Discover the Music You Want:

Building a Music Search Engine Using Audio Content and Social Context

Douglas Turnbull

Computer Audition Lab

UC San Diego

dturnbul@cs.ucsd.edu

Work with Luke Barrington, David Torres, and Gert Lanckriet

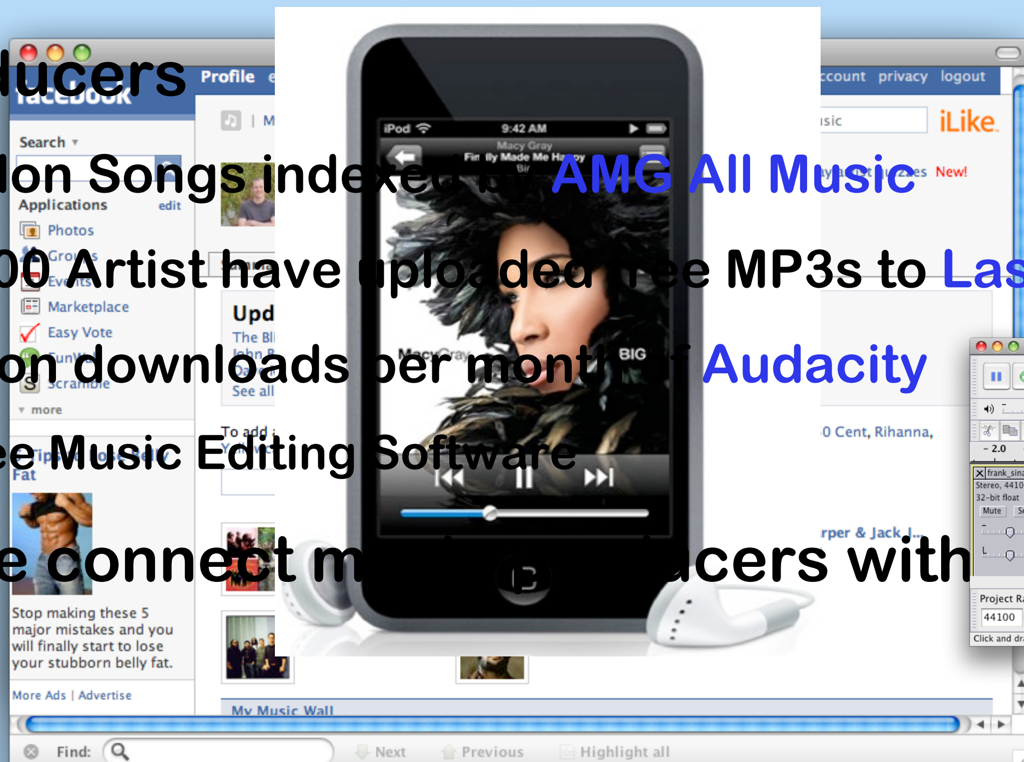
‘Age of Music Proliferation’

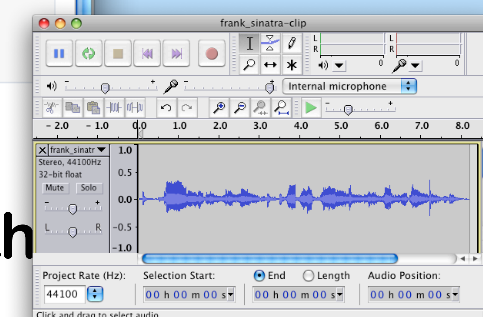
More Consumers

- 110 Million Apple iPods sold worldwide
  - 40,000 Songs on a 160 GB handheld device
- 7 Million Users on Pandora
- 700K daily Facebook iLike users

More Producers

- 12 Millon Songs indexed by AMG All Music
- 100,000 Artist have uploaded free MP3s to LastFM
- 1 million downloads per month of Audacity
  - Free Music Editing Software





How do we connect m… …cers with

How do we find music?

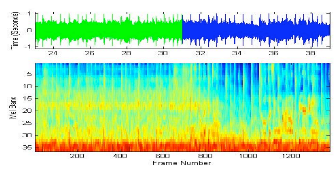
- **Query-by-Metadata - artist, song, album, year**
  - We must know what we want
- **Query-by-(Humming, Tapping, Beatboxing)**
  - Requires talent
- **Query-by-Song-Similarity**
  - Collaborative Filtering, Acoustic Similarity
  - Lacks interpretablilty
- **Query-by-Semantic-Description**
  - **Google** seems to work pretty well for **text**
  - **Semantic Image Labeling** is a hot topic in **Computer Vision**
  - Can it work for **music?**

Semantic Music Annotation and Retrieval

Our goal is build a system that can

1. **Annotate** a song with meaningful tags
2. **Retrieve** songs given a text-based query

Frank Sinatra
'Fly Me to the Moon'



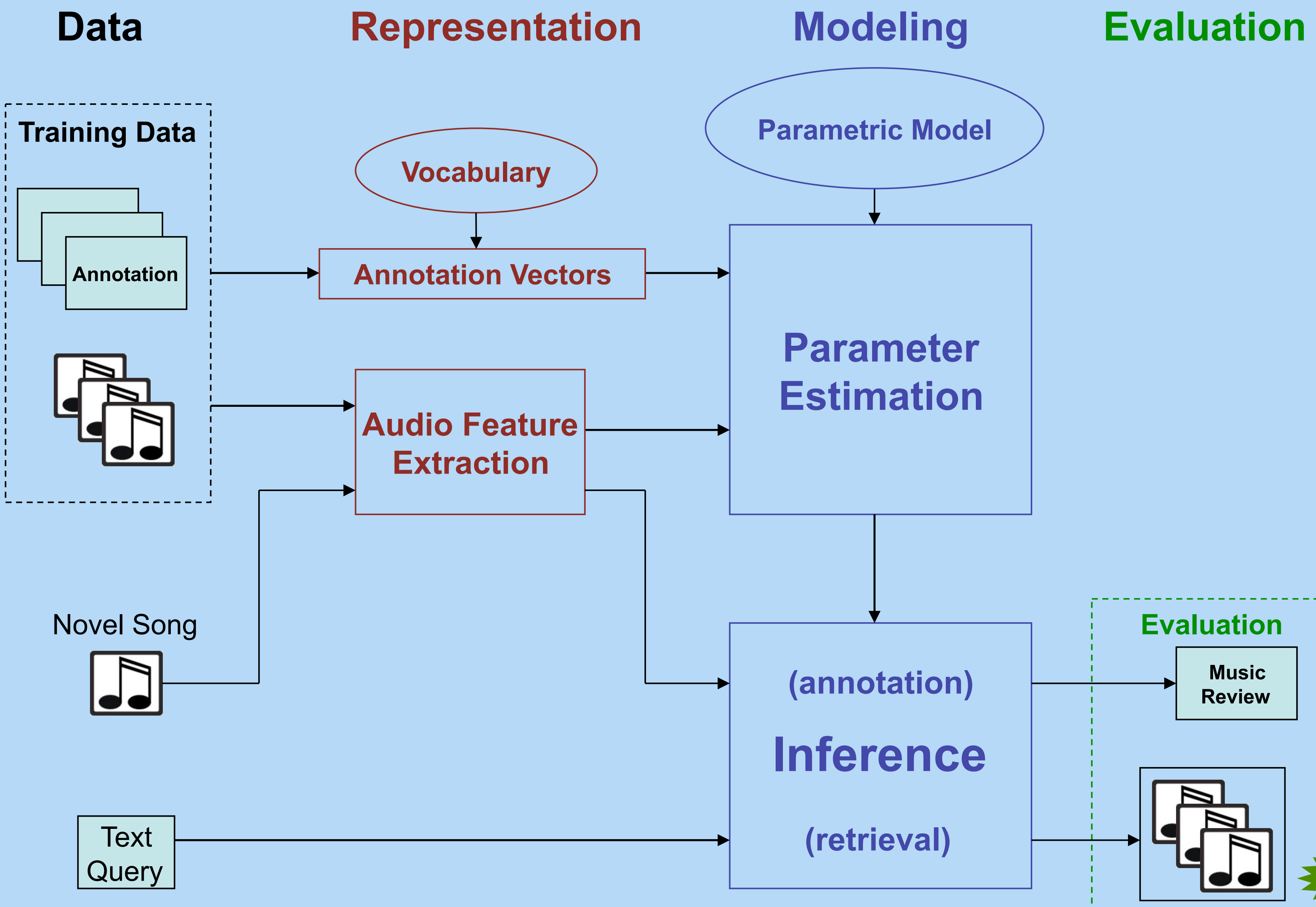
Annotation →

← **Retrieval**

'Jazz'
'Male Vocals'
'Sad'
'Mellow'
'Slow Tempo'

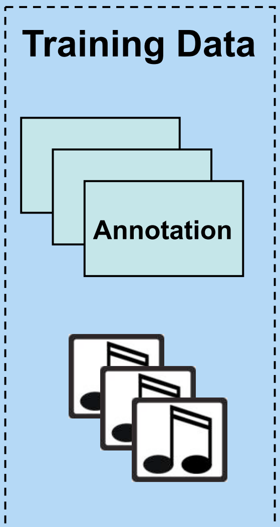
Plan: Learn a probabilistic model that captures a relationship between **audio content** and **tags**.

