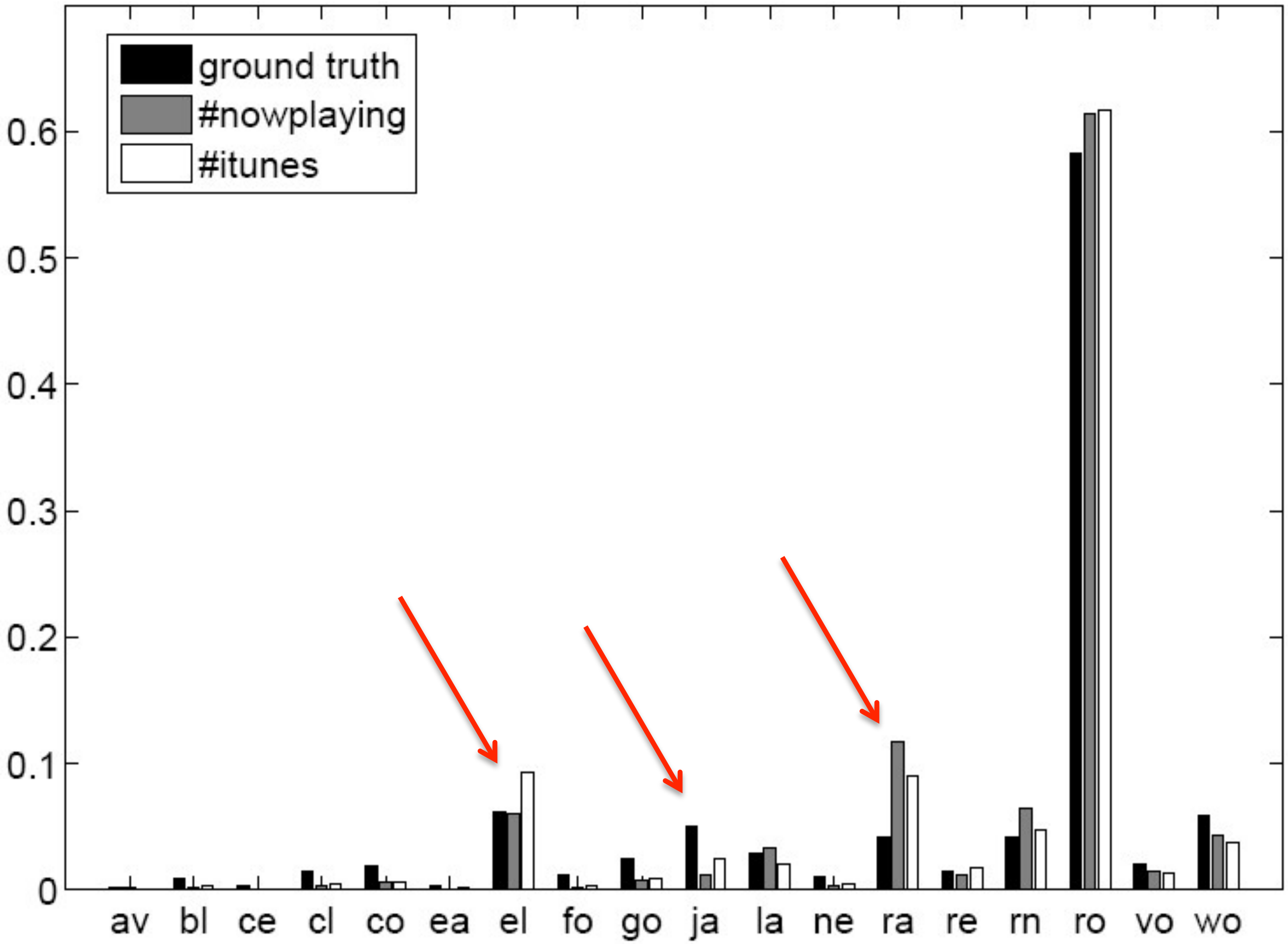
# #nowplaying approaches: Music taste analysis
# Least mainstreamy countries

(Schedl, Hauger; 2012)



# #nowplaying approaches: Music taste analysis
# Usage of specific products

(Schedl, Hauger; 2012)

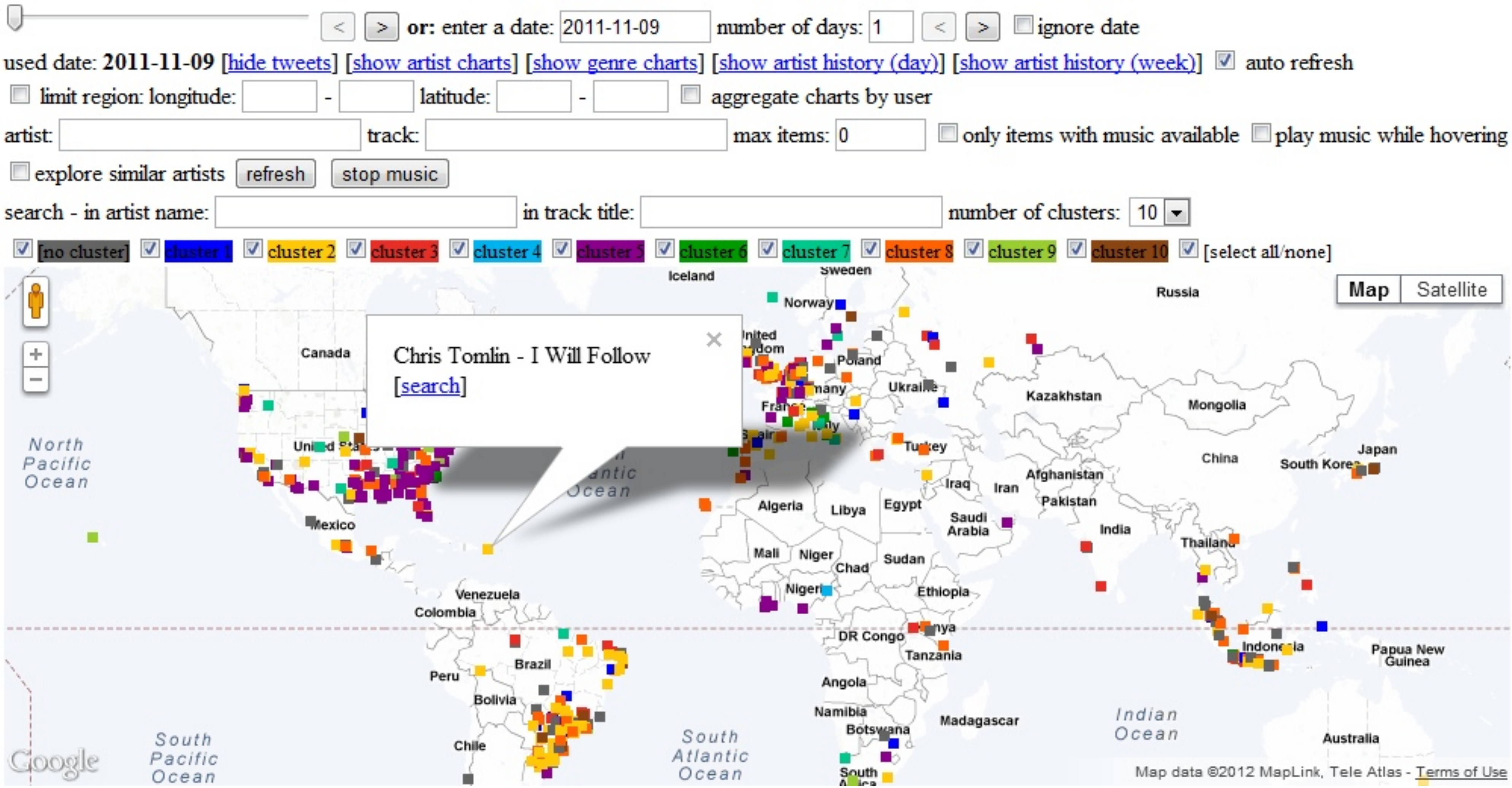


# #nowplaying approaches: Browsing the world of music on the microblogosphere

- *"MusicTweetMap"*
  - Info: http://www.cp.jku.at/projects/MusicTweetMap
  - App: http://songwitch.cp.jku.at/cp/maps/tweetMapOverlay.php
  - Features:
    - browse by specific date/day or time range
    - show similar artists (based on co-occurrences in tweets)
    - restrict to country, state, city, and longitude/latitude coordinates
    - metadata-based search (artist, track)
    - clustering based on Non-negative Matrix Factorization (NMF) on Last.fm tags → genres
    - artist charts, genre charts
    - artist histories on plays

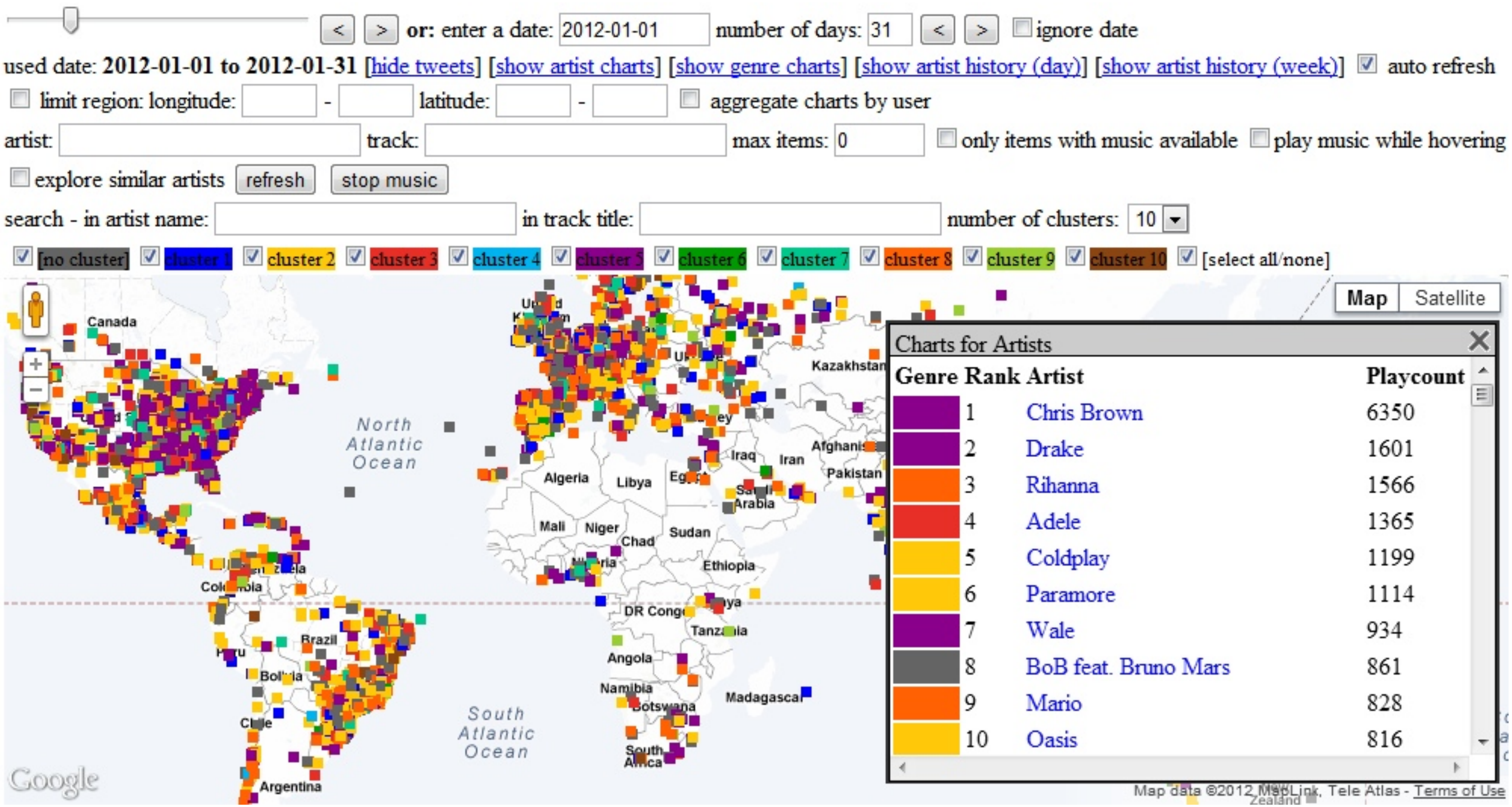
# #nowplaying approaches: Browsing the world of music on the microblogosphere

