

Music Retrieval and Recommendation

Part III: Music Context-based Similarity

Peter Knees
Markus Schedl

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

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

Context- and Web-Based Methods

In the following, contextual data refers to extended meta-data, usually

- Generated by users
- Stemming from unstructured data-sources
- Accessible via the Web

Text processing methods used for:

- Music similarity (again, the central concept for retrieval)
- Music indexing and retrieval
- Tagging of music
- Complementing audio-based methods

Text-Based Similarity and Genre Classification

Use Web data to transform the music similarity task into a text similarity task

Allows to use the full armory of IR methods, typically...

- Bag-of-words, Vector Space Model
- Stopword removal, dictionaries, term selection
- $TF \cdot IDF$
- Latent Semantic Indexing, Latent Dirichlet Allocation
- Part-of-Speech tagging
- Named Entity Detection
- Sentiment analysis

Large range of possible similarity measures

- Overlap, Manhattan, Euclidean, Cosine, BM25, etc.

Levels of Music Context Feature Extraction



High-level

Examples:
style, mood, intentions, meaning (e.g., lyrics), “semantics”



Mid-level

Examples:
tags, concept clusters, topics, sentiments



Low-level

Examples:
characters, words, n-grams, part of speech, co-occurrences

Sources of Extended Music Meta-data

Text-based data sources used in MIR:

- Web pages retrieved via Web search engines



- microblogs on *Twitter*

twitter



- product reviews

Epinions 

- semantic tags

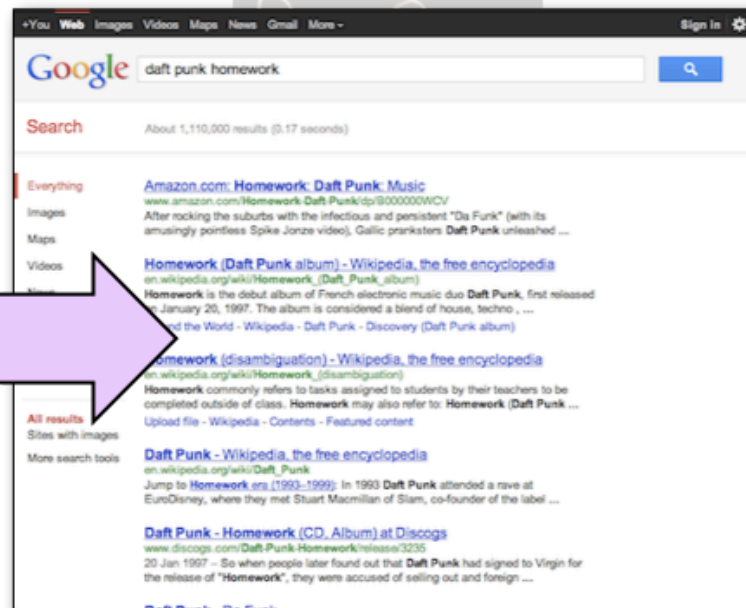
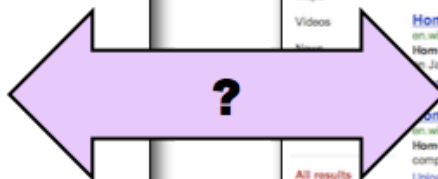
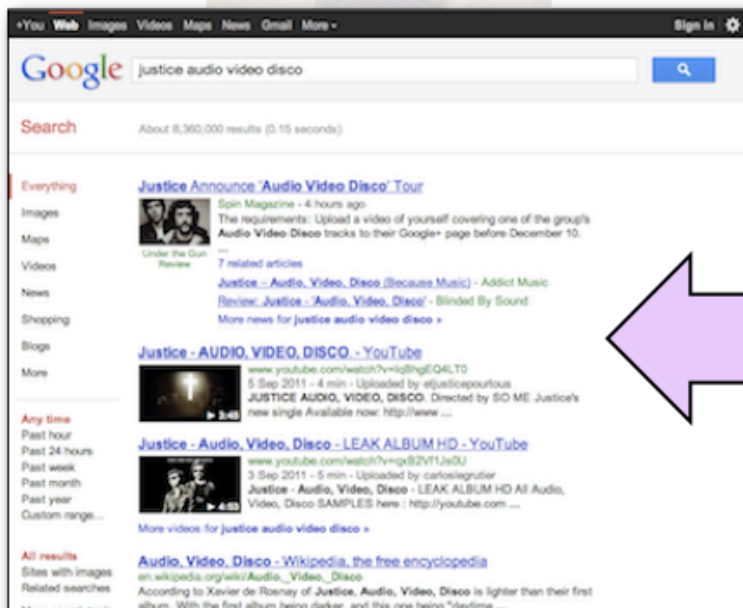
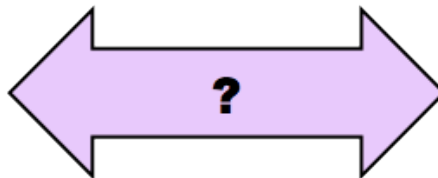
last.fm
the social music revolution

- lyrics

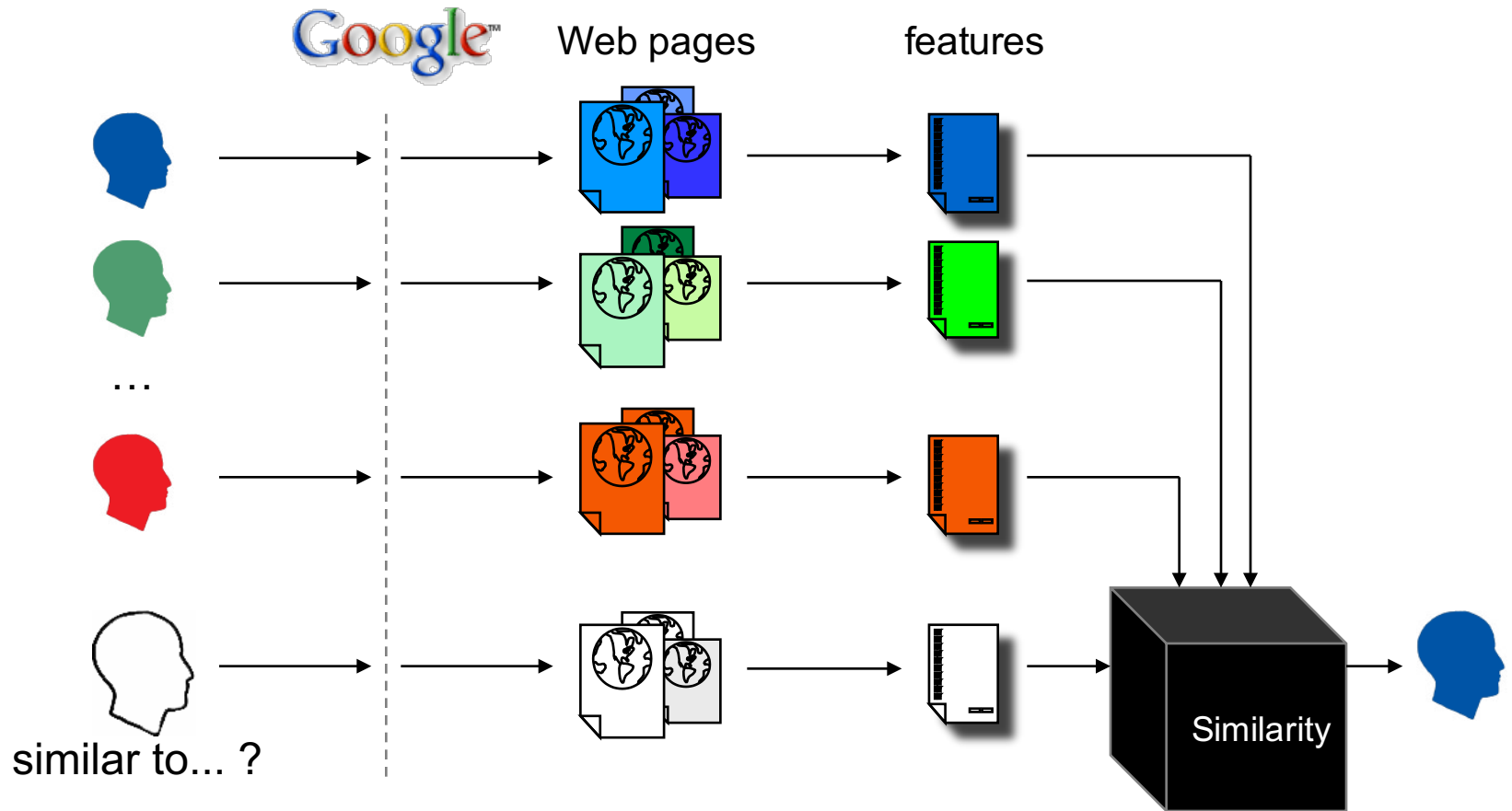
Lyrics.com

- biographies

Related Web Pages as Text Source



Related Web Pages as Text Source



Related Web Pages as Text Source

- Using search engines and queries such as

“artist” +music

“artist” +music +review

(Whitman, Lawrence; 2002) (Baumann, Hummel; 2003) (Knees et al.; 2004)

- Analyze

- result page directly or

- download up to top 100 Web pages (combine into one “virtual document” or analyze separately)

- Apply “IR magic”

- Applicable for similarity estimation, classification, retrieval, annotation

(NB: Most discriminating terms between genres are artist names and album/track titles)