System Overview



System Overview

Data



Collecting an Annotated Music Corpus

1. Text-mining web documents
   - 2,100 song reviews from AMG All Music
   - Extracted a vocab of 317 words

Strawberry Fields Forever

The Beatles

Send to Friend

Composed By	Other Links
John Lennon/Paul McCartney	All Performers that have performed this Title

Song Review

by Richie Unterberger

A hit single (as part of a double A-side with "Penny Lane") in early 1967, "Strawberry Fields Forever" is one of the Beatles' peak achievements and one of the finest Lennon- McCartney songs. At the beginning of its quite complicated evolution, the tune, principally the work of John Lennon, was a folky ballad evoking a dream world where nothing was real and there was nothing to get hung about. As is well known, Strawberry Fields is a real location in the city where Lennon grew up, Liverpool. Strawberry Field (the actual name is singular) was a Salvation Army orphanage where he used to play with his friends as a child. The song "Strawberry Fields Forever," however, is not so much about a physical place as it is about a state of mind, drug-influenced almost surely. As is the case with several other Lennon songs of the period, there could be an implication that this attractively lethargic, peaceful state is an inner state of being preferable to that of the straight world. "Strawberry Fields" is not necessarily a utopia, though, as the references to living being easy with eyes closed convey; it could be interpreted as an unhealthy escapist withdrawal from reality or even an inference that death is preferable to life. The song's effect was greatly enhanced by its production, with an arrangement that would undergo numerous changes during the course of the recording process. Two finished versions were completed, one closer to the song's folky origins, one more orchestrated. John Lennon wanted to use portions from each, and asked George Martin to combine them, despite the fact that they had been recorded in different keys and tempos. In his most famous production feat of all, Martin slowed one portion down and sped up the other, finding that -- miraculously -- both the tempos and the keys then matched. That accounts in large part for the magnificently varied moods and textures of the final track, with its unearthly opening mellotron introduction (mistaken by some for flutes) leading into the gently delivered initial chorus and verse. The atmosphere becomes tenser and effectively underscores the ambiguity of the song's lyrics, as funereal marching brass is introduced, making the promise of Strawberry Fields' dream world seem more ominous, even menacing (especially when a noise like that of a spade digging a grave is hard). The song comes to a glorious close with several repetitions of the title chorus and a fadeout with lovely guitar, cello, and swordmandel (played by George Harrison). But wait -- that isn't quite the end. After a few seconds of dead air, a totally unrelated and quite dissonant piece of musique concrète fades in, with crunchy Ringo Starr drum rolls, nightmarish flute mellotrons, and a noise that sounds like muted ambulance horns. This coda again could signify that the placid dreamscape of Strawberry Fields isn't all it seems, and that a hellish whirlwind could be lurking just underneath. At the

Collecting an Annotated Music Corpus

1. Text-mining web documents
   - ✓ Cheap, tons of data
   - ✗ Noisy, opinionated, unnatural disconnect

Collecting an Annotated Music Corpus

1. Text-mining web documents
2. Conducting a survey
   - 174-tag hierarchical vocab - genre, emotion, usage, …
   - Paid 55 undergrads to annotate music for 120 hours
   - CAL500: 500 songs annotated by a minimum of 3 people