Visualization and browsing of geospatial music taste



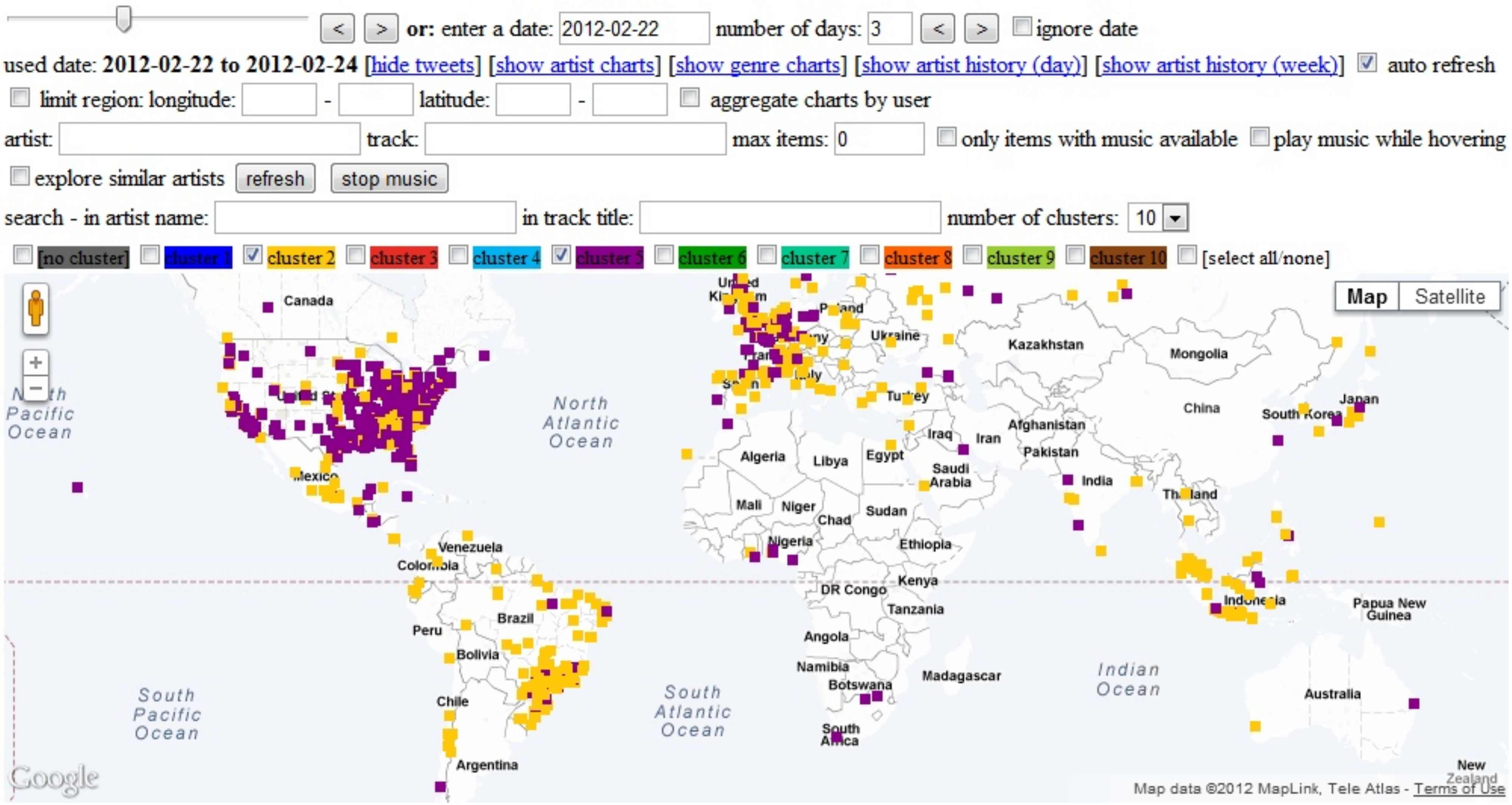
# #nowplaying approaches: Browsing the world of music on the microblogosphere

Investigating geospatial music taste: 1 month



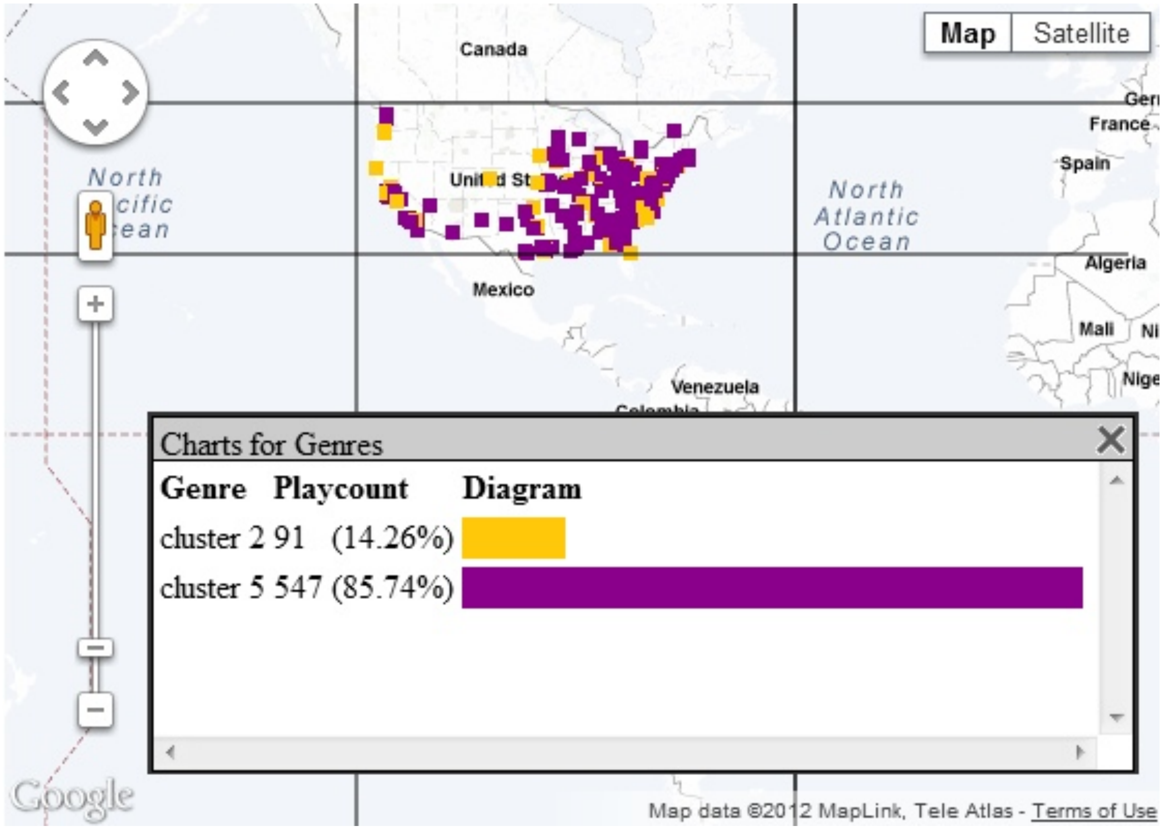
# #nowplaying approaches: Browsing the world of music on the microblogosphere

Geospatial music taste: “hip-hop” vs. “rock”



# #nowplaying approaches: Browsing the world of music on the microblogosphere

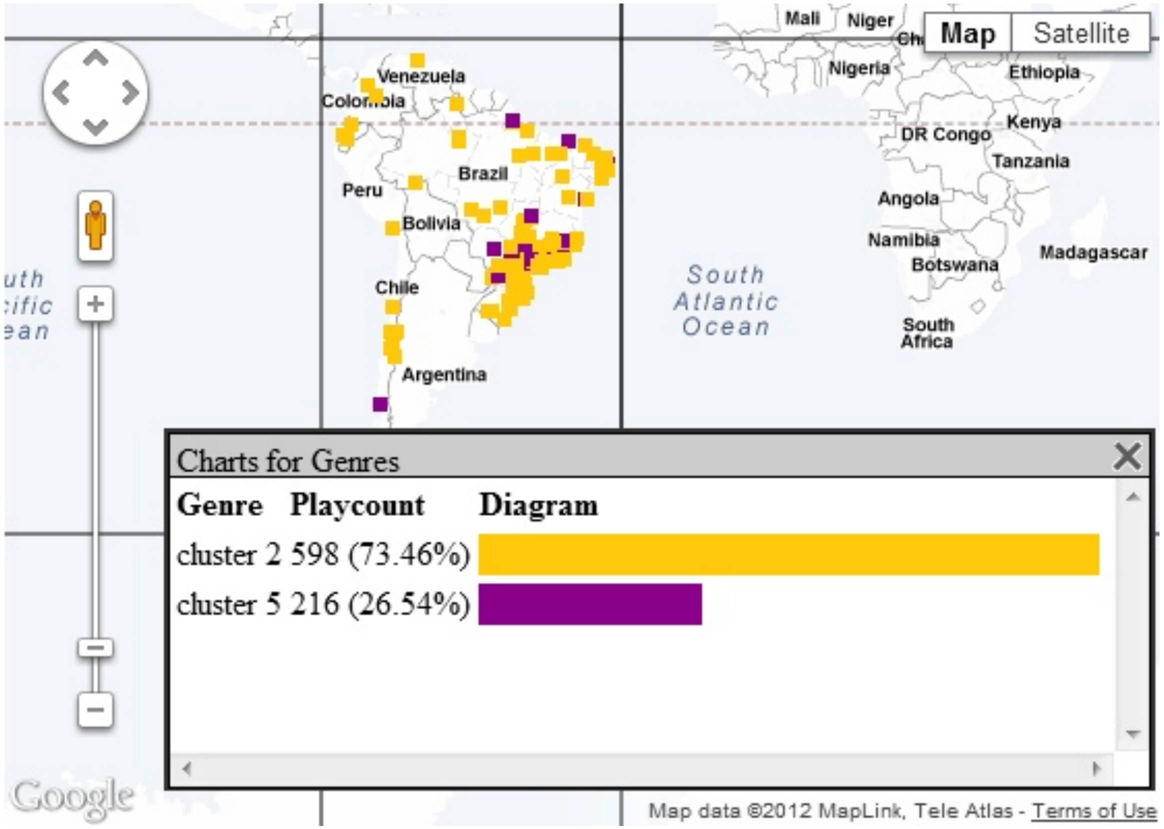
## Geospatial music taste: “hip-hop” vs. “rock” (USA)




Department of Computational Perception

# #nowplaying approaches: Browsing the world of music on the microblogosphere

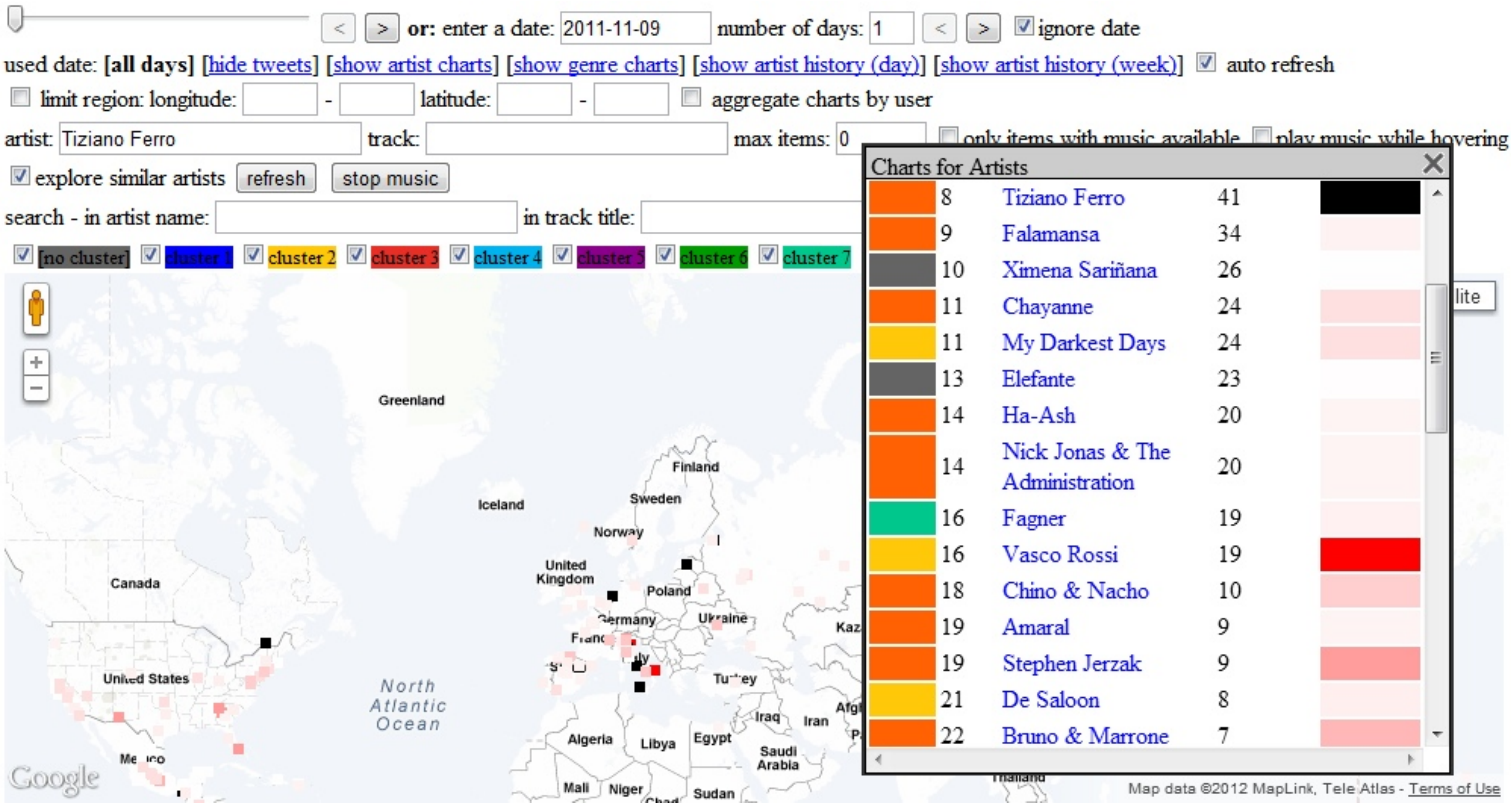
Geospatial music taste: “hip-hop” vs. “rock” (South America)




Department of Computational Perception

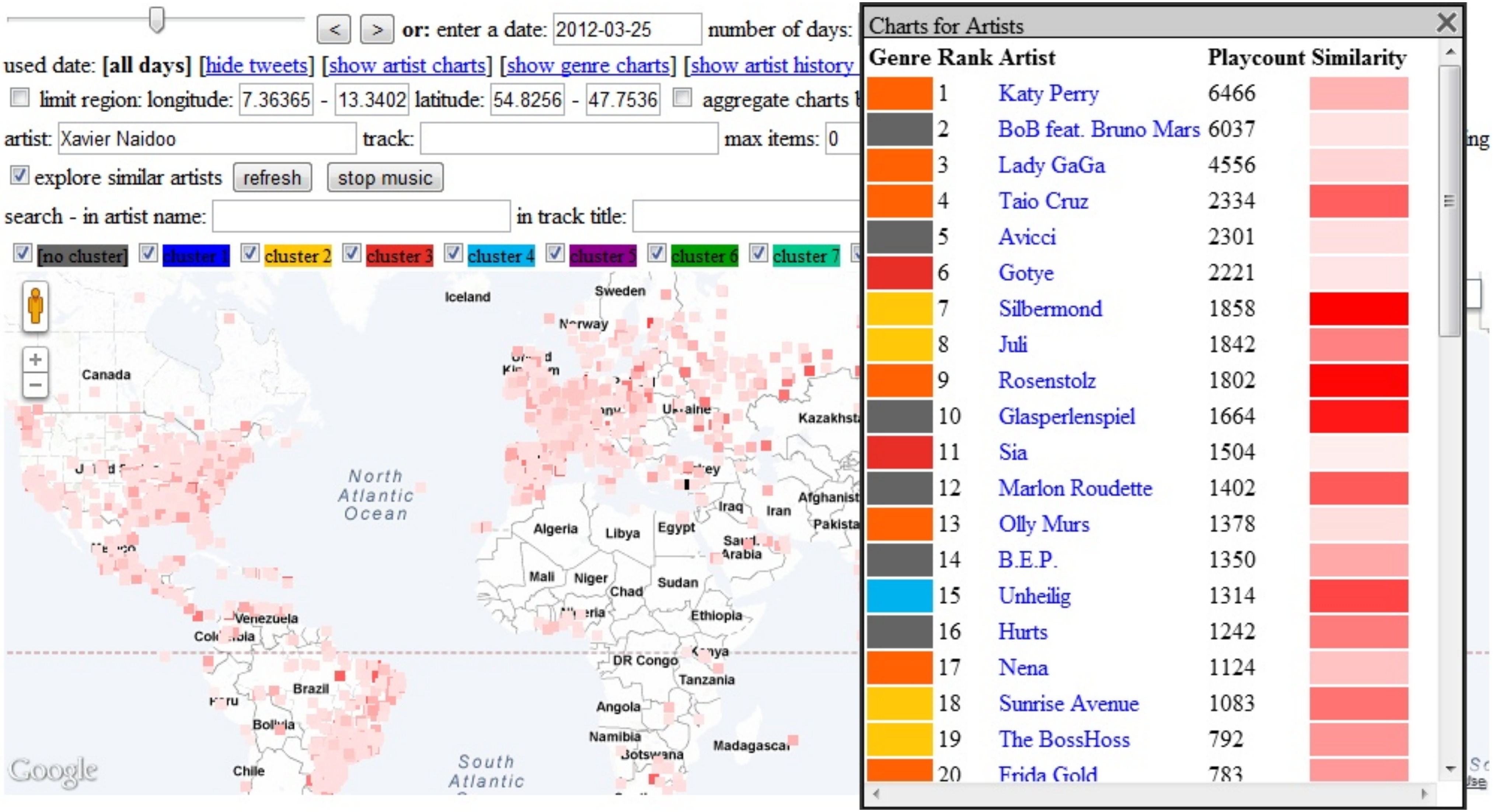
# #nowplaying approaches: Browsing the world of music on the microblogosphere

## Exploring similar artists: Example “Tiziano Ferro”



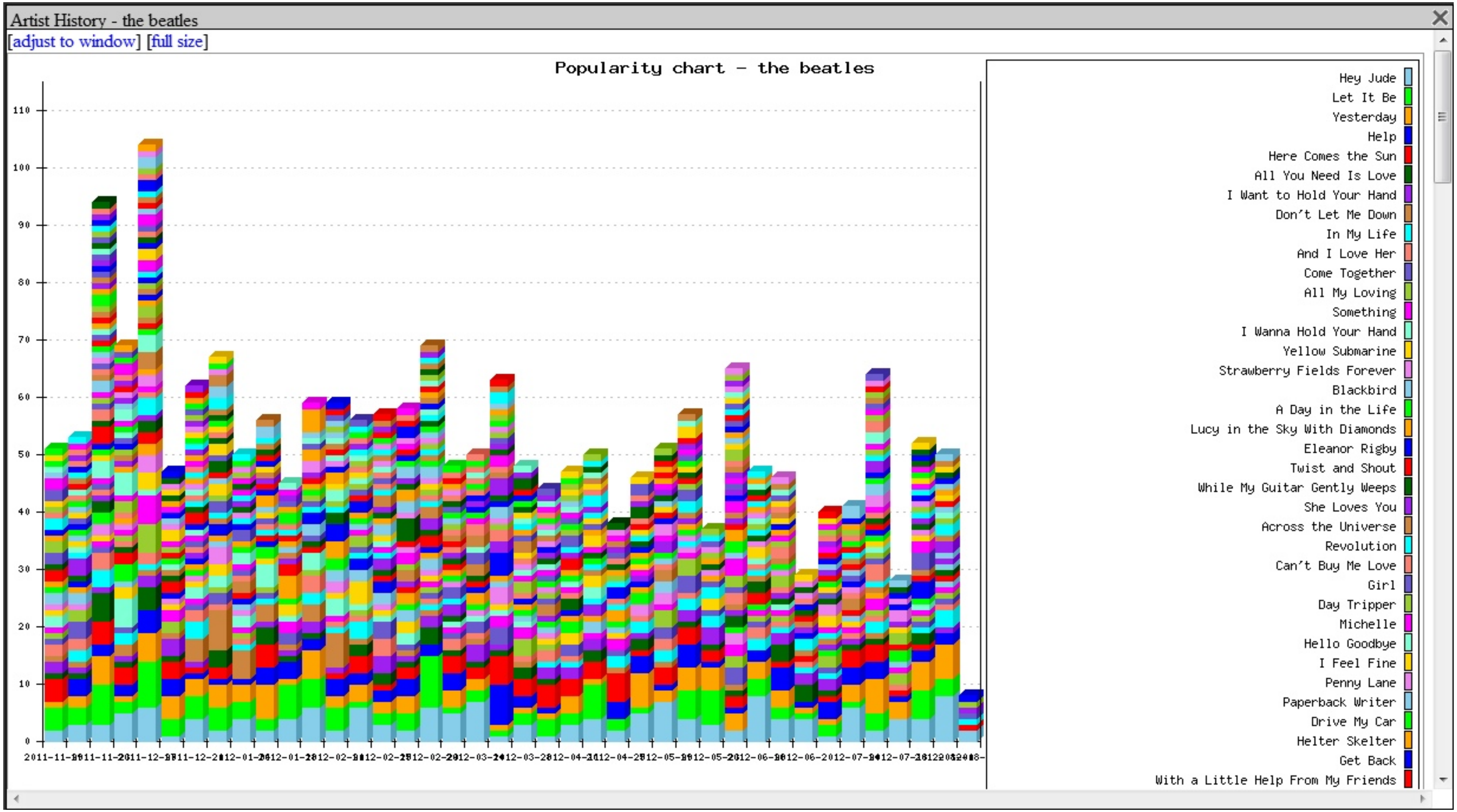
# #nowplaying approaches: Browsing the world of music on the microblogosphere

## Exploring similar artists: Example “Xavier Naidoo”



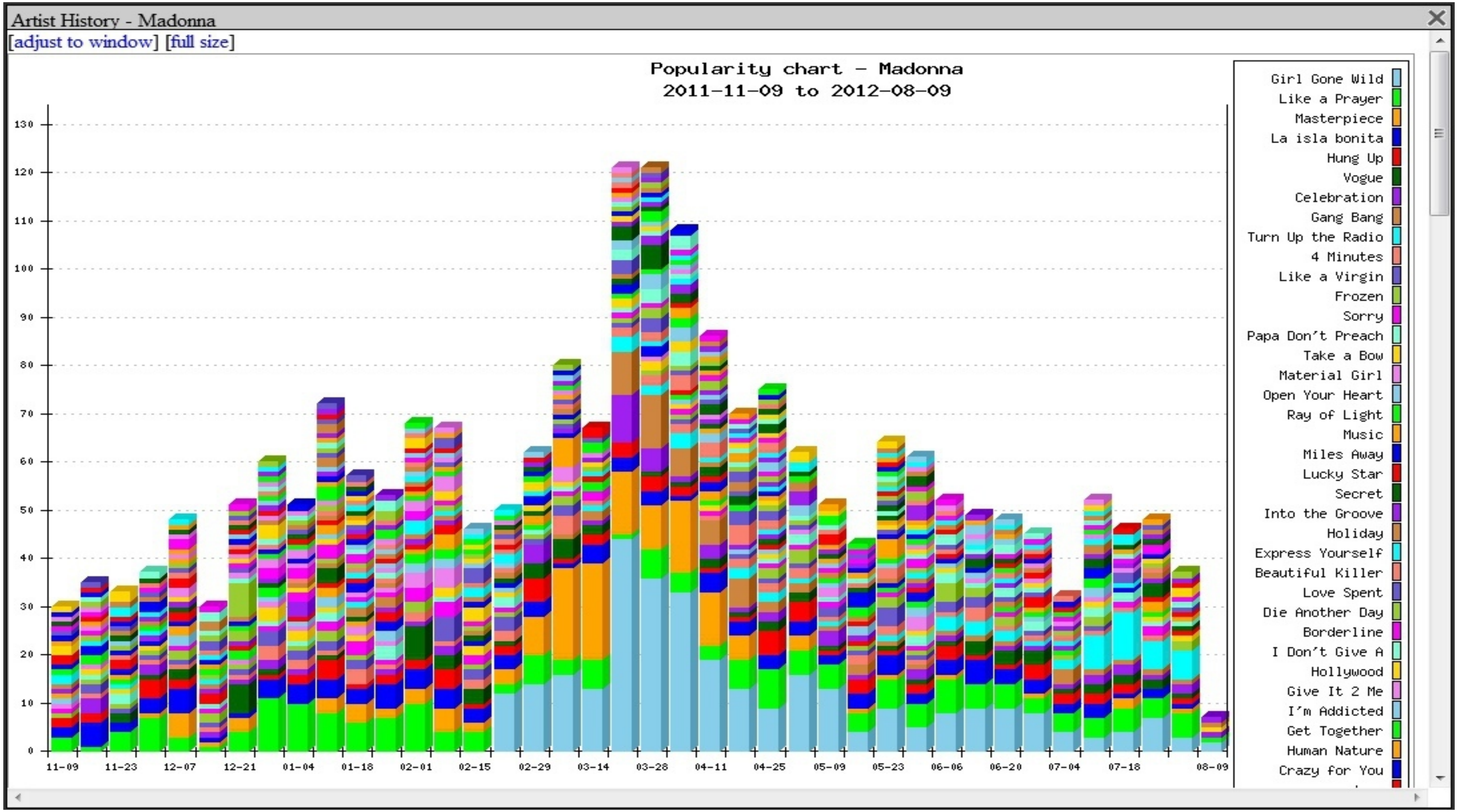
# #nowplaying approaches: Browsing the world of music on the microblogosphere

## Exploring music trends: Example “The Beatles”



# #nowplaying approaches: Browsing the world of music on the microblogosphere

## Exploring music trends: Example "Madonna"



# Geospatial Music Recommendation

(Schedl, Schnitzer; SIGIR 2013)

- Combining music content + music context features
  - audio features: PS09 award-winning feature extractors (rhythm and timbre)
  - text/web: TFIDF-weighted artist profiles from artist-related web pages
- Using collection of geo-located music tweets (cf. (Schedl; ECIR 2013))
- Aims:

  (i) determining ideal combination of music content and –context

  (ii) ameliorate music recommendation by user's location information

# Ideal combination of music content and context

(Schedl, Schnitzer; SIGIR 2013)

| $\xi$ | $K=1$ | $K=3$ | $K=5$ |
|---|---|---|---|
| **web only – 0.00** | **.5829** | **.5753** | **.5774** |
| .05 | .6421 | .6280 | .6257 |
| **.15** | .6432 | **.6286** | **.6261** |
| .25 | **.6433** | .6275 | .6258 |
| .35 | .6430 | .6275 | .6257 |
| .45 | .6408 | .6266 | .6252 |
| .55 | .6394 | .6259 | .6244 |
| .65 | .6379 | .6255 | .6232 |
| .75 | .6368 | .6234 | .6221 |
| .85 | .6330 | .6202 | .6188 |
| .95 | .6215 | .6083 | .6059 |
| **audio only – 1.00** | **.5436** | **.5302** | **.5247** |

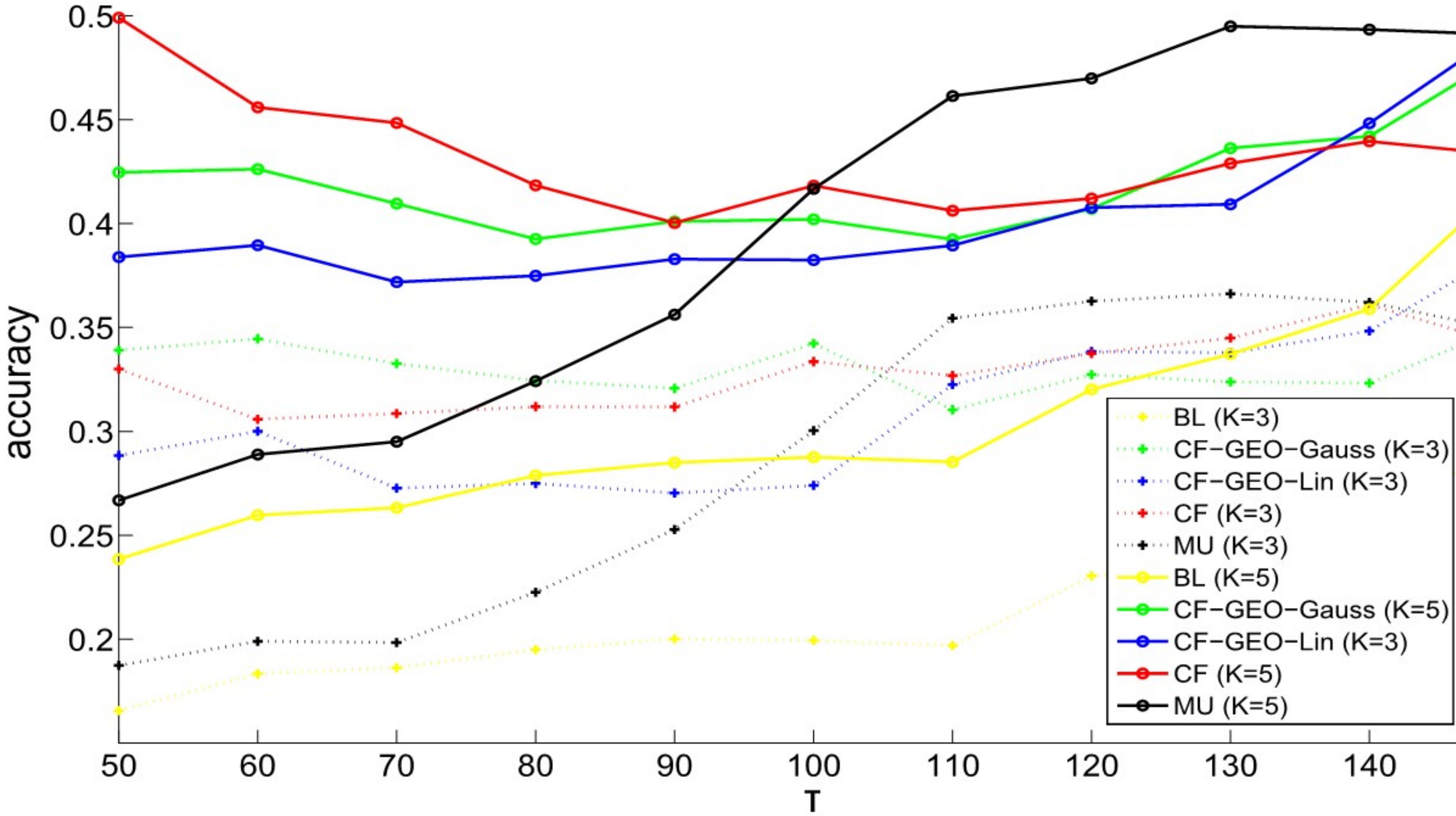
# Adding user context (different approaches)

(Schedl, Schnitzer; SIGIR 2013)

| **Abbreviation** | **Description** |
|---|---|
| BL | random baseline |
| MU | hybrid music model |
| CF | collaborative filtering model |
| CF-GEO-Lin | CF model: geospatial user weighting using linear spatial distances |
| CF-GEO-Gauss | CF model: geospatial user weighting weighting using a Gauss kernel |

(Schedl, Schnitzer; SIGIR 2013)

# Evaluation Results




T: minimum number of distinct artists a users must have listened to to be included


Department of
Computational
Perception

# User-Aware Music Recommendation on Smart Phones

(Breitschopf; 2013)

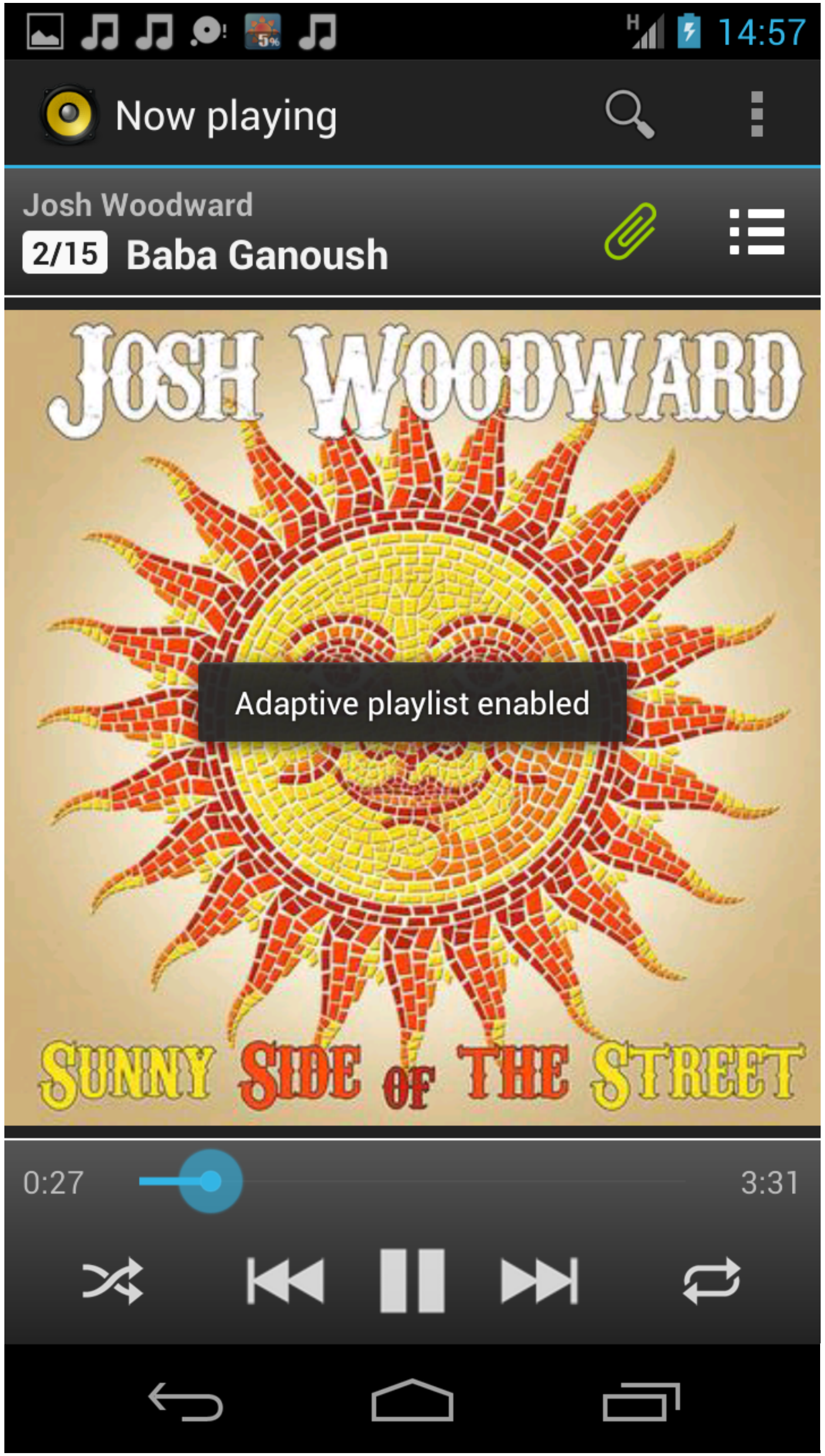
*"Mobile Music Genius"*: music player for the Android platform

- collecting user context data while playing
- adaptive system that learns user taste/preferences from implicit feedback (player interaction: play, skip, duration played, playlists, etc.)
- ultimate aim: dynamically and seamlessly update the user's playlist according to his/her current context

# Mobile Music Genius: Approach

*"Mobile Music Genius"*: music player for the Android platform

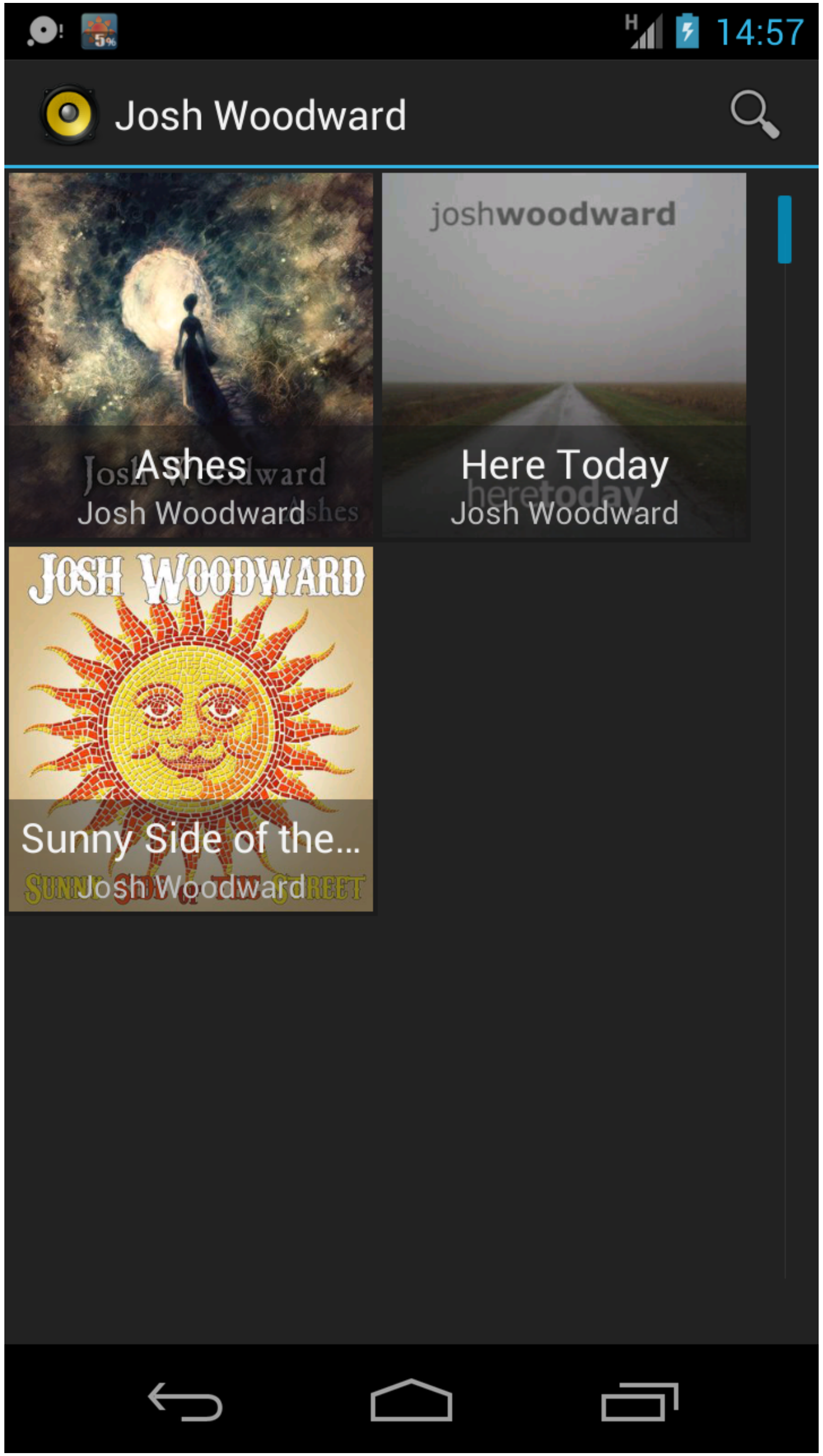
- standard, non-context-aware playlists are created using Last.fm tag features (weighted tag vectors on artists and tracks); cosine similarity between linear combination (of artist and track features) used for playlist generation
- learning and adapting a user model via relations
  {user context – music preference}
  on the level of genre, mood, artist, and song
- playlist is adapted when change in similarity between current user context and earlier user context is above threshold




# Mobile Music Genius

Music player in adaptive playlist generation mode

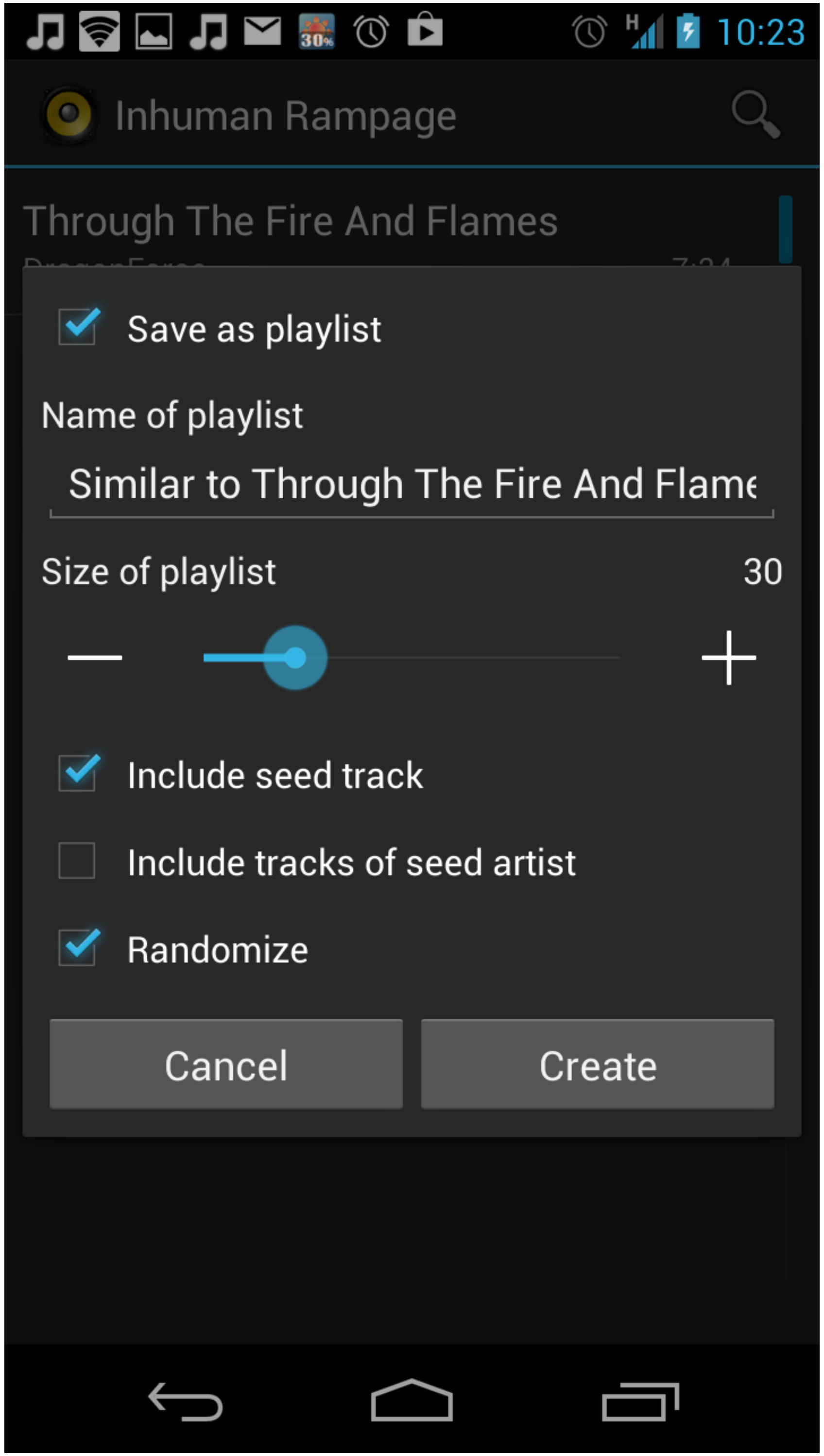
Department of Computational Perception




# Mobile Music Genius

Album browser
in cover view

Department of
Computational
Perception

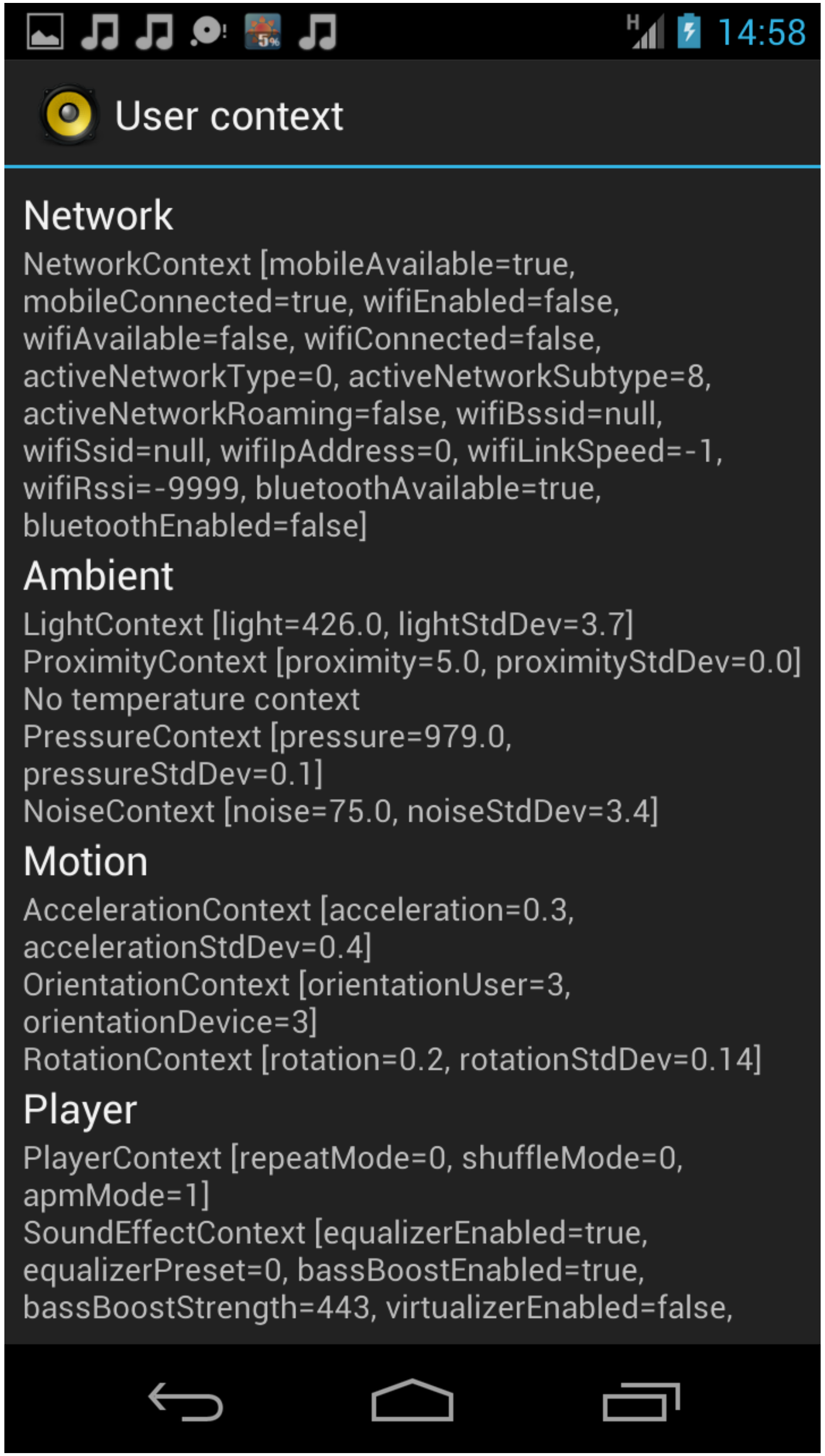



# Mobile Music Genius

Automatic playlist generation based on music context (features and similarity computed based on Last.fm tags)

Department of
Computational
Perception

# Mobile Music Genius




Some user context features gathered while playing

Department of Computational Perception

# User Context Features from Android Phones

*Time:* timestamp, time zone

*Personal:* userID/eMail, gender, birthdate

*Device:* devideID (IMEI), sw version, manufacturer, model, phone state, connectivity, storage, battery, various volume settings (media, music, ringer, system, voice)

*Location:* longitude/latitude, accuracy, speed, altitude

*Place:* nearby place name (populated), most relevant city

*Weather:* wind direction, speed, clouds, temperature, dew point, humidity, air pressure

*Ambient:* light, proximity, temperature, pressure, noise, digital environment (WiFi and BT network information)

*Activity:* acceleration, user and device orientation, screen on/off, running apps

*Player:* artist, album, track name, track id, track length, genre, plackback position, playlist name, playlist type, player state (repeat, shuffle mode), audio output (headset plugged)

**mood and activity (direct user feedback)**

Department of Computational Perception

# Preliminary Evaluation

- collected user context data from 12 participants over a period of 4 weeks
- age: 20-40 years, gender: male
- user context vectors recoded whenever a “sensor” records a change
- 166k data points
- assess different classifiers (Weka) for the task of predicting artist/track/genre/mood given a user context vector: k-nearest neighbor (kNN), decision tree (C4.5), Support Vector Machine (SVM), Bayes Network (BN)
- cross-fold validation (10-CV)

To be analyzed:

(i) Which granularity/abstraction level to choose for representation/learning?

(ii) Which user context features are the most important to predict music preference?

# Preliminary Evaluation: Results

(i) Which granularity/abstraction level to choose for representation/learning?

Predicting class *track*

Results barely above baseline.

Predicting particular tracks is hardly feasible with the amount of data available.

| Dataset | 0-R | KNN | C4.5 | SVM | BN | Max.rel. | Avg.rel. |
|---|---|---|---|---|---|---|---|
| Time | 1.13 | 1.26 | 0.90 | 1.30 | 1.31 | 116.04% | 105.84% |
| Location | 1.13 | 1.40 | 1.57 | 1.42 | 1.58 | 139.76% | 132.06% |
| Location - state | 1.13 | 1.36 | 1.69 | 0.96 | 0.82 | 150.26% | 107.28% |
| Location - place | 1.13 | 1.31 | 1.47 | 1.46 | 2.23 | 197.49% | 143.46% |
| Weather | 1.13 | 1.17 | 0.91 | 1.19 | 1.07 | 105.25% | 96.21% |
| Ambient | 1.13 | 0.79 | 0.63 | 1.08 | 1.12 | 98.97% | 79.99% |
| Ambient - no n. | 1.13 | 0.64 | 0.63 | 0.97 | 1.10 | 97.49% | 73.97% |
| Ambient - noise | 1.13 | 0.45 | 0.67 | 1.28 | 1.11 | 113.38% | 77.77% |
| Motion | 1.13 | 0.54 | 0.97 | 1.06 | 1.32 | 117.15% | 86.25% |
| Motion - acc. | 1.13 | 0.58 | 0.58 | 1.39 | 1.10 | 123.50% | 80.75% |
| Motion - orient. | 1.13 | 1.09 | 1.33 | 0.94 | 1.41 | 124.76% | 105.78% |
| Task | 1.13 | 1.43 | 1.96 | 1.57 | 1.73 | 173.61% | 148.36% |
| Task - display | 1.13 | 1.75 | 1.68 | 1.76 | 1.76 | 156.47% | 154.21% |
| Task - tasks | 1.13 | 1.16 | 1.60 | 1.13 | 1.53 | 141.76% | 120.03% |
| Phone | 1.13 | 1.12 | 0.97 | 0.70 | 0.99 | 99.41% | 83.85% |
| Network | 1.13 | 1.43 | 1.34 | 1.26 | 1.82 | 161.79% | 129.88% |
| Network - state | 1.13 | 1.31 | 1.75 | 1.58 | 1.82 | 161.79% | 143.27% |
| Network - env. | 1.05 | 1.79 | 1.45 | 1.44 | 1.08 | 170.20% | 137.07% |
| Device | 1.13 | 1.07 | 1.56 | 1.12 | 1.24 | 138.14% | 110.74% |
| Device - battery | 1.13 | 0.71 | 1.12 | 1.23 | 1.12 | 109.39% | 92.78% |
| Device - storage | 1.13 | 0.95 | 1.07 | 1.44 | 1.42 | 127.49% | 108.09% |
| Device - memory | 1.13 | 0.92 | 0.79 | 1.24 | 1.30 | 115.59% | 94.46% |
| Device - audio | 1.13 | 0.46 | 0.63 | 0.96 | 1.30 | 114.93% | 74.26% |
| Player | 1.13 | 1.29 | 1.36 | 1.35 | 1.35 | 120.77% | 118.46% |
| All | 1.13 | 0.90 | 1.78 | 1.14 | 1.14 | 158.02% | 110.05% |

# Preliminary Evaluation: Results

(i) Which granularity/abstraction level to choose for representation/learning?

Predicting class *artist*

Best results achieved, significantly outperforming baseline.

Relation {context → artist} seems to be predictable.

| Dataset | 0-R | KNN | C4.5 | SVM | BN | Max.rel. | Avg.rel. |
|---|---|---|---|---|---|---|---|
| Time | 28.54 | 60.83 | 57.10 | 59.68 | 58.70 | 213.15% | 207.01% |
| Location | 28.54 | 42.69 | 41.42 | 37.80 | 40.04 | 149.58% | 141.86% |
| Location - state | 28.54 | 41.71 | 41.83 | 33.11 | 37.05 | 146.55% | 134.64% |
| Location - place | 28.54 | 35.74 | 36.99 | 36.07 | 36.28 | 129.62% | 127.09% |
| Weather | 28.54 | 63.46 | 63.25 | 56.06 | 61.34 | 222.35% | 213.84% |
| Ambient | 28.54 | 34.70 | 36.83 | 31.17 | 35.18 | 129.03% | 120.77% |
| Ambient - no n. | 28.54 | 33.54 | 34.87 | 31.43 | 34.46 | 122.19% | 117.65% |
| Ambient - noise | 28.54 | 26.12 | 30.55 | 28.75 | 29.81 | 107.04% | 100.94% |
| Motion | 28.54 | 35.08 | 36.10 | 37.14 | 35.11 | 130.15% | 125.65% |
| Motion - acc. | 28.54 | 26.54 | 27.87 | 28.93 | 28.62 | 101.36% | 98.07% |
| Motion - orient. | 28.54 | 36.22 | 35.63 | 36.54 | 35.17 | 128.02% | 125.75% |
| Task | 28.54 | 60.75 | 60.65 | 59.63 | 56.20 | 212.86% | 207.81% |
| Task - display | 28.54 | 28.12 | 28.31 | 28.62 | 28.34 | 100.29% | 99.33% |
| Task - tasks | 28.54 | 61.35 | 61.28 | 60.28 | 55.23 | 214.97% | 208.60% |
| Phone | 28.54 | 37.30 | 38.74 | 31.33 | 33.74 | 135.74% | 123.61% |
| Network | 28.54 | 36.38 | 36.44 | 37.93 | 34.87 | 132.90% | 127.56% |
| Network - state | 28.54 | 34.95 | 33.14 | 34.58 | 34.17 | 122.45% | 119.86% |
| Network - env. | 21.90 | 25.01 | 26.42 | 27.43 | 22.69 | 125.26% | 115.92% |
| Device | 28.54 | 70.42 | 68.68 | 54.95 | 65.31 | 246.76% | 227.20% |
| Device - battery | 28.54 | 39.10 | 47.15 | 36.41 | 46.02 | 165.23% | 147.76% |
| Device - storage | 28.54 | 61.17 | 60.37 | 40.96 | 57.92 | 214.33% | 193.08% |
| Device - memory | 28.54 | 39.22 | 40.56 | 32.11 | 36.53 | 142.10% | 130.01% |
| Device - audio | 28.54 | 47.92 | 47.71 | 41.42 | 42.76 | 167.90% | 157.50% |
| Player | 28.54 | 38.18 | 38.36 | 38.30 | 38.25 | 134.41% | 134.10% |
| All | 28.54 | 69.56 | 69.01 | 69.87 | 67.66 | 244.83% | 241.86% |

# Preliminary Evaluation: Results

(i) Which granularity/abstraction level to choose for representation/learning?

Predicting class *genre*

Prediction on more general level than for artist.

Still genre is an ill-defined concept, hence results inferior to artist prediction.

| Dataset | 0-R | KNN | C4.5 | SVM | BN | Max.rel. | Avg.rel. |
|---|---|---|---|---|---|---|---|
| Time | 29.80 | 46.75 | 44.99 | 46.46 | 46.27 | 156.88% | 154.76% |
| Location | 29.80 | 32.92 | 34.17 | 34.45 | 32.05 | 115.61% | 112.08% |
| Location - state | 29.80 | 32.25 | 33.41 | 32.48 | 30.44 | 112.12% | 107.87% |
| Location - place | 29.80 | 29.75 | 32.54 | 32.38 | 32.45 | 109.19% | 106.65% |
| Weather | 29.80 | 49.68 | 50.61 | 43.77 | 46.70 | 169.83% | 160.03% |
| Ambient | 29.80 | 28.30 | 34.12 | 31.38 | 33.27 | 114.50% | 106.61% |
| Ambient - no n. | 29.80 | 31.52 | 33.39 | 31.42 | 33.34 | 112.04% | 108.79% |
| Ambient - noise | 29.80 | 23.38 | 29.92 | 29.67 | 29.77 | 100.40% | 94.57% |
| Motion | 29.80 | 32.23 | 34.34 | 34.56 | 34.39 | 115.98% | 113.69% |
| Motion - acc. | 29.80 | 25.67 | 28.55 | 30.50 | 30.41 | 102.35% | 96.59% |
| Motion - orient. | 29.80 | 34.49 | 35.22 | 34.28 | 34.39 | 118.21% | 116.10% |
| Task | 29.80 | 43.89 | 46.47 | 44.55 | 41.85 | 155.95% | 148.29% |
| Task - display | 29.80 | 28.57 | 29.04 | 28.78 | 28.78 | 97.44% | 96.61% |
| Task - tasks | 29.80 | 44.71 | 47.62 | 44.94 | 42.31 | 159.81% | 150.66% |
| Phone | 29.80 | 31.17 | 33.43 | 31.33 | 30.13 | 112.20% | 105.77% |
| Network | 29.80 | 32.31 | 31.96 | 33.93 | 31.73 | 113.85% | 109.00% |
| Network - state | 29.80 | 31.70 | 31.14 | 32.07 | 31.26 | 107.63% | 105.85% |
| Network - env. | 26.10 | 26.17 | 27.02 | 29.78 | 27.28 | 114.08% | 105.58% |
| Device | 29.80 | 49.65 | 50.03 | 43.16 | 48.00 | 167.88% | 160.11% |
| Device - battery | 29.80 | 31.58 | 38.03 | 33.42 | 35.85 | 127.61% | 116.51% |
| Device - storage | 29.80 | 47.76 | 47.55 | 37.25 | 46.56 | 160.29% | 150.28% |
| Device - memory | 29.80 | 30.79 | 36.87 | 31.76 | 36.60 | 123.73% | 114.11% |
| Device - audio | 29.80 | 40.19 | 41.12 | 38.16 | 37.02 | 137.99% | 131.29% |
| Player | 29.80 | 35.79 | 36.34 | 36.08 | 35.59 | 121.96% | 120.65% |
| All | 29.80 | 46.75 | 49.22 | 50.41 | 48.51 | 169.15% | 163.50% |

# Preliminary Evaluation: Results

(i) Which granularity/abstraction level to choose for representation/learning?

Predicting class *mood*

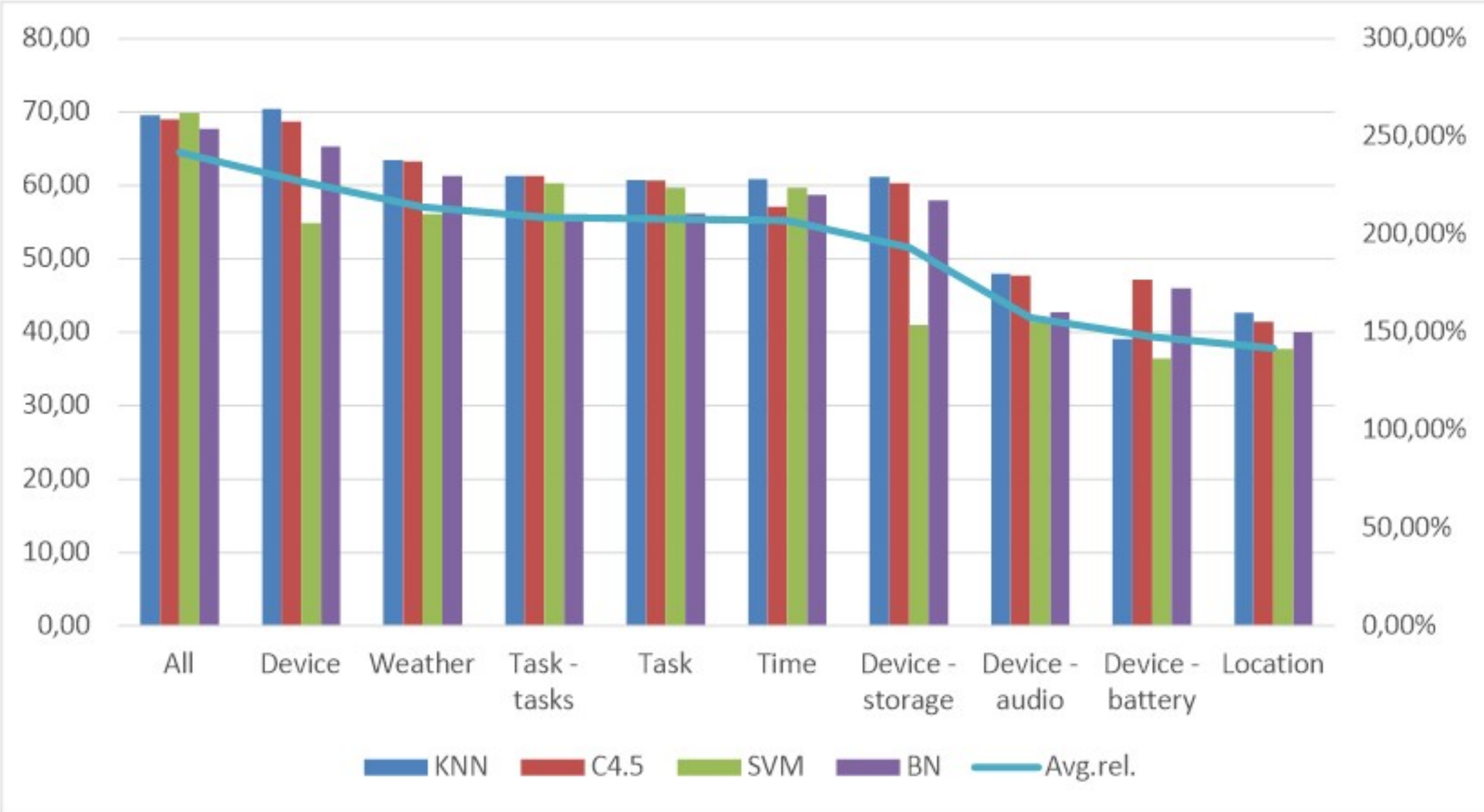
| Dataset | 0-R | KNN | C4.5 | SVM | BN | Max.rel. | Avg.rel. |
|---|---|---|---|---|---|---|---|
| Time | 24.00 | 24.79 | 27.73 | 24.56 | 24.29 | 115.53% | 105.59% |
| Location | 24.00 | 23.27 | 23.89 | 25.05 | 24.62 | 104.38% | 100.86% |
| Location - state | 24.00 | 23.44 | 23.97 | 25.25 | 24.79 | 105.20% | 101.51% |
| Location - place | 24.00 | 21.99 | 23.99 | 23.80 | 23.67 | 99.94% | 97.33% |
| Weather | 24.00 | 25.13 | 27.05 | 27.86 | 25.39 | 116.07% | 109.82% |
| Ambient | 24.00 | 17.04 | 19.41 | 23.59 | 24.04 | 100.17% | 87.58% |
| Ambient - no n. | 24.00 | 21.14 | 23.18 | 23.87 | 24.00 | 100.00% | 96.03% |
| Ambient - noise | 24.00 | 16.70 | 21.38 | 23.79 | 23.96 | 99.83% | 89.40% |
| Motion | 24.00 | 19.88 | 26.54 | 24.78 | 24.65 | 110.56% | 99.84% |
| Motion - acc. | 24.00 | 20.86 | 22.75 | 24.32 | 23.96 | 101.34% | 95.72% |
| Motion - orient. | 24.00 | 23.99 | 27.82 | 24.99 | 24.65 | 115.91% | 105.68% |
| Task | 24.00 | 22.94 | 24.32 | 24.58 | 25.00 | 104.18% | 100.87% |
| Task - display | 24.00 | 24.45 | 24.58 | 24.97 | 24.88 | 104.06% | 103.00% |
| Task - tasks | 24.00 | 23.56 | 25.20 | 24.99 | 24.13 | 105.00% | 101.95% |
| Phone | 24.00 | 19.34 | 24.64 | 26.75 | 26.74 | 111.45% | 101.52% |
| Network | 24.00 | 22.81 | 24.20 | 23.92 | 24.28 | 101.17% | 99.17% |
| Network - state | 24.00 | 23.48 | 24.39 | 24.01 | 24.28 | 101.64% | 100.17% |
| Network - env. | 27.78 | 27.68 | 28.36 | 29.24 | 27.78 | 105.26% | 101.74% |
| Device | 24.00 | 21.45 | 24.72 | 25.79 | 24.86 | 107.46% | 100.86% |
| Device - battery | 24.00 | 16.09 | 26.31 | 23.94 | 24.06 | 109.64% | 94.17% |
| Device - storage | 24.00 | 25.57 | 26.69 | 25.36 | 24.48 | 111.19% | 106.36% |
| Device - memory | 24.00 | 13.92 | 21.39 | 23.59 | 23.81 | 99.22% | 86.16% |
| Device - audio | 24.00 | 26.33 | 26.50 | 25.48 | 24.43 | 110.43% | 107.03% |
| Player | 24.00 | 24.81 | 25.57 | 25.37 | 25.45 | 106.54% | 105.41% |
| All | 24.00 | 22.43 | 26.16 | 24.81 | 26.11 | 109.00% | 103.66% |

Poor results as mood in music is quite subjective and hence hard to predict.

Which mood anyway: composers intention? mood expressed by performers? mood evoked in listeners?

# Preliminary Evaluation: Results

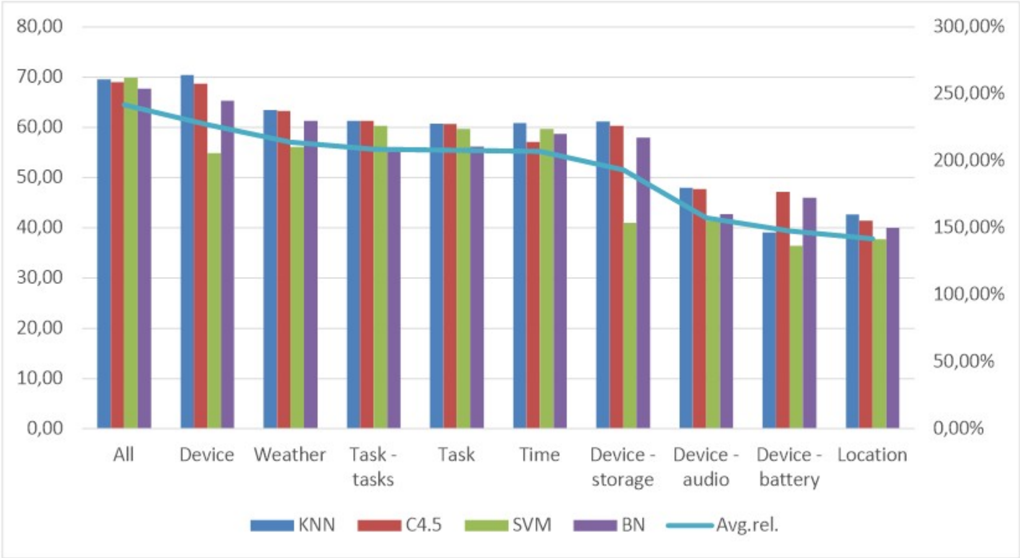
(ii) Which user context features are the most important to predict music preference?




Making use of all features yields best results.

Department of Computational Perception

# Preliminary Evaluation: Results




(ii) Which user context features are the most important to predict music preference?

Weka-feature selection confirms most important attributes:

*time:* weekday, hour of day

*location*: nearest populated place (better than longitude, and latitude)

*weather:* temperature, humidity, air pressure, wind speed/direction, and dew point

*device*: music and ringer volume, battery level, available storage and memory

*task*: running tasks/apps

Department of
Computational
Perception

# Preliminary Evaluation: Results

- Problems:

– too little data to make significant predictions on the quality of the approach

→ need more data from more participants over a longer period of time

→ large-scale study

– dataset does not incorporate features potentially highly relevant to music listening inclination (user activity and mood)

# Large-scale Evaluation

- collected user context data from JKU students over a period of 2 months
- about 8,000 listening data items and corresponding user context gathered

To be analyzed:

(i) How well does our approach perform to predict the preferred artist based on a given user context vector?

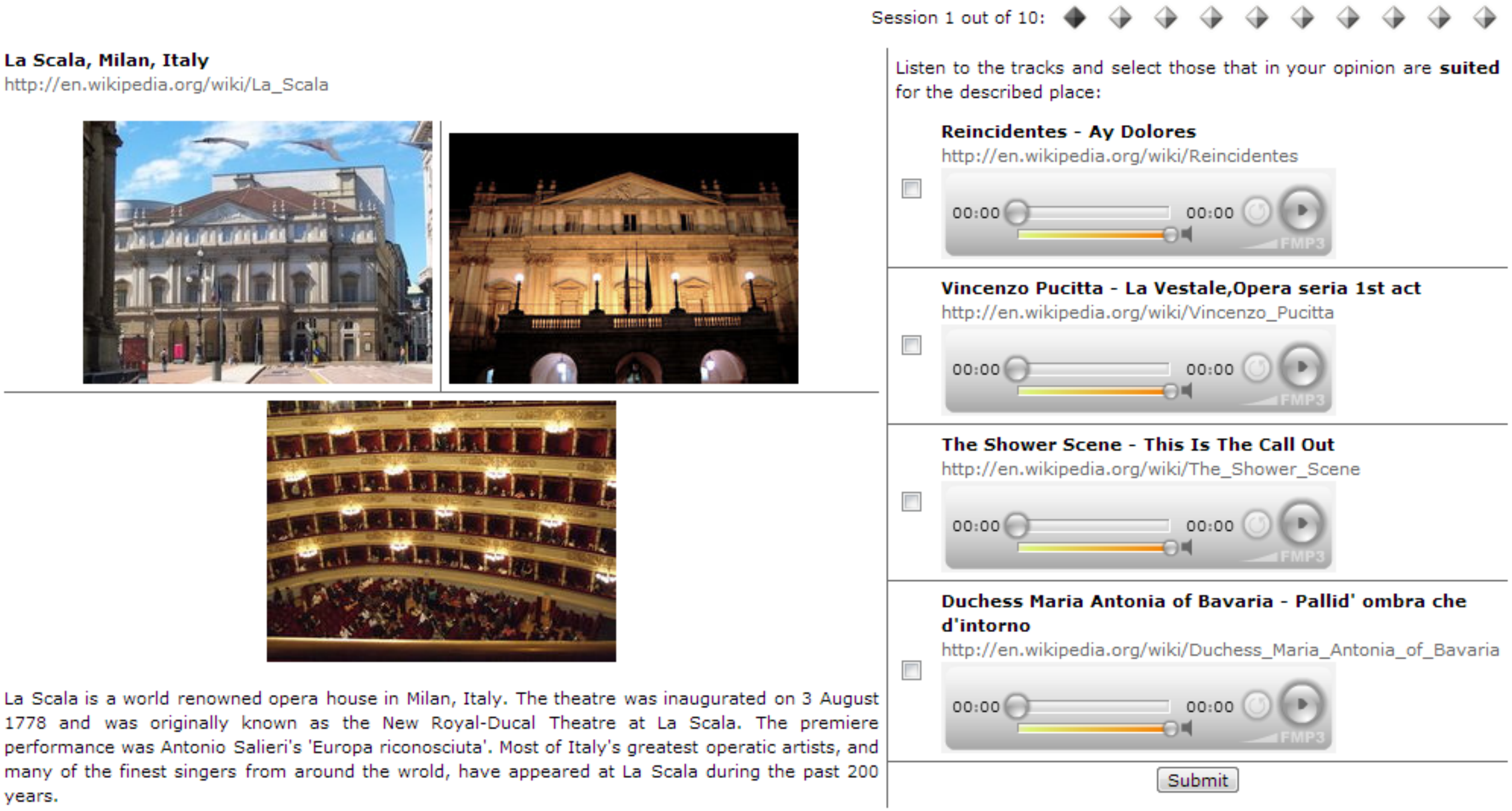
Results for predicting class "artist":

| | |
|---|---|
| ZeroR (baseline) classifier | 15% accuracy |
| k-nearest neighbors | 42% accuracy |
| JRip rule learner | 51% accuracy |
| J48 decision tree | 55% accuracy |

# Matching Places of Interest and Music

(Kaminskas et al.; RecSys 2013)

recommend music that is suited to a place of interest (POI) of the user (context-aware)



# Matching Places of Interest and Music

(Kaminskas et al.; RecSys 2013)

Approaches:

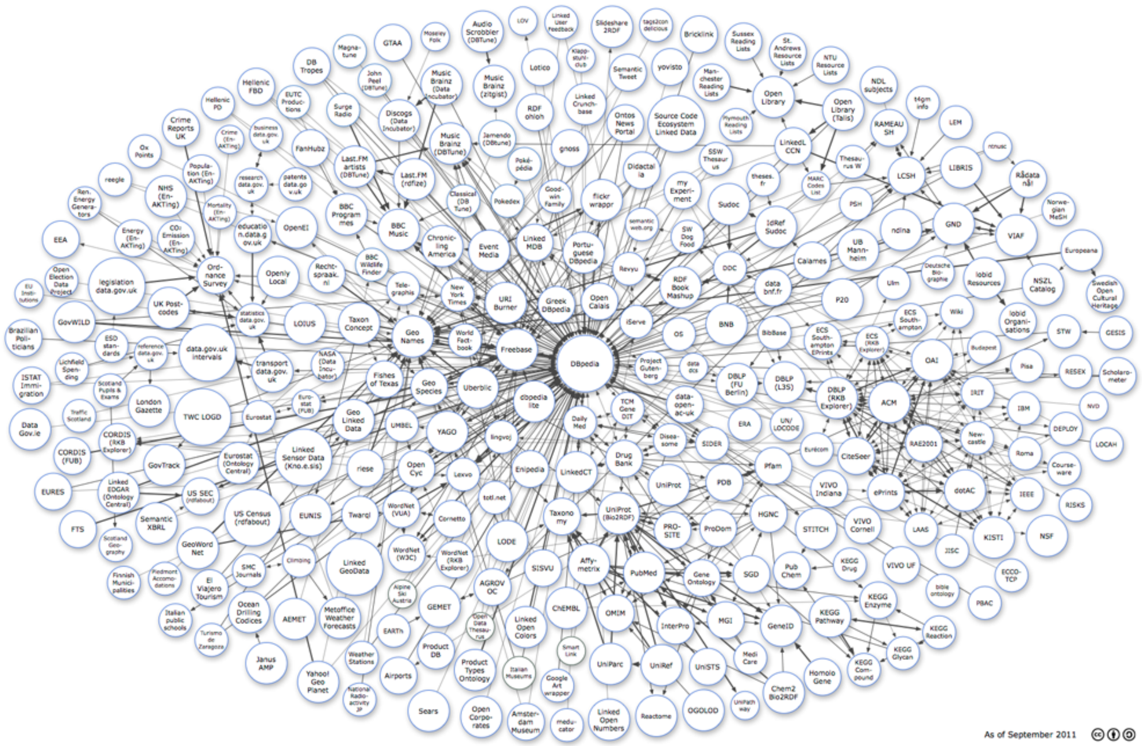
- *genre-based*: only play music belonging to the user's preferred genres (baseline)

# Matching Places of Interest and Music

(Kaminskas et al.; RecSys 2013)

Approaches:

- *knowledge-based*: use the DBpedia knowledge base (relations between POIs and musicians)



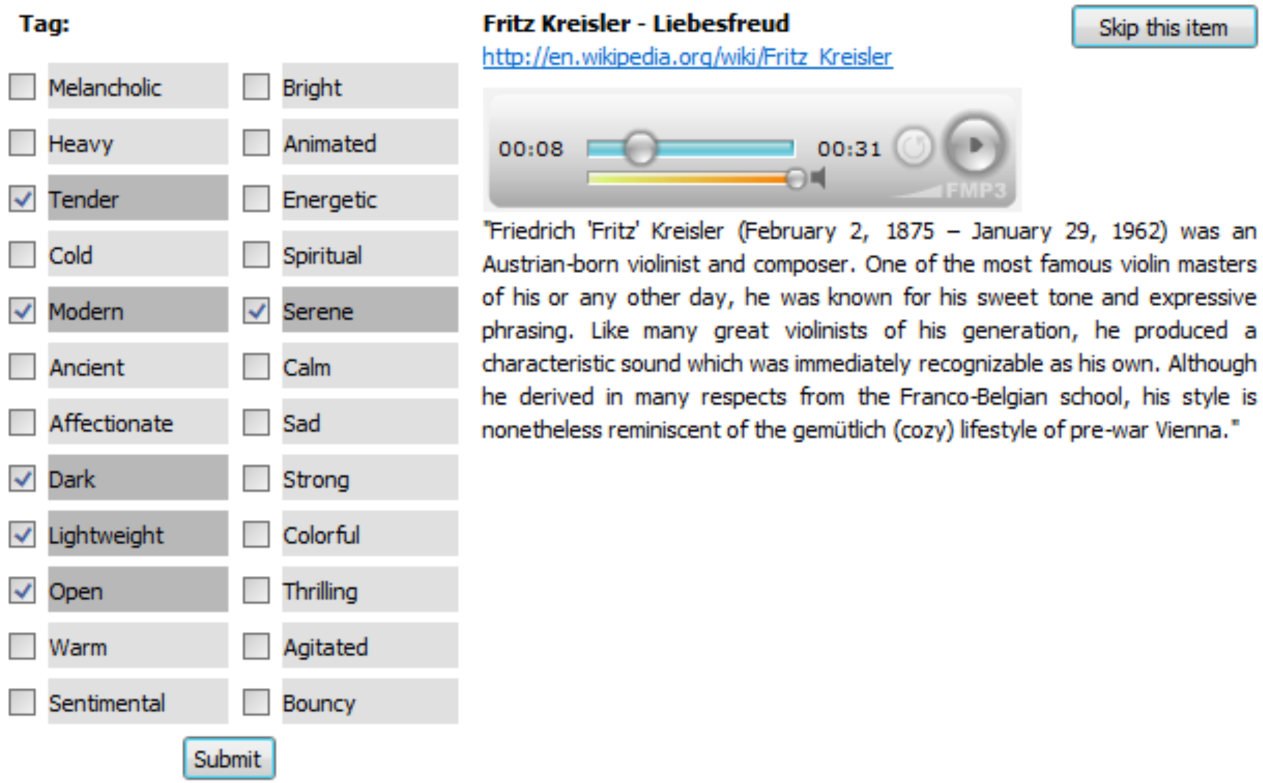
As of September 2011

# Matching Places of Interest and Music

(Kaminskas et al.; RecSys 2013)