Large-Scale Study

Investigating different aspects in modeling artist term profiles from Web pages
(9,200 experiments): (Schedl et al.; 2011)

- **term frequency**

Abbr.	Description	Formulation
TF_A	Formulation used for binary match SB = b	$r_{d,t} = \begin{cases} 1 & \text{if } t \in \mathcal{T}_d \\ 0 & \text{otherwise} \end{cases}$
TF_B	Standard formulation SB = t	$r_{d,t} = f_{d,t}$
TF_C	Logarithmic formulation	$r_{d,t} = 1 + \log_e f_{d,t}$
TF_C2	Alternative logarithmic formulation suited for $f_{d,t} < 1$	$r_{d,t} = \log_e(1 + f_{d,t})$
TF_C3	Alternative logarithmic formulation as used in <i>ltc</i> variant	$r_{d,t} = 1 + \log_2 f_{d,t}$
TF_D	Normalized formulation	$r_{d,t} = \frac{f_{d,t}}{f_d^m}$
TF_E	Alternative normalized formulation. Similar to [55] we use $K = 0.5$. SB = n	$r_{d,t} = K + (1 - K) \cdot \frac{f_{d,t}}{f_d^m}$
TF_F	Okapi formulation, according to [55, 36]. For W we use the vector space formulation, i.e., the Euclidean length.	$r_{d,t} = \frac{f_{d,t}}{f_{d,t} + W_d / \text{av}_{d \in D}(W_d)}$
TF_G	Okapi BM25 formulation, according to [35].	$r_{d,t} = \frac{(k_1 + 1) \cdot f_{d,t}}{f_{d,t} + k_1 \cdot \left[(1-b) + b \cdot \frac{W_d}{\text{av}_{d \in D}(W_d)} \right]}$ $k_1 = 1.2, b = 0.75$



Large-Scale Study

Investigating different aspects in modeling artist term profiles from Web pages
(9,200 experiments): (Schedl et al.; 2011)

- term frequency
- inverse document frequency

Abbr.	Description	Formulation
IDF_A	Formulation used for binary match $SB = x$	$w_t = 1$
IDF_B	Logarithmic formulation $SB = f$	$w_t = \log_e \left(1 + \frac{N}{f_t} \right)$
IDF_B2	Logarithmic formulation used in <i>ltc</i> variant	$w_t = \log_e \left(\frac{N}{f_t} \right)$
IDF_C	Hyperbolic formulation	$w_t = \frac{1}{f_t}$
IDF_D	Normalized formulation	$w_t = \log_e \left(1 + \frac{f_m}{f_t} \right)$
IDF_E	Another normalized formulation $SB = p$	$w_t = \log_e \frac{N - f_t}{f_t}$
	The following definitions are based on the term's noise n_t and signal s_t .	$n_t = \sum_{d \in \mathcal{D}_t} \left(-\frac{f_{d,t}}{F_t} \log_2 \frac{f_{d,t}}{F_t} \right)$ $s_t = \log_2(F_t - n_t)$
IDF_F	Signal	$w_t = s_t$
IDF_G	Signal-to-Noise ratio	$w_t = \frac{s_t}{n_t}$
IDF_H		$w_t = \left(\max_{t' \in T} n_{t'} \right) - n_t$
IDF_I	Entropy measure	$w_t = 1 - \frac{n_t}{\log_2 N}$
IDF_J	Okapi BM25 IDF formulation, according to [35, 31]	$w_t = \log \frac{N - f_t + 0.5}{f_t + 0.5}$



Large-Scale Study

Investigating different aspects in modeling artist term profiles from Web pages
(9,200 experiments): (Schedl et al.; 2011)

- **term frequency**
- **inverse document frequency**
- **virtual document modeling:** *concatenate* all Web pages/posts of the artist or perform *aggregation* via mean, max, etc.

Large-Scale Study

Investigating different aspects in modeling artist term profiles from Web pages
(9,200 experiments): (Schedl et al.; 2011)

- **term frequency**
- **inverse document frequency**
- **virtual document modeling:** *concatenate* all Web pages/posts of the artist or perform *aggregation* via mean, max, etc.
- **normalization with respect to document length**

Abbr.	Description	Formulation
NORM_NO	No normalization.	
NORM_SUM	Normalize sum of each virtual document's term feature vector to 1.	$\sum_{t \in T_d} r_{d,t} = 1$
NORM_MAX	Normalize maximum of each virtual document's term feature vector to 1.	$\max_{t \in T_d} r_{d,t} = 1$

Large-Scale Study

Investigating different aspects in modeling artist term profiles from Web pages
(9,200 experiments): (Schedl et al.; 2011)

- **term frequency**
- **inverse document frequency**
- **virtual document modeling:** *concatenate* all Web pages/posts of the artist or perform *aggregation* via mean, max, etc.
- **normalization with respect to document length**
- **similarity measure**

Abbr.	Description	Formulation
SIM_INN	Inner Product	$S_{d_1, d_2} = \sum_{t \in \mathcal{T}_{d_1, d_2}} (w_{d_1, t} \cdot w_{d_2, t})$
SIM_COS	Cosine Measure	$S_{d_1, d_2} = \frac{\sum_{t \in \mathcal{T}_{d_1, d_2}} (w_{d_1, t} \cdot w_{d_2, t})}{W_{d_1} \cdot W_{d_2}}$
SIM_DIC	Dice Formulation	$S_{d_1, d_2} = \frac{2 \sum_{t \in \mathcal{T}_{d_1, d_2}} (w_{d_1, t} \cdot w_{d_2, t})}{W_{d_1}^2 + W_{d_2}^2}$
SIM_JAC	Jaccard Formulation	$S_{d_1, d_2} = \frac{\sum_{t \in \mathcal{T}_{d_1, d_2}} (w_{d_1, t} \cdot w_{d_2, t})}{W_{d_1}^2 + W_{d_2}^2 - \sum_{t \in \mathcal{T}_{d_1, d_2}} (w_{d_1, t} \cdot w_{d_2, t})}$
SIM_OVL	Overlap Formulation	$S_{d_1, d_2} = \frac{\sum_{t \in \mathcal{T}_{d_1, d_2}} (w_{d_1, t} \cdot w_{d_2, t})}{\min(W_{d_1}^2, W_{d_2}^2)}$
SIM_EUC	Euclidean Similarity	$D_{d_1, d_2} = \sqrt{\sum_{t \in \mathcal{T}_{d_1, d_2}} (w_{d_1, t} - w_{d_2, t})^2}$ $S_{d_1, d_2} = \left(\max_{d'_1, d'_2} (D_{d'_1, d'_2}) \right) - D_{d_1, d_2}$
SIM_JEF	Jeffrey Divergence-based Similarity	$S_{d_1, d_2} = \left(\max_{d'_1, d'_2} (D_{d'_1, d'_2}) \right) - D_{d_1, d_2}$ $D(F, G) = \sum \left(f_i \log \frac{f_i}{m_i} + g_i \log \frac{g_i}{m_i} \right)$

Large-Scale Study

Investigating different aspects in modeling artist term profiles from Web pages (9,200 experiments): (Schedl et al.; 2011)

- **term frequency**
- **inverse document frequency**
- **virtual document modeling:** *concatenate* all Web pages/posts of the artist or perform *aggregation* via mean, max, etc.
- **normalization with respect to document length**
- **similarity measure**
- **index term set**

Abbr. / Term Set	Cardinality	Description
TS_A - all_terms	C224a, QS_A: 38,133 C224a, QS_M: 19,133 C3ka, QS_A: 1,489,459 C3ka, QS_M: 437,014	All terms (stemmed) that occur in the corpus of the retrieved Twitter posts.
TS_S - scowl_dict	698,812	All terms that occur in the entire SCOWL dictionary.
TS_N - artist_names	224 / 3,000	Names of the artists for which data was retrieved.
TS_D - dictionary	1,398	Manually created dictionary of musically relevant terms.
TS_L - last.fm_toptags	250	Overall top-ranked tags returned by last.fm 's <i>Tags.getTopTags</i> function.
TS_F - freebase	3,628	Music-related terms extracted from Freebase (genres, instruments, emotions).

Large-Scale Study

Investigating different aspects in modeling artist term profiles from Web pages
(9,200 experiments): (Schedl et al.; 2011)

- **term frequency**
- **inverse document frequency**
- **virtual document modeling:** *concatenate* all Web pages/posts of the artist or perform *aggregation* via mean, max, etc.
- **normalization with respect to document length**
- **similarity measure**
- **index term set**
- **query scheme**

Abbr.	Query Scheme
QS_A	“artist name”
QS_M	“artist name”+music

Large-Scale Study

Investigating different aspects in modeling artist term profiles from Web pages
(9,200 experiments): (Schedl et al.; 2011)

- **term frequency**
- **inverse document frequency**
- **virtual document modeling:** *concatenate* all Web pages/posts of the artist or perform *aggregation* via mean, max, etc.
- **normalization with respect to document length**
- **similarity measure**
- **index term set**
- **query scheme**

Interesting Findings

- modeling artists as *virtual documents* is preferable (Schedl et al.; 2011)
- using query scheme “*artist*”+*music* outperforms “*artist*”
- *normalization* does not yield a statistically significant difference
- standard *cosine similarity* measure does not yield the very best results, but the most stable ones (varying other parameters)
- *consistent results* among the (top-ranked) variants for two collections
- *minor change in one parameter* can have a huge impact on performance
- overall winners in terms of term weighting functions:

TF_C3.IDF_I

TF_C3.IDF_H → logarithmic formulations for TF and IDF

TF_C2.IDF_I



Search Engine Page Count for Similarity

(Schedl et al.; 2005)

Idea: expect entities that occur frequently in the same context to be similar → search engine page counts estimates indicate this

For all pairs of artists: query “**artist 1**” “**artist 2**” +**music** +**review**

For each artist: query “**artist**” +**music** +**review**



Use **page counts** for sim. (results in quadratic page count matrix)

$$sim_{pc_cp}(A_i, A_j) = \frac{1}{2} \cdot \left(\frac{pc(A_i, A_j)}{pc(A_i)} + \frac{pc(A_i, A_j)}{pc(A_j)} \right)$$

To avoid quadratic number of queries: download top 100 pages for each artist and parse for occurrences of other artists (linear complexity)

NB: asymmetry of *pc* matrix can be used to identify prototypical artists!

Music Information Extraction from Web Pages

Web data is a rich source for all types of meta-data and semantic relations

Methods from NLP, IE, Named Entity Detection for data extraction

- **Genres, Moods, Similarities** using Rule Patterns
(Geleijnse, Korst; 2006)
- **Band Members and Line-Up** using Rule Patterns
(Schedl, Widmer; 2007)
- **Band Members, Discography, Artist Detection** (rule based)
(Krenmair; 2010)
- **Band Members, Discography** using Supervised Learning
(Knees, Schedl; 2011)
- **Album cover** detection and extraction
(Schedl et al., ECIR 2006)

Web-Based Texts for Indexing and Retrieval

- Use Web data to transform music retrieval into a text retrieval task
- Find associated (or associable) texts and use them instead of music
- Allows for diverse and semantic queries
(e.g, “chilled music”, “great riffs”)

Search Sounds (Celma et al.; 2006)

Crawl lists of RSS feeds and use Weblog entries to index pieces

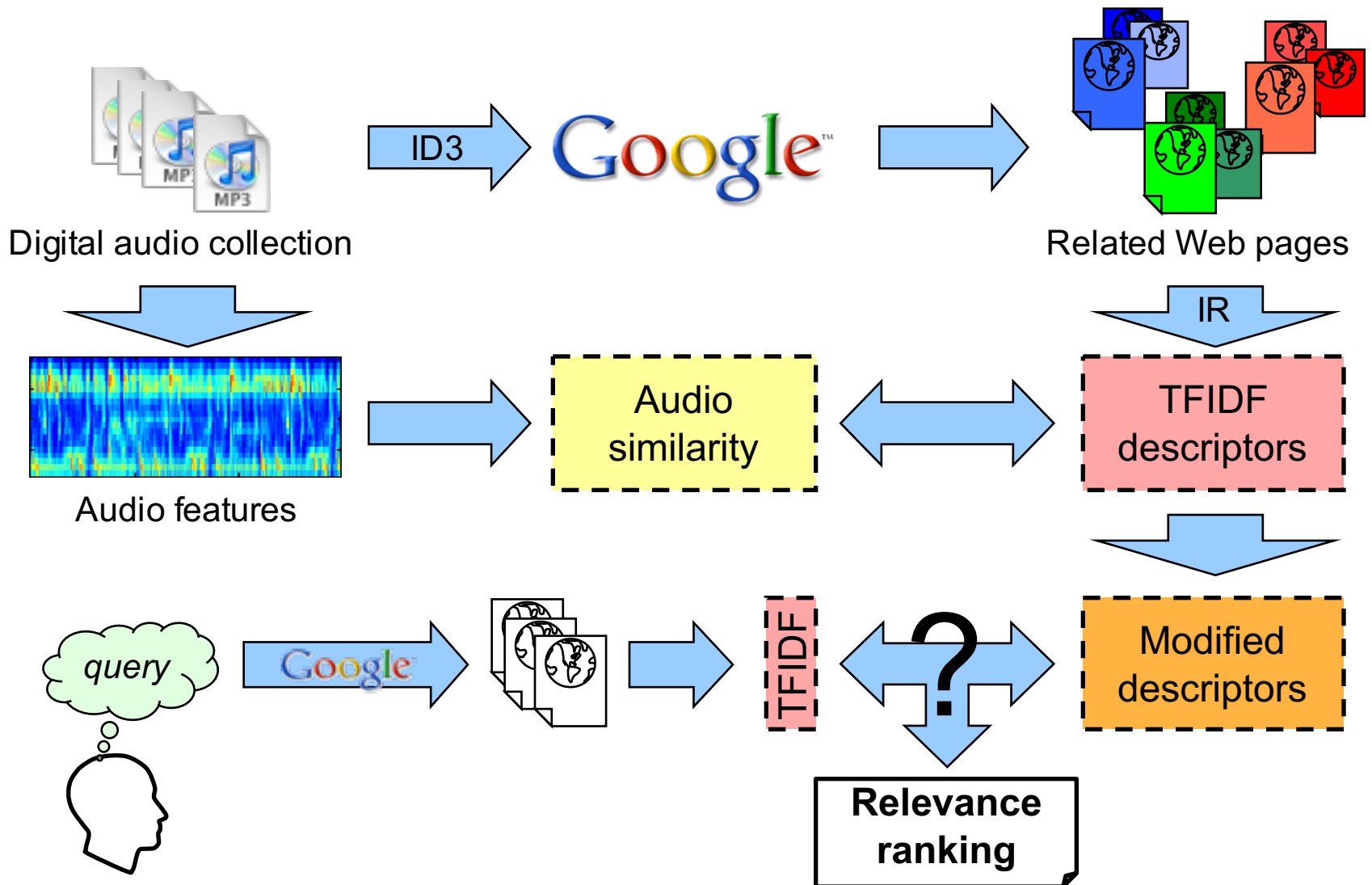
Squiggle (Celino et al.; 2006)

Combine meta-data databases (like MusicBrainz) for rich indexing

Gedoodle (Knees et al.; SIGIR 2007)

Query Google and combine Web pages to index pieces

Gedoodle




Gedoodle

(Knees et al.; SIGIR 2007)

- For each track: join 100 Google results of
 - “artist” music
 - “artist” “album” music review
 - “artist” “title” music review -lyrics
- Combine all pages into one virtual document
- Create normalized TFIDF vector for each track
- Include audio similarity for vector modification and dimensionality reduction

Gedoodle (Example queries)



Search for Music

Results 18 - 27 of 1691 for **damon albarn**. (0.02 seconds)

Music Is My Radar
by blur
from the album: *blur: the best of*
Genre: Alternative - 192 kBit/s - length: 5:29 min. [Listen](#)

On Your Own
by blur
from the album: *blur: the best of*
Genre: Alternative - 192 kBit/s - length: 4:27 min. [Listen](#)

Girls & Boys
by blur
from the album: *blur: the best of*
Genre: Alternative - 192 kBit/s - length: 4:19 min. [Listen](#)

There's No Other Way
by blur
from the album: *blur: the best of*
Genre: Alternative - 192 kBit/s - length: 3:14 min. [Listen](#)

Feel Good Inc.
by Gorillaz
from the album: *FM4 Soundselection 12*
Genre: Alternative - 192 kBit/s - length: 4:20 min. [Listen](#)

Say So What
by Graham Coxon
from the album: *Uncut - 2006.05*
Genre: Rock - 192 kBit/s - length: 3:05 min. [Listen](#)

Slash Dot Dash
by Fatboy Slim
from the album: *The Greatest Hits: Why Try Harder*
Genre: Electronic - 192 kBit/s - length: 2:55 min. [Listen](#)



Search for Music

Results 1 - 10 of 1774 for **smooth and relaxing**. (0.02 seconds)

Joy And Pain
by Count Basic
from the album: *Moving In The Right Direction*
Genre: Acid Jazz - 168 kBit/s - length: 6:25 min. [Listen](#)

Higher
by Count Basic
from the album: *Bigger & Brighter*
Genre: Acid Jazz - 192 kBit/s - length: 4:00 min. [Listen](#)

Sweet Luis
by Count Basic
from the album: *Moving In The Right Direction*
Genre: Acid Jazz - 158 kBit/s - length: 5:11 min. [Listen](#)

Got To Do
by Count Basic
from the album: *Moving In The Right Direction*
Genre: Acid Jazz - 167 kBit/s - length: 4:58 min. [Listen](#)

John Lee Huber
by Tosca
from the album: *J.A.C.*
Genre: Electronica/Dance - 192 kBit/s - length: 4:33 min. [Listen](#)

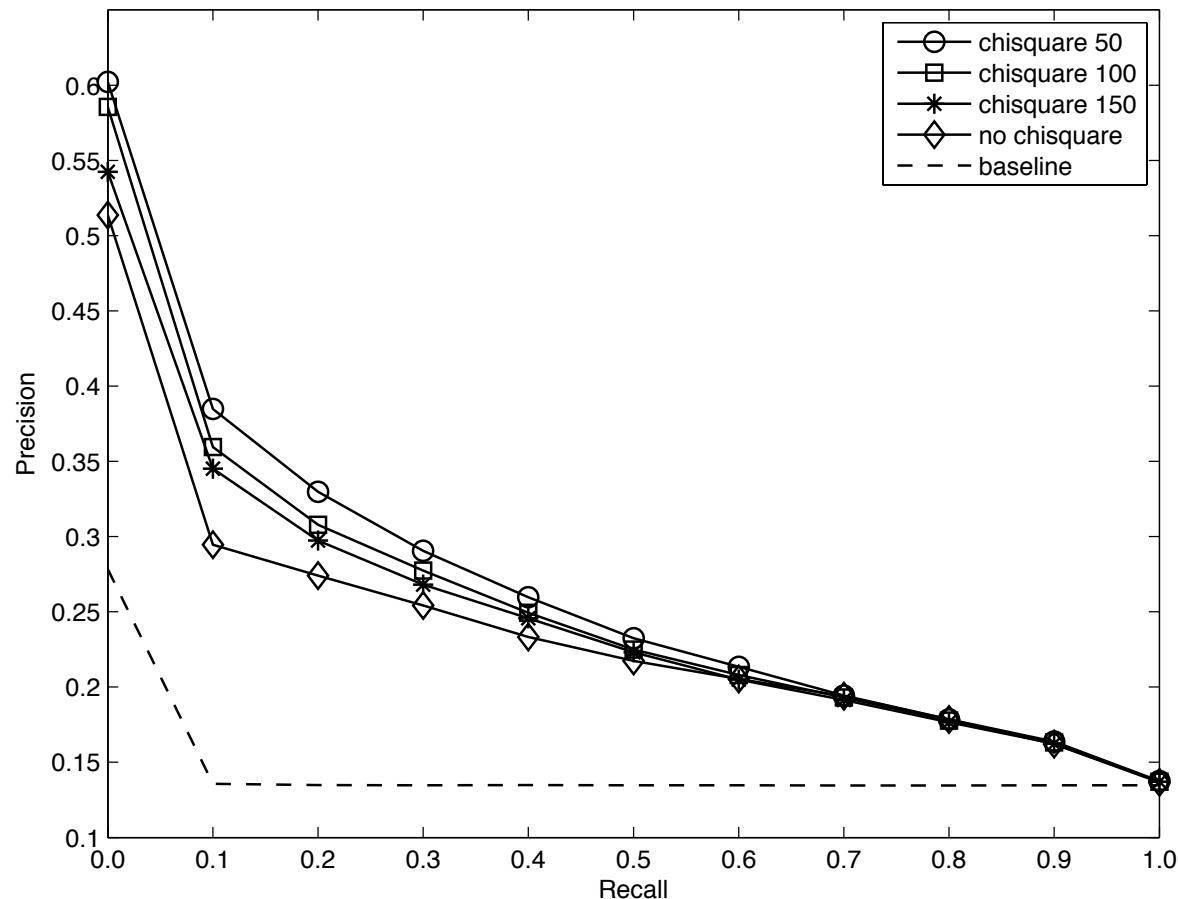
No More Olives
by Tosca
from the album: *J.A.C.*
Genre: Electronica/Dance - 192 kBit/s - length: 6:02 min. [Listen](#)

Naschkatze
by Tosca
from the album: *J.A.C.*
Genre: Electronica/Dance - 192 kBit/s - length: 4:34 min. [Listen](#)



Gedoodle Results

Effects of TFIDF feature space pruning using content-similarity-based χ^2 -test (Knees et al.; SIGIR 2007)



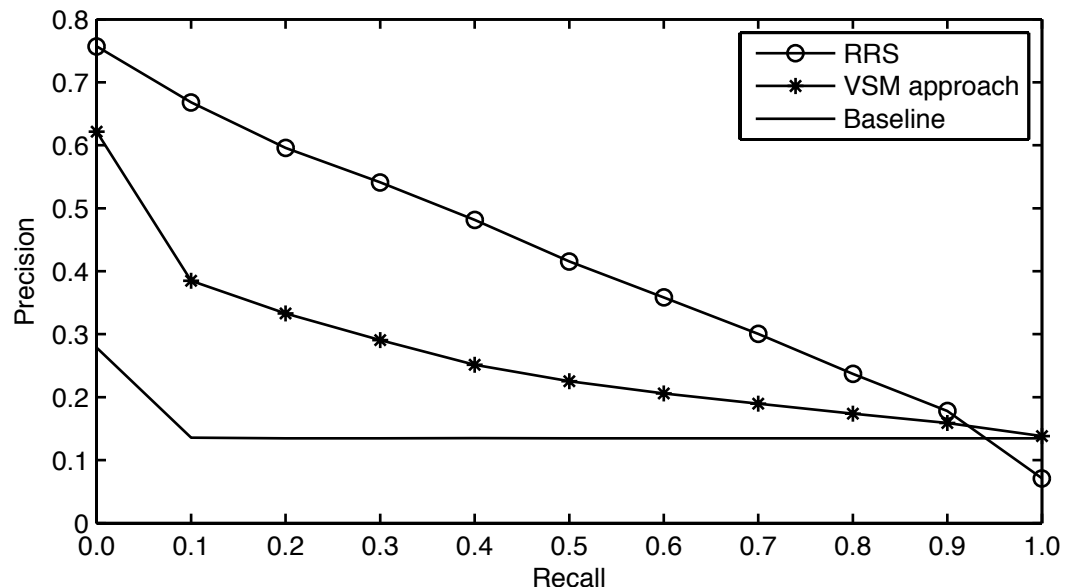
Gedoodle Results

Alternative: Document-centered ranking (Knees et al.; ECIR 2008)

- Indexing of all web documents in standard index
- Music query addresses this index
- Music ranking calculated from web doc ranking according to

$$RRS(m, q) = \sum_{p \in D_m \cap D_q} \frac{1}{1 + |D_q| - rank(p, D_q)}$$

Comparison with
vector space model



Semantic Querying via Auto-Tagging

- Use machine learning techniques to predict tags (labels) based on song features (content, context, or combination)
- Automatic description of music (browsing) and automatic generation of indexing terms for retrieval
- Mitigates “cold-start problem” in social tagging

Automatic Record Reviews (Whitman, Ellis; 2004)

Regularized least squares learning on TFIDF-Web and cepstral features

Autotagger (Bertin-Mahieux et al.; 2008)

Ensemble classifier to map MFCCs, autocorrelation, Const-Q. to Web tags

Semantic Music Discovery (Turnbull et al.; SIGIR 2007, 2009)

Combines timbre, harmony, Web texts, and Web tags to predict user labels

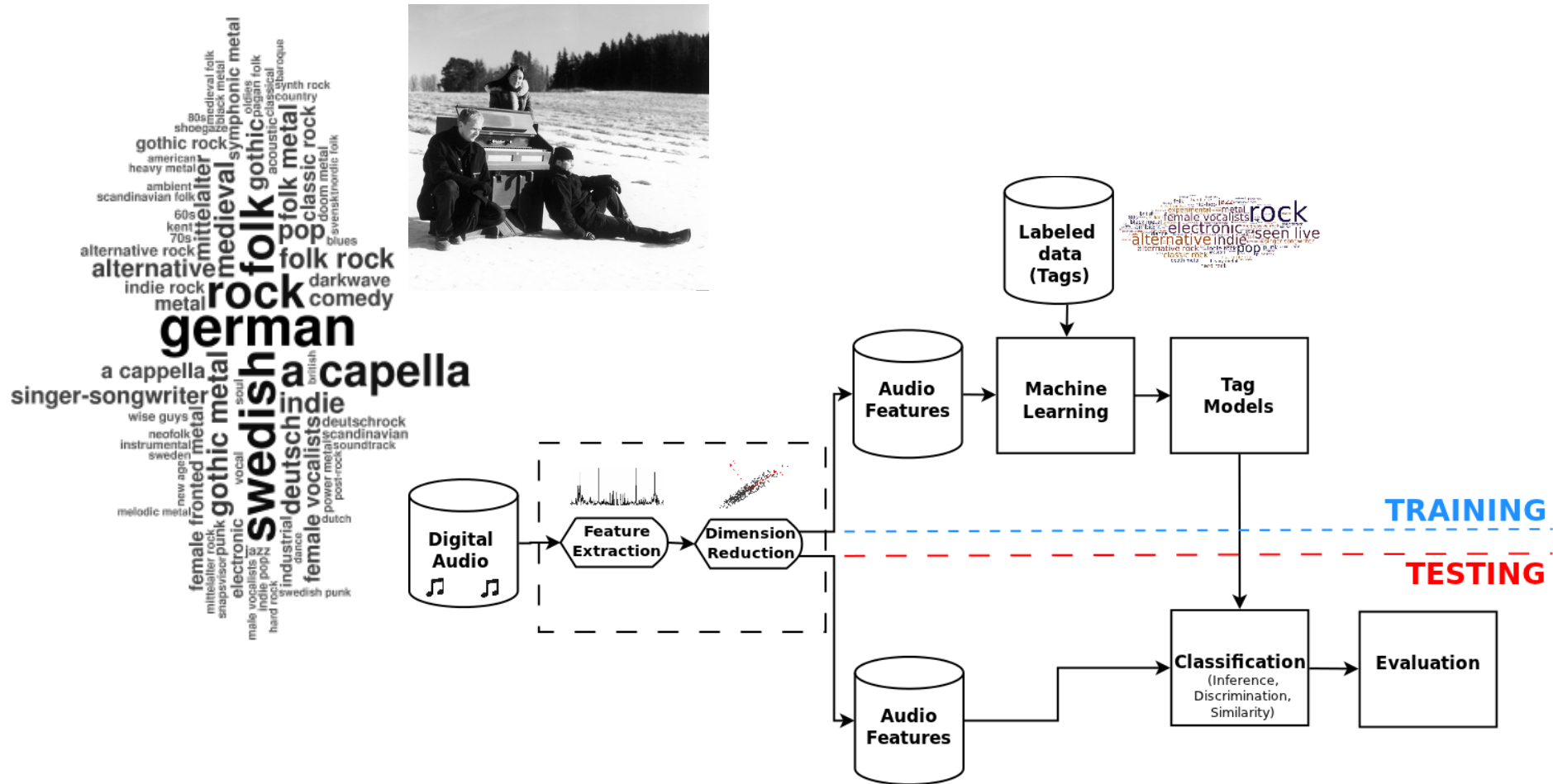
Semantic Annotation of Music Collections (Sordo; 2012)

Propagation of tags through audio similarity

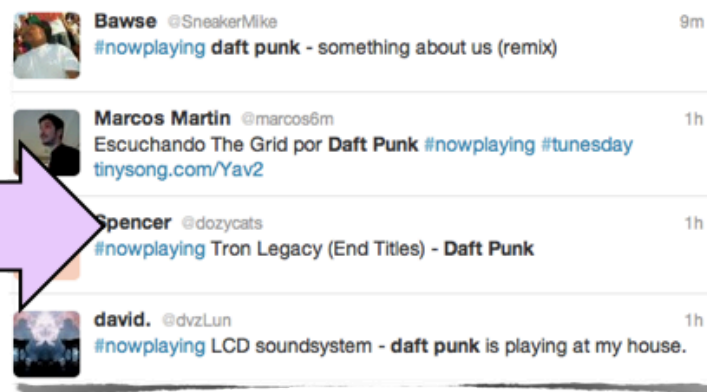
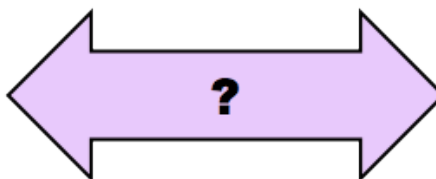
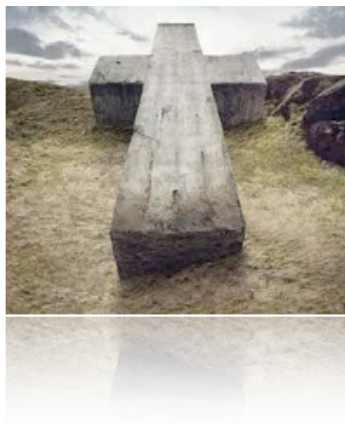
Auto-Tagging/Retrieval by Tag

Learning indexing labels from content features

(Sordo; 2012)



Microblogs as Text Sources



(Schedl; 2012a)

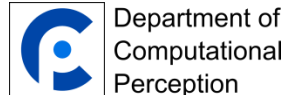
(+music)



artist B



similarity estimate



Microblogs as Text Sources: Scheme

(Schedl; 2012a)

Large-scale study similar to (Schedl et al.; TOIS 2011)

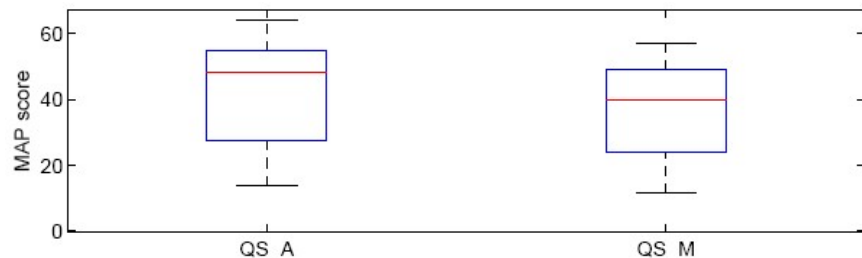
Investigating different aspects in modeling artist term profiles from microblogs
(23,100 experiments):

- **query scheme**
- **index term set**
- **term frequency**
- **inverse document frequency**
- **normalization with respect to document length**
- **similarity measure**

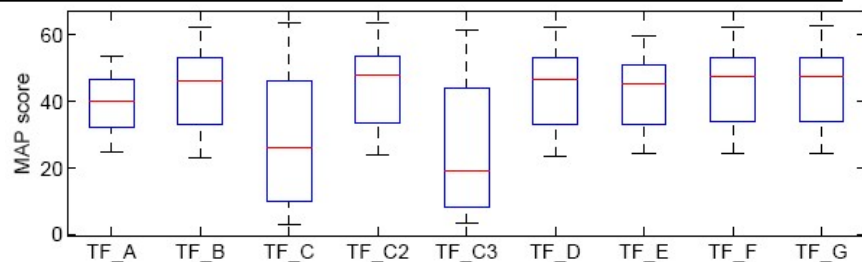


Microblogs as text-based source: Results

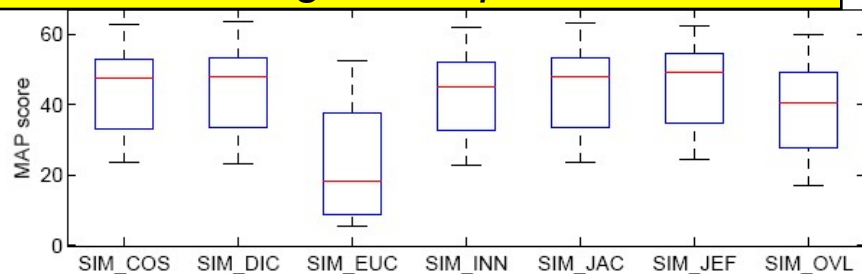
(Schedl; 2012a)



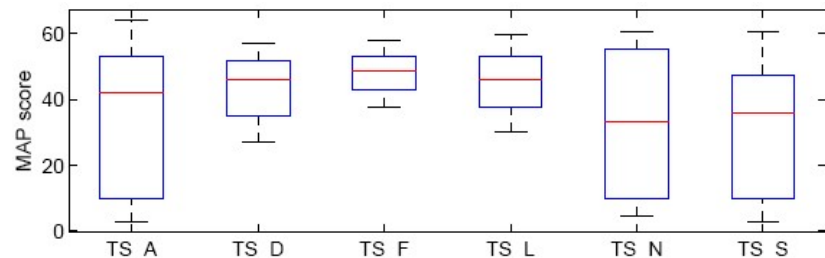
use query scheme “artist name”



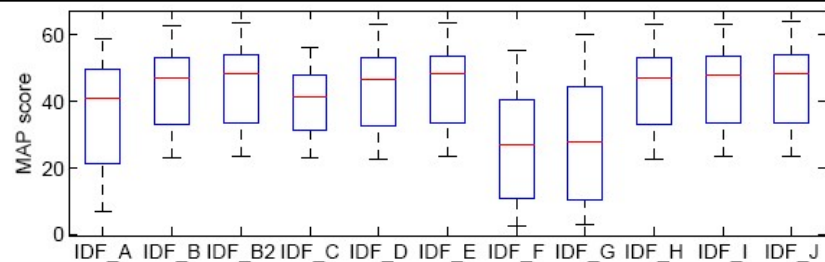
use *log* or *Okapi BM25*



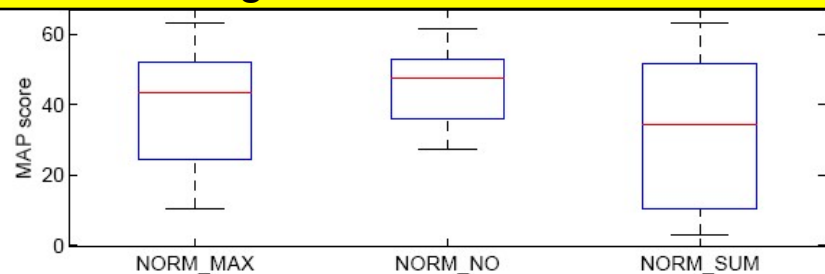
don't use Euclid; use Jeffrey or Inner Prod.



music-specific dictionary favorable



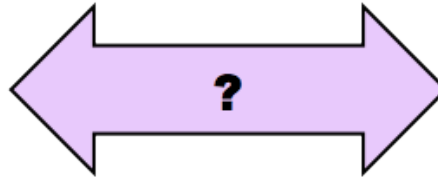
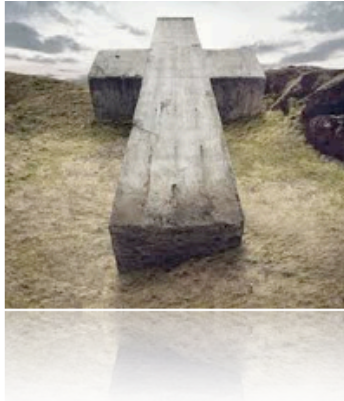
use *logarithmic* formulations



no document length normalization



Product Reviews as Text Sources



4 of 6 people found the following review helpful:

★★★★★ **Bolder than Cross: prog-dance in the making..**, 25 Oct 2011
By [Kieren Thomson "Kieren"](#) (Brading, Isle Of Wight) - [See all my reviews](#)
This review is from: Audio, Video, Disc. (Audio CD)
Imagine if the Bee Gees decided to make a prog-rock album, or that Jeff Wayne's War Of The Worlds was conducted in a disco. That's how Justice have played out on their follow up to one of the greatest dance albums of the last 10 years, Cross. They've dropped the samples and have made an electro-instrumental album with tinges of progressive rock.

A wonder to behold, Audio Video Disco contains nods to some of the greatest rock of the 70s, but keeps the great elements of experimental dance from the 60s. Highlights include Canon - a club-stomper built for Daft Punk, and Helix - a nod to the last album but with bigger and bolder synths.

It's not Cross, but it doesn't need to be. It's a bold, guitar-laden album built on rock instead of experimental-dance. Regicide.

Help other customers find the most helpful reviews
Was this review helpful to you? [Yes](#) [No](#) [Report abuse](#) [Permalink](#) [Comment](#)

0 of 1 people found the following review helpful:

★★★★★ **great great great**, 12 Nov 2011
By [KJ Coleman \(england\)](#) - [See all my reviews](#)
This review is from: Audio, Video, Disc. (Audio CD)
the first album is my favourite dance album ever initially a bit of a shock - the prog rock/ heavy metal direction but after several listens its still quite cool! spinal tap it is not

Help other customers find the most helpful reviews
Was this review helpful to you? [Yes](#) [No](#) [Report abuse](#) [Permalink](#) [Comment](#)

Most Recent Customer Reviews

★★★★★ **Try all of this...**
The negative reviews dragging this album down are silly. You can not compare Justice to anyone. No longer do they present as the angry high pitched mates of Daft Punk.
[Read more](#)
Published 1 month ago by A. Livingstone

★★★★★ **Brilliantly Innovative, but in the same streak**
From the very moment I heard "Civilisation" in the Adidas advert, I got excited about this album. Listening to it did not leave me disappointed at all!
[Read more](#)
Published 1 month ago by Raser Taylor

★★★★★ **On'n'On'n'On'n'On...**
Justice has seriously tamed the "Audio, Video, Disc."

If you are looking for the most soundscapes and swashbuckling "Cross" in it's...
[Read more](#)
Published 1 month ago by Drapace

★★★★★ **"New" Justice**
This album is different from their first album, but although I am a huge fan of "Cross", I do enjoy this new album a lot, too.
[Read more](#)
Published 1 month ago by Christian Schmeier

★★★★★ **Such a disappointment for a hardcore Justice fan**
Having seen Justice live on at least 5 occasions and being a big fan and proud owner of Cross and A Cross the Universe, I am disappointed to say that there is no such

Help other customers find the most helpful reviews
Was this review helpful to you? [Yes](#) [No](#) [Report abuse](#) [Permalink](#) [Comment](#)

11 of 13 people found the following review helpful:

★★★★★ **debut daft punk**, 22 Jan 2004
By [E. P. "spideredd"](#) (Suffolk, England) - [See all my reviews](#)
This review is from: Audio, Video, Disc. (Audio CD)
After discovery, so my expectations were a little high. I have with this album is that the songs are extreme.

Compare the two albums, but I feel that homework has the better layout and appeal.

One or two friends of mine won't listen to this album because there is little to break it up. This is the only reason that I haven't given this the whole five stars.

All in all a good, if somewhat strange album. I'd recommend that anyone should at least listen to it.

Help other customers find the most helpful reviews
Was this review helpful to you? [Yes](#) [No](#) [Report abuse](#) [Permalink](#) [Comment](#)

1 of 1 people found the following review helpful:

★★★★★ **Terrible!!!**, 19 Oct 2011

Most Recent Customer Reviews

★★★★★ **Superior house music**
This album will never be beaten, much much imitated but never equalled. Play it loud and proud as this was released in 1996 and still "Around the world" sounds as "fresh" as it...
[Read more](#)
Published 4 months ago by Mr. G. Bellinger

★★★★★ **If summer was a sound it would sound like this**
I think I got Daft Punk backwards. Beyond hearing the odd single and track in a bar I didn't really pay them a lot of mind.
[Read more](#)
Published 6 months ago by Christopher Long

★★★★★ **Quality**
Had this album on vinyl when it first came out. Since then lost that so had to get it on cd. Still sounds as fresh as it did back then!! Absolute quality music!!!!
[Read more](#)
Published 13 months ago by Craig J. Gendron

★★★★★ **Such a great album!**
This CD puts a smile on my face. This is sooooo good. Look up the video from around the world and you're sold. The rest of the album is just as good.
[Read more](#)
Published 13 months ago by Xee

★★★★★ **Brilliant**
As a born-again Daft Punk fan I bought this having not long ago bought Discovery, and I love it.

Other people can express what's great about this more eloquently than...
[Read more](#)
Published 21 months ago by Mrs Whitmore

Product Reviews as Text Sources

Exploiting sources such as *Amazon.com* or *Epinions.com* (Hu et al.; 2005)

This review is from: [Ray of Light \(Audio CD\)](#)

This is Madonna's work of art. And this CD is the very best collection of any music she has ever produced since "Erotica." Madonna's lyrics are beautiful and strong because even after 9 years it still stands the test of time. It's completely impossible for this CD to be dated; with the electronica feel to it and fast moving dance numbers, such as the title-track this CD was way ahead of its time. Even in the double-00's "Ray of Light" is still very important as both a dance record and a record of reflection and interpersonal renewal.

This review is from: [Never Gonna Give You Up \(MP3 Download\)](#)

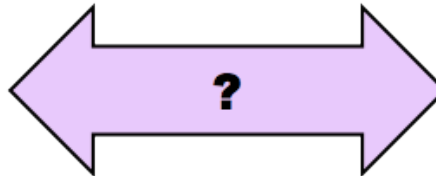
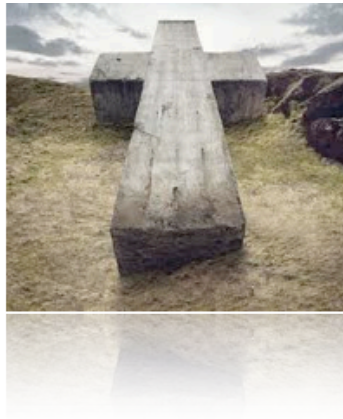
This is truly Astley's greatest opus.

The track is flawless. It is instantly accessible, but features many hidden layers and pleasures that cannot be discovered upon the first listen alone. With this and all of his other fantastic work, it's no wonder that Radiohead calls Astley their "greatest inspiration."

Allows for sentiment analysis and associated rating prediction

Very prone to attacks (remedy: consider “helpfulness” ratings)

Community Tags as Text Sources



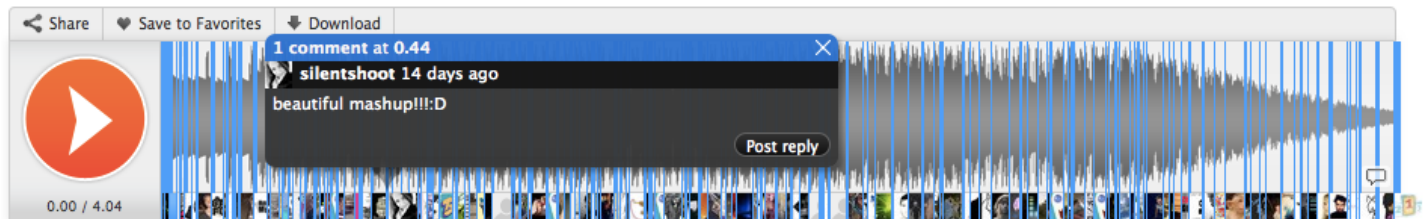
00s alternative ambient chillout club cool **dance** dance punk dance-punk death 00s 80s 90s alternative alternative rock ambient awesome big beat blues chillout classic
metal digital dirty electro disco distortion ed banger **electro** electro dance rock club daft punk **dance** disco dub **electro** electro house electroclash
electro house electroclash **electronic** **electronica** electropop **electronic** **electronica** electropop experimental favorites
elektro eletronic experimental favourite france **french** french electro french house french touch funk funky great
french touch funk funky german glitch hardcore hardcore punk indie industrial instrumental japanese jazz love metal
indietronica instrumental justice love metal new rave noise nu rave party pop progressive house psychedelic psytrance punk robots rock soul
psychedelic punk rock sexy synthpop techno thrash metal trance want to see live soundtrack synth synthpop **techno** trance trip-hop

Tag Sources

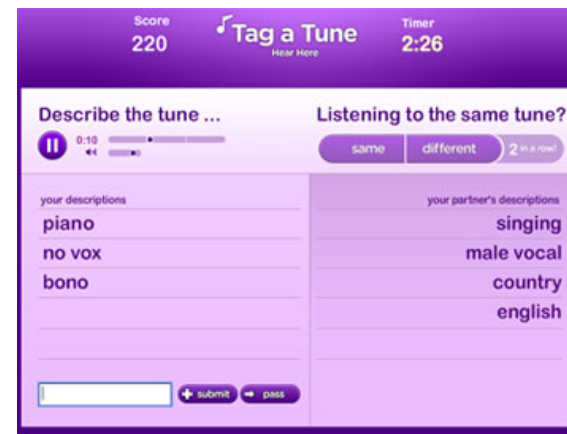
- **Community**
e.g., Last.fm

1960s 60s acoustic american bacharach baroque baroque pop
boltonesque brill building pop burt bacharach chill classic composer disco driving
easy easy listening everything favorite artists favorites
film music film score fusion genius god great innovators guitar hal david inspirerande
instrumental jazz lounge male male vocalists master melancholy music to
warm the heart and hands my ancients my tag oldies outstanding pop relax
rock score sexy singer-songwriter smooth songwriter sophistopop soul
soundtrack space age pop swing symphonic pop us usa virtuoso vocal 2005

e.g., Soundcloud (annotations along timeline)



- **Games with a purpose (GWAP)**
e.g., Tag-a-Tune
(Law, von Ahn; 2009)
- **Autotags** (see before)



Community Tags as Text Sources

Treating collections of tags (e.g., from Last.fm) as documents

(Pohle et al.; 2007) (Levy, Sandler; 2008) (Hu et al.; 2009)

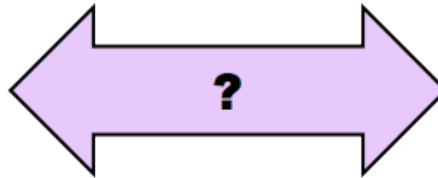
- Retrieve tags for artist or track from Last.fm
- Cleaning of noisy and redundant tags:
manually or automatically (Geleijnse et al.; 2007)
- List of collected terms is treated as text document and TF·IDF'd
(Levy, Sandler; 2007)
- Optionally, LSA to reduce dimensionality
- Comparison of vectors via cosine similarity (or overlap score)
- Data often available in standardized fashion, dedicated terms for music
- Lower dimensionality
e.g., 13,500 tags vs. >200,000 Web terms (Levy, Sandler; 2007)
- Depends on community, needs annotators
- Hacking and Attacks!



Lyrics as Text Source



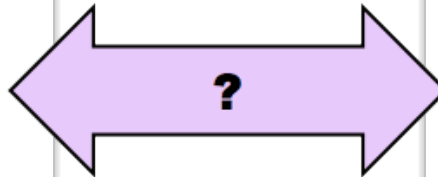
Before day break there was none
And as it broke there was one
The Moon, the sun, it goes on 'n' on
The winter battle was won
The summer children were born
And so the story goes on 'n' on
Come woman if your life beats
Those we buried with the house keys
Smoke and feather where the fields are green
From here to eternity
Come woman on your own time



Around the world, around the world
Around the world, around the world
Around the world, around the world
Around the world, around the world

Around the world, around the world
Around the world, around the world
Around the world, around the world
Around the world, around the world

Around the world, around the world
Around the world, around the world
Around the world, around the world



Lyrics as Text Source

Topic Features (Logan et al.; 2004)

- Typical topics for lyrics are distilled from a large corpus using (P)LSA (“Hate”, “Love”, “Blue”, “Gangsta”, “Spanish”)
- Lyrics are transformed to topic-based vectors, similarity is calculated via L_1 distance
- Alternative approaches use TF·IDF with optional LSA and Stemming for **Mood Categorization** (Laurier et al.; 2009) (Hu et al.; 2009)

Rhyme Features (Mayer et al.; 2008) (Hirjee, Brown; 2009)

- Phonetic transcription is searched for patterns of rhyming lines (AA, ABAB, AABB)
- Frequency of patterns + statistics like *words per minute*, *punctuation freq.* etc.

Other Features (Mahedero et al.; 2005) (Hirjee, Brown; 2009, 2010)

- Language, structure, artist identification

Text-based Similarity Approaches: Summary

	Web-Terms	Microblogs	Reviews	Tags	Lyrics
Source	Web pages	platform	shops, platform	Web service	portal
Community-based	depends	depends	yes	yes	no
Level	artists	artists (tracks)	albums	artists (tracks)	tracks (artists)
Feature Dimensionality	very high	high	possibly high	moderate	possibly high
Specific Bias	low	low	personal	community	none
Potential Noise	high	high	low	moderate	low

Combining Music Content and Context

Music content and context features

- Stem from different sources
- Describe different dimensions
- Can be combined beneficially

Different approaches shown in three applications in the following

- Rank-based similarity combination → *MusicSun*
- Context-based filtering of content similarity → “*The Wheel*”
- Description of content-based clustering → *nepTune*

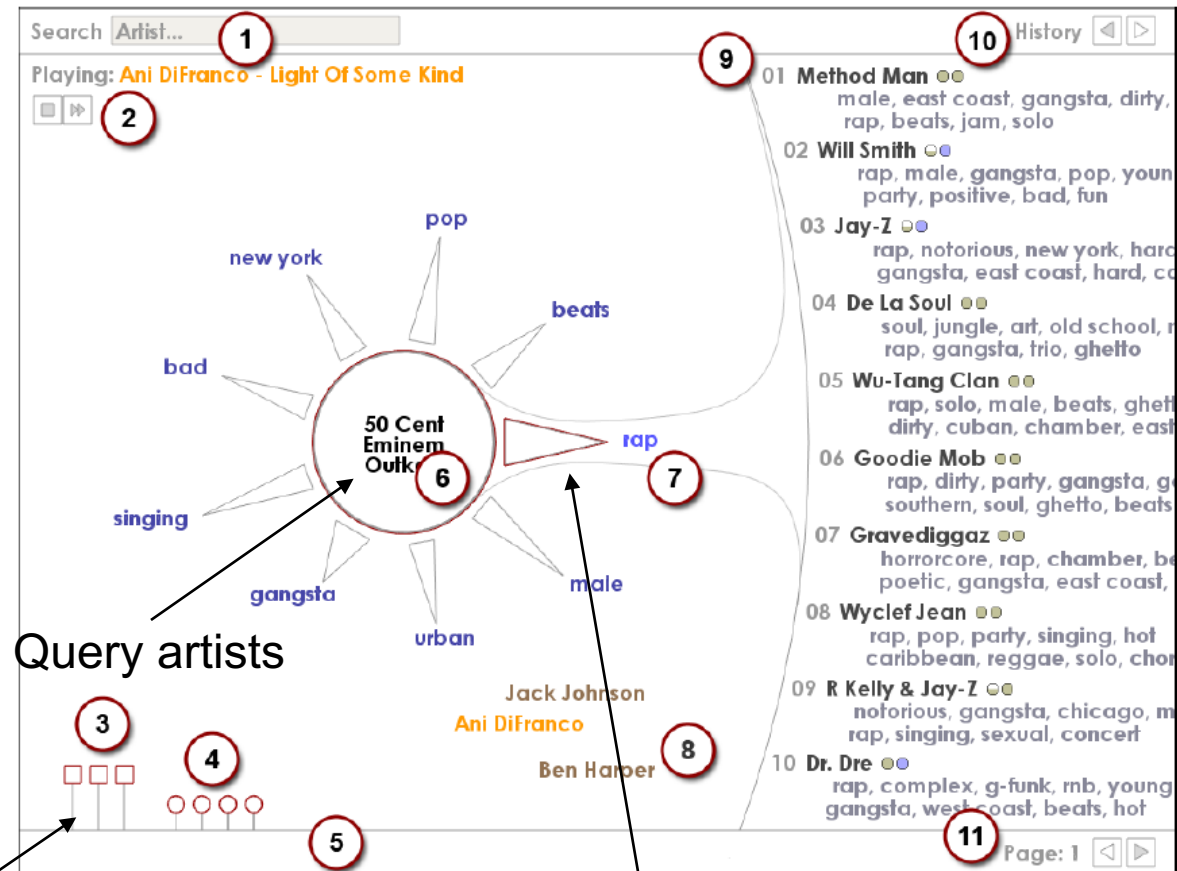
Many more exist, e.g., transformed, “semantic” audio space learned from tag embedding

Similarity Combination for Browsing

“MusicSun”

(Pampalk, Goto; 2007)

- Interactive “Artist Recommender”
- Recommendation is influenced/directed by selecting relevant similarity dimensions
- Combines different similarity measures



3 types of similarity: audio, web-based, word
overall similarity = weighted average of ranks

Relevant word dimension

Contextual Filtering for Playlist Generation


(Knees et al.; 2006)

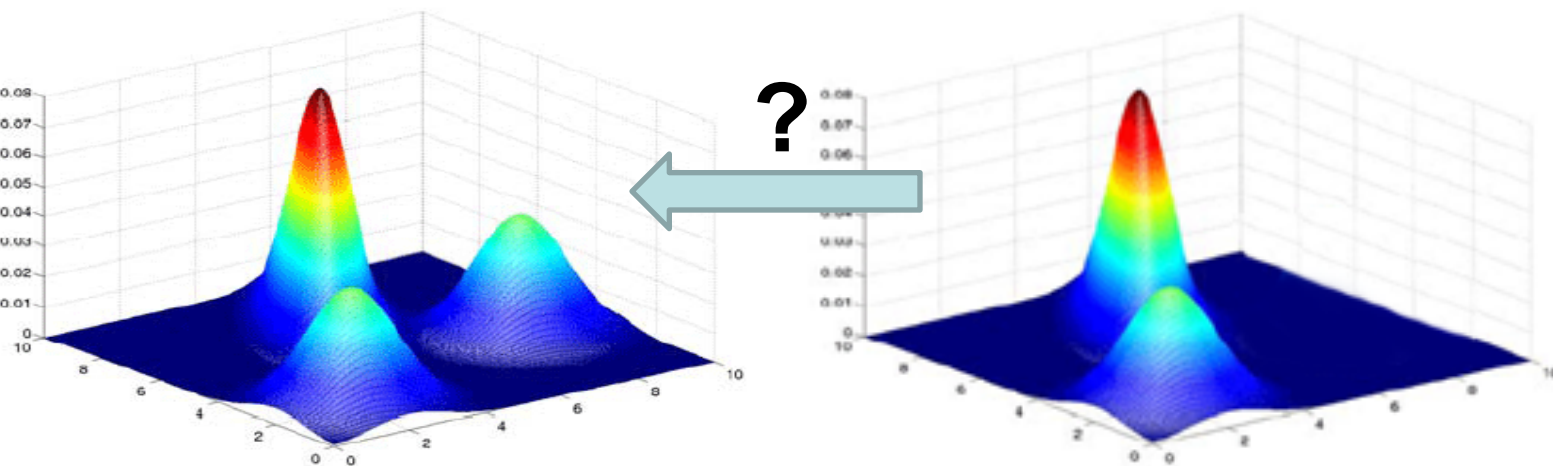
- Idea: combine music content + music context features to improve and speed up playlist generation
- Application scenario: “The Wheel” – create a circular playlist containing all tracks in a user’s collection (consecutive tracks as similar as possible)
- Approach: use Web features to confine search for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



Contextual Filtering for Playlist Generation

(Knees et al.; 2006)

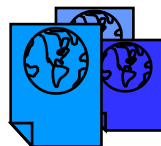
- Audio/content features: 
 - compute Mel-Frequency Cepstral Coefficients (MFCC)
 - model song's distribution of MFCCs via Gaussian Mixture Models (GMM)
 - estimate similarity between two songs A and B by sampling points from A 's GMM and computing probability that points “belong to” GMM of B



Contextual Filtering for Playlist Generation

(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



Contextual Filtering for Playlist Generation

(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



Contextual Filtering for Playlist Generation

(Knees et al.; 2006)



To obtain the final, hybrid similarity measure:

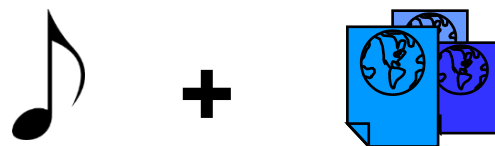


Folk-Rock(4) Rap(4) Jazz(1) Punk-Rock(1)	Electronica(2)	Electronica(5)	Electronica(1)	Electronica(16) Acid Jazz(1)
Folk-Rock(1) Italian(1)	Electronica(1)	Acid Jazz(1)		Acid Jazz(1) Electronica(1)
Italian(3) Electronica(1)		Reggae(2) Italian(1)		Rap(2) A Cappella(1) Acid Jazz(1) Electronica(1)
Punk-Rock(4)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

Train Self-Organizing Map
(SOM) on artist Web features

Contextual Filtering for Playlist Generation

(Knees et al.; 2006)



To obtain the final, hybrid similarity measure:

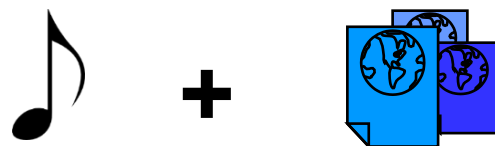


Folk-Rock(4) Rap(4) Jazz(1) Punk-Rock(1)	Electronica(2)	Electronica(5)	Electronica(1)	Electronica(16) Acid Jazz(1)
Folk-Rock(1) Italian(1)	Electronica(1)	Acid Jazz(1)		Acid Jazz(1) Electronica(1)
Italian(3) Electronica(1)		Reggae(2) Italian(1)		Rap(2) A Cappella(1) Acid Jazz(1) Electronica(1)
Punk-Rock(4)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

Contextual Filtering for Playlist Generation

(Knees et al.; 2006)



To obtain the final, hybrid

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

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

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



Contextual Filtering for Playlist Generation

(Knees et al.; 2006)

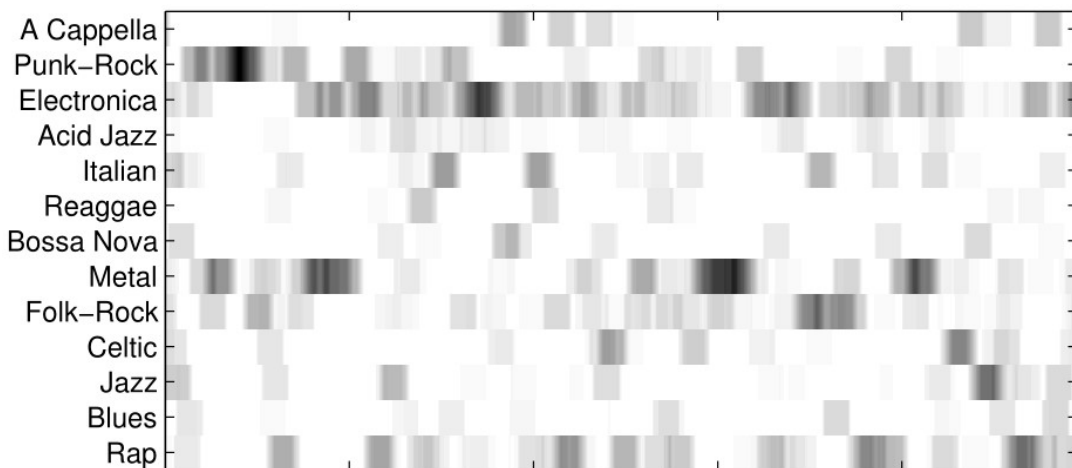
- Evaluation:

- 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)

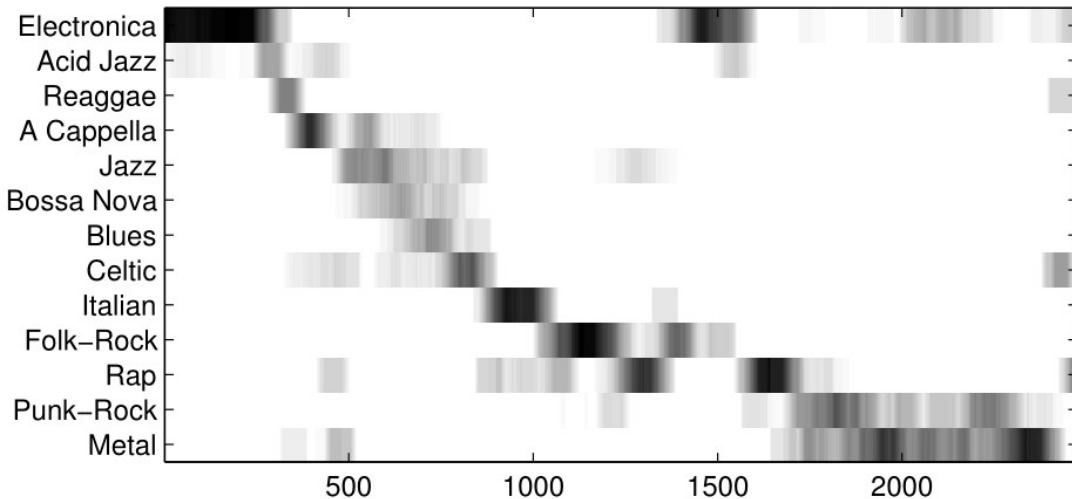


Contextual Filtering for Playlist Generation

(Knees et al.; 2006)



music content
similarity only



hybrid approach

Content Description for Music Browsing



nepTune – Structuring the Music Space

(Knees et al.; MM 2006)

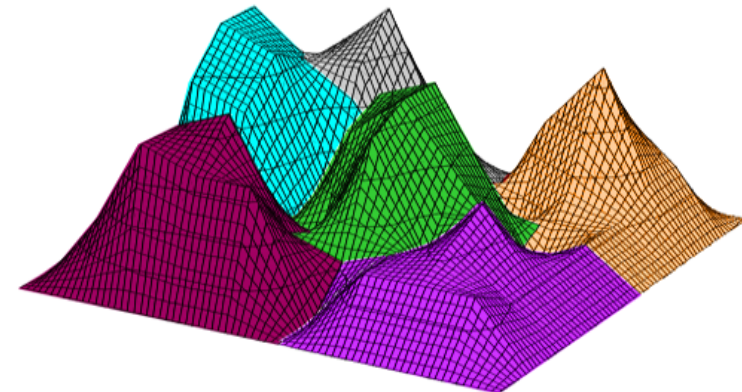
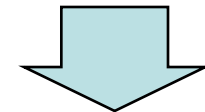
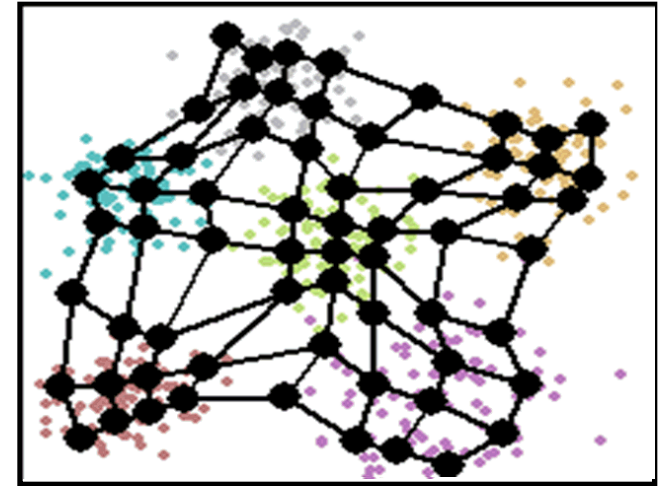
Clustering of music pieces

Each song corresponds to point in feature
(similarity) space

Self-organizing Map

High-dimensional data (content-based
features) is projected to 2-dim. plane

Number of pieces per cluster
→ landscape height profile



nepTune – Web-based Augmentation

(Knees et al.; MM 2006)

Automatic description of landscape via Web term extraction