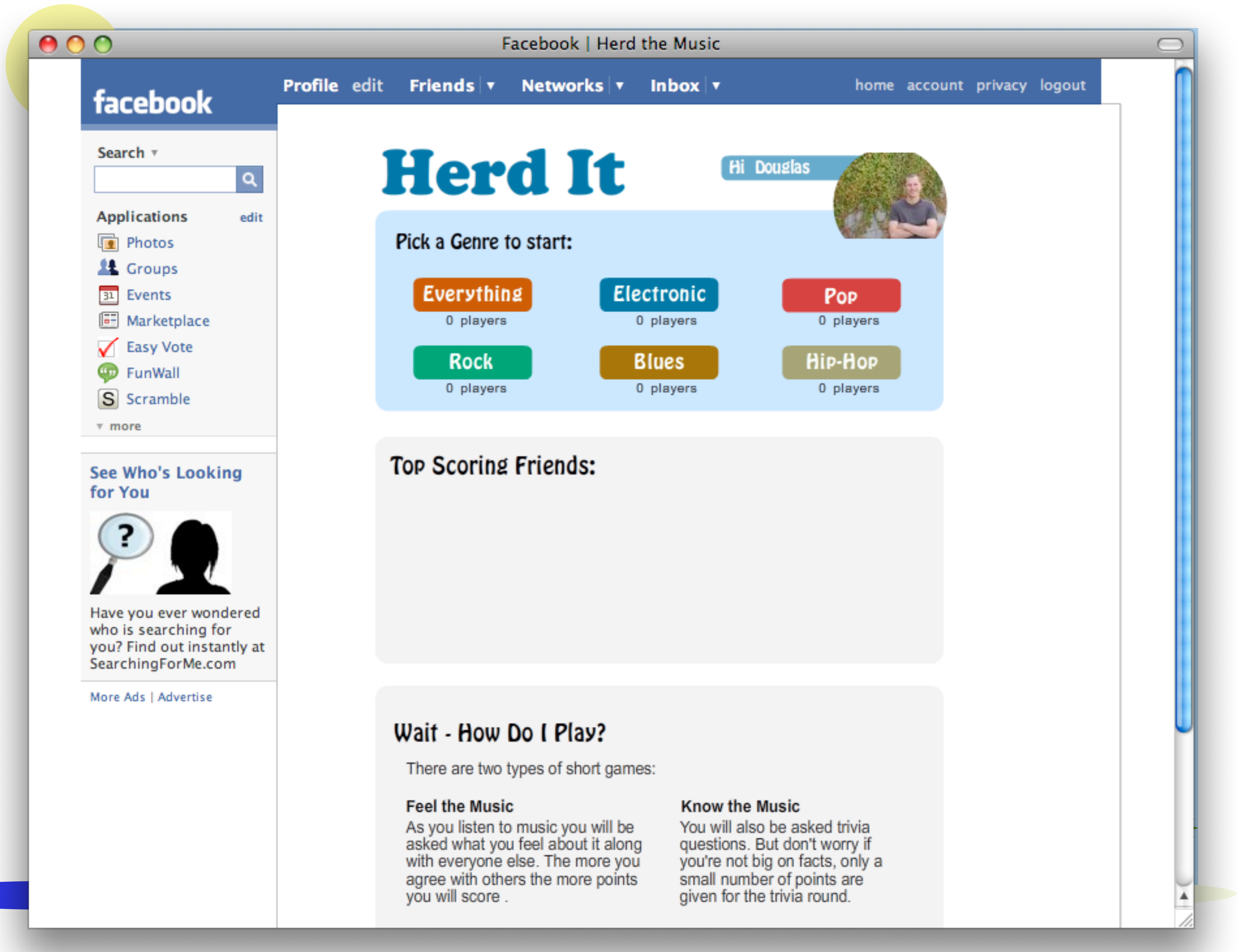
INSTRUMENTATION												
Which instruments are present, are prominent, or are featured in a solo.												
Instrument	None	Uncertain	Present	Prominent	Solo		Instrument	None	Uncertain	Present	Prominent	Solo
Voice							- String Ensemble	○	○	○	○	☐
- Male Lead Vocals	○	○	○	○	☐		- Orchestra	○	○	○	○	☐
- Female Lead Vocals	○	○	○	○	☐		Wind Instruments					
- Backing vocals	○	○	○	○	☐		- Harmonica	○	○	○	○	☐
- Choir	○	○	○	○	☐		- Trumpet	○	○	○	○	☐
Guitar Family							- Trombone	○	○	○	○	☐
- Acoustic Guitar	○	○	○	○	☐		- Saxophone	○	○	○	○	☐
- Electric Guitar (clean)	○	○	○	○	☐		- Horn Section	○	○	○	○	☐
- Electric Guitar (distorted)	○	○	○	○	☐		Electronics					
- Slide Guitar	○	○	○	○	☐		- Samples	○	○	○	○	☐
- Bass	○	○	○	○	☐		- Ambient Sounds	○	○	○	○	☐
- Banjo	○	○	○	○	☐		- Scratches	○	○	○	○	☐
Keyboards							- Sequencer	○	○	○	○	☐
- Piano	○	○	○	○	☐		Percussion					
- Organ	○	○	○	○	☐		- Drum Set	○	○	○	○	☐
- Synthesizer	○	○	○	○	☐		- Drum Machine	○	○	○	○	☐
String Instruments							- Hand Drums	○	○	○	○	☐
- Violin/Fiddle	○	○	○	○	☐		- Tambourine	○	○	○	○	☐
Instrument	None	Uncertain	Present	Prominent	Solo		Instrument	None	Uncertain	Present	Prominent	Solo

Collecting an Annotated Music Corpus

1. Text-mining web documents
2. Conducting a survey
   - ✓ Reliable, Precise, Tailored to Application
   - X Expensive, Laborious, Not Scalable

Collecting an Annotated Music Corpus