Approaches:

- *tag-based*: user-assigned emotion tags describing images of POIs and music, Jaccard similarity between music-tag-vectors and POI-tag-vectors

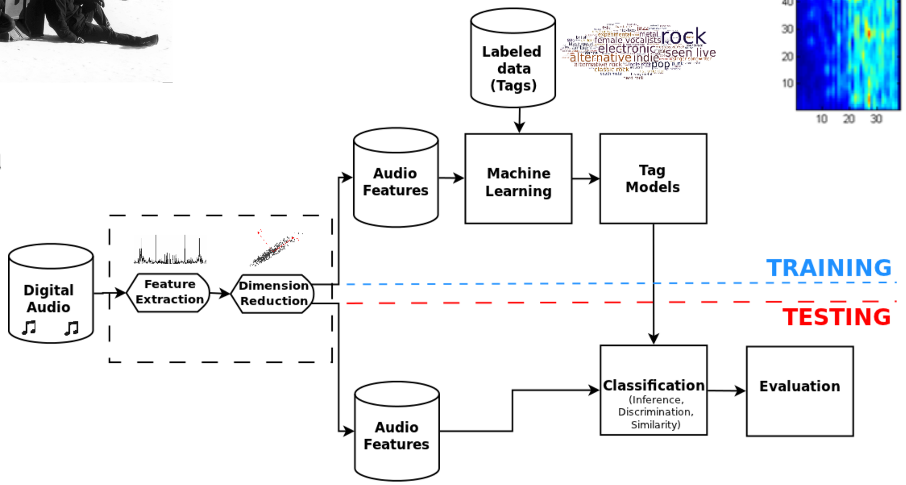


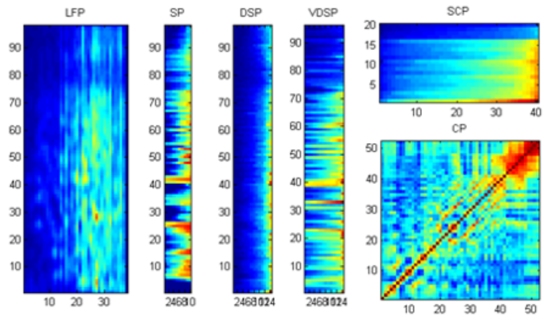
# Matching Places of Interest and Music

(Kaminskas et al.; RecSys 2013)

Approaches:

- *auto-tag-based*: use state-of-the-art music auto-tagger based on the Block-level Feature framework to automatically label music pieces; then again compute Jaccard similarity between music-tag-vectors and POI-tag-vectors

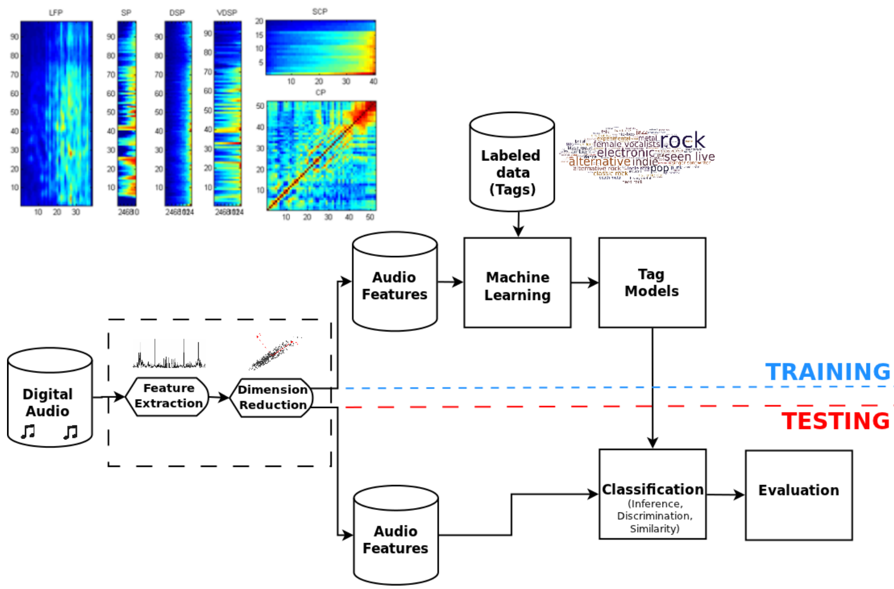

Department of Computational Perception

# Matching Places of Interest and Music

(Kaminskas et al.; RecSys 2013)

Approaches:

- *combined*: aggregate music recommendations w.r.t. ranks given by knowledge-based and auto-tag-based approaches

# Matching Places of Interest and Music

(Kaminskas et al.; RecSys 2013)

Approaches:

- *genre-based*: only play music belonging to the user's preferred genres (baseline)
- *knowledge-based*: using the DBpedia knowledge base (relations between POIs and musicians)
- *tag-based*: user-assigned emotion tags describing images of POIs and music, Jaccard similarity between music-tag-vectors and POI-tag-vectors
- *auto-tag-based*: using state-of-the-art music auto-tagger based on the Block-level Feature Framework to automatically label music pieces; then again use Jaccard similarity between music-tag-vectors and POI-tag-vectors
- *combined*: aggregate music recommendations w.r.t. ranks given by knowledge-based and auto-tag-based approaches

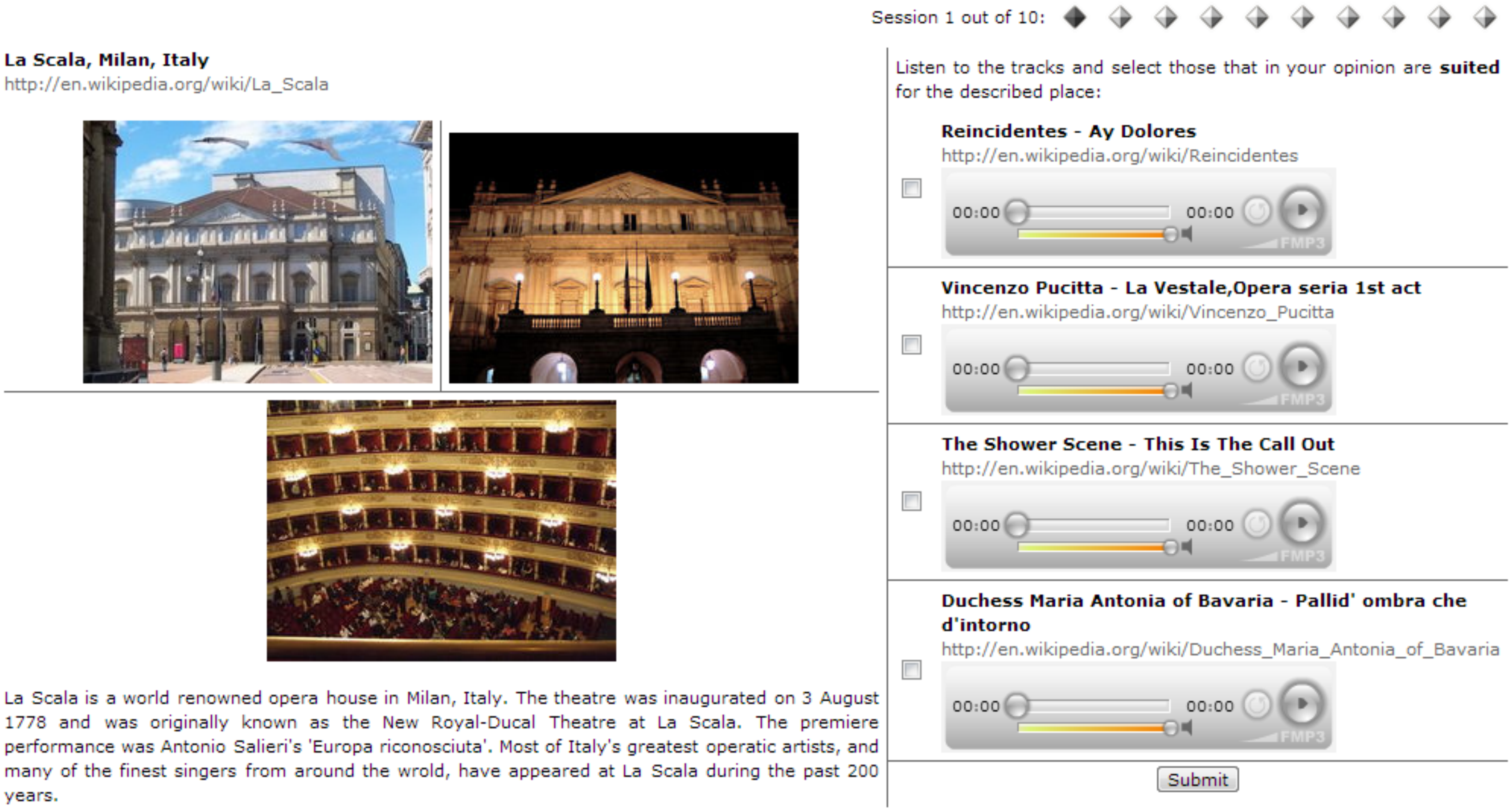
Department of Computational Perception

# Matching Places of Interest and Music

(Kaminskas et al.; RecSys 2013)

Evaluation:

- user study via web interface (58 users, 564 sessions)

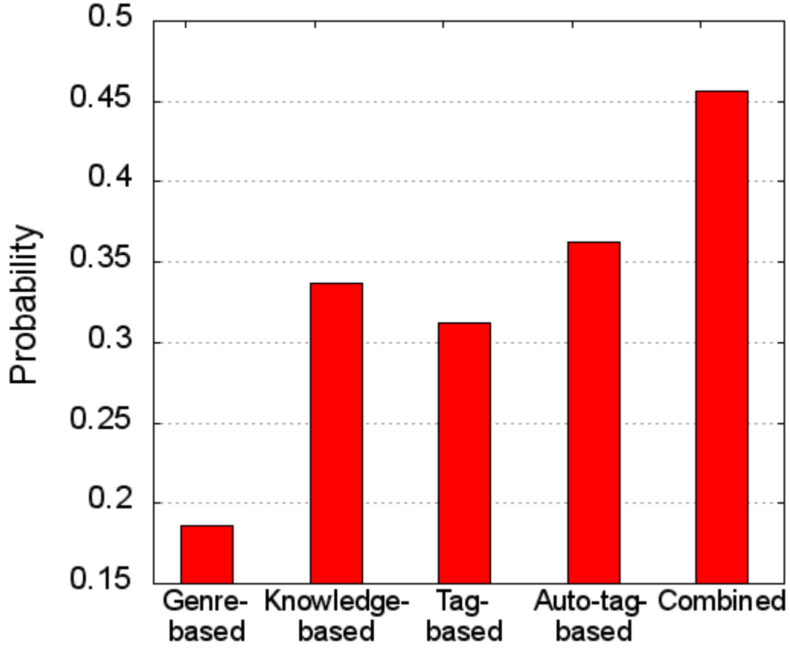


# Matching Places of Interest and Music

(Kaminskas et al.; RecSys 2013)

Evaluation:

- Performance measure: number of times a track produced by each approach was considered as well-suited in relation to total number of evaluation sessions, i.e. probability that a track marked as well-suited by a user was recommended by each approach



# SUMMARY

Department of
Computational
Perception

# Music Information Retrieval is a great field

Various approaches to extract information from the audio signal

Various sources and approaches to extract contextual data and similarity information from the Web

Multi-modal modeling and retrieval is important and allows for exciting applications

**Next big challenges:**

- modeling user properties and context
- improve personalization and context-awareness
- situation-based retrieval
- new and better suited evaluation strategies