# Five Approaches to Collecting Tags for Music

Approach

Summary

Strengths

Weaknesses

Example

Algorithm

Survey

Experts are paid to annotate songs using a stardard form

**Custom-tailored vocabulary** High-quality annotations Strong labeling

Small, predetermined vocabulary Human-labor intensive Time consuming approach lacks scalability

> PANDORA created by the Music Genome Project™.

### CAL 500 Data Set [1]

Paid 55 undergraduates to annotate 500 songs by 500 artists using a vocabulary of tags. Each song was annoated my a minimum of 3 individuals.

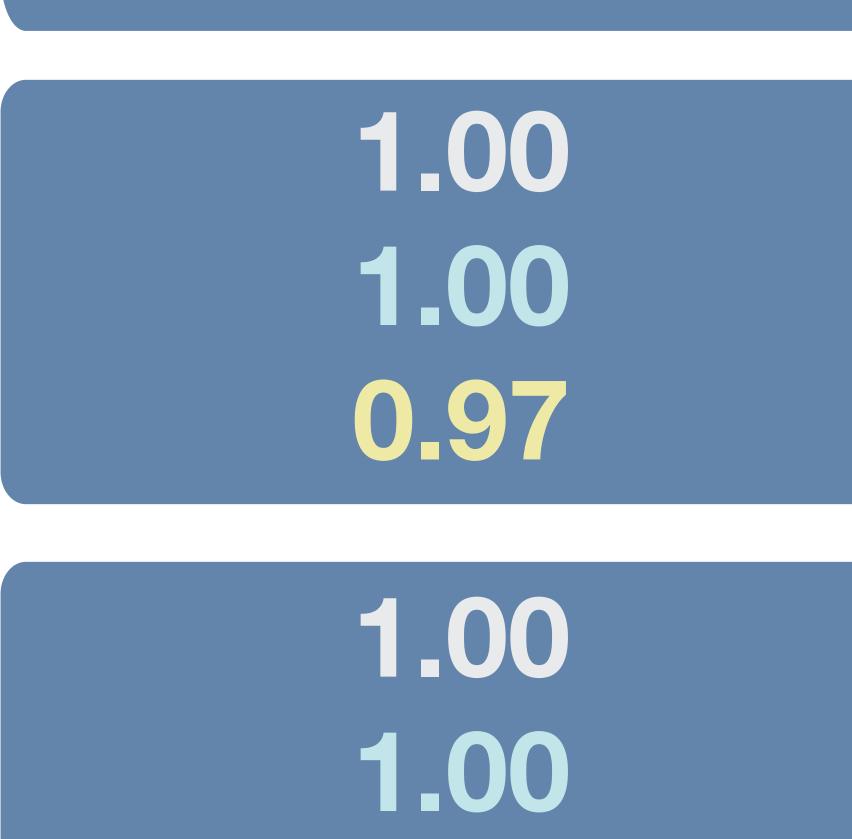
This data serves as the ground truth. There are 87 "long tail" songs from Magnatunes.

The vocabulary consists of 109 tags that relate to genre, instrumentation, emotion, usage, rhythm, vocals, and other musical characteristics.

S Density S AUC-ROC Top 10 Prec

a Density **o** AUC-ROC

**S** Top 10 Prec



0.57

# Social Tags

Large community contributes tags using a social network

Collective wisdom of crowds Unlimited vocabulary **Provides social context** 

Ad-hoc annotation behavior Produces weak labeling Sparse/missing data in long tail



#### Audioscrobbler

Attempted to collect a list of tags associated with each CAL500 song and each CAL500 artist from Last.fm's Audioscrobbler website. For each song, the song and artist lists were combined.

The combined list was matched to the CAL500 vocabulary. We attempted to use synonyms, alternative spellings, and wildcard matching to improve coverage.

Game Tags

Players produce tags as they play a video game

**Collective wisdom of the crowds** ast paced for rapid data collection Intertaining incentives produce high-quality tage

"Gaming" the system Difficult to create viral gaming experience Based on short clips, rather than songs



ListenGame [2] During a two week pilot study of ListenGame, we collected 16,500 tags for 250 of the CAL500 songs from 440 players.

Each of the 27,250 song-tag pairs were present 1.8 times on average.

 $\mathbf{0.23}$ 0.62 0.37

0.03 0.54 0.19

0.37 0.65  $\mathbf{0.32}$ 

Based on our experimental setup, the long tail results are misleading because the selection of songs in ListenGame is independent of popularity.

Please Contact: Douglas Turnbull <turnbull@cs.swarthmore.edu> Data, Papers, and additional Information can be found at: http://cosmal.ucsd.edu/cal/

# **Douglas Turnbull, Luke Barrington, Gert Lanckriet Computer Audition Laboratory, UC San Diego**

# Webtags

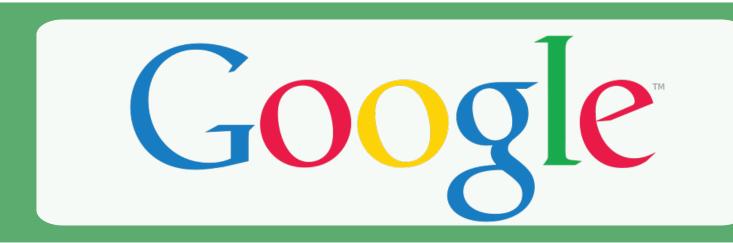




#### Analyze a corpus of music reviews, artist bios, blogs, discussion boards

Large corpus of publicly-available documents No direct human involvement Provides social context

**Text-mining introduces noise Produces weak labeling** Sparce/missing data in long tail



## **Relevance Scoring** [3]

- 1) Collect Corpus query google with song, artist and album
- 2) Query Corpus with Tag find most relevant song given tag

#### Site-Specific

use web documents from sites that are known to have high-quality content (Rolling Stones, AMG AllMusic, etc

0.67

0.66

0.32

Weighted RS weight pages by tag relevance

# Autotags

Annotate audio content using signal processing and machine learning

Not affected by cold-start problem No direct human involvement Produces strong labeling

**Computationally intensive** Limited by training data Based solely on audio content, no context



## Supervised Multilabel Model [1]

MFCC+Delta Feature Space One GMM per tag **Mixture Hiearchies EM to train GMMs** 

Produces "Semantic Multinomial" distribution over tag vocabulary for each novel song

**Top performing system in 2008 MIREX Audio Tag Classification Task** 

> 1.00 0.69  $\mathbf{0.33}$

1.00 0.70 0.27

0.18

0.25 0.56

[1] Turnbull, Barrington, Torres, Lanckriet Semantic Annotation and Retrieval Music and Sound Effects. TASLP 2008 [2] Turnbull, Liu Barrington, Lanckriet Using Games to Collect Semantic Information About Music. ISMIR 2007 [3] Knees, Pohle, Schedl, Schnitzer, Seyerlehner A Document-Centered Approach to a Natural Language Music Search Engine. ECIR 2008 [4] Barrington, Yazdani, Turnbull, Lanckriet Combining Feature Kernels for Semantic Music Retrieval. ISMIR 2008 [5] Turnbull Design and Development of a Semantic Music Discovery Engine. Ph.D. Thesis, UC San Diego 2008

# Hybrid

**Combination of approaches** 

Combine social context and audio content Use strengths, remove weaknesses Multi-tiered approach to cold-start problem

Increased system complexity Combining data sources can be tricky



Rank-Based Interleaving Given a tag, rank songs based on their best rank according to other appoaches

#### Other hybrid approachs

- 1) Kernel Combination [4]
- 2) RankBoost [5]
- 3) Calibarated Score Averaging [5]

1.00 0.74 0.38

1.00 0.71 0.28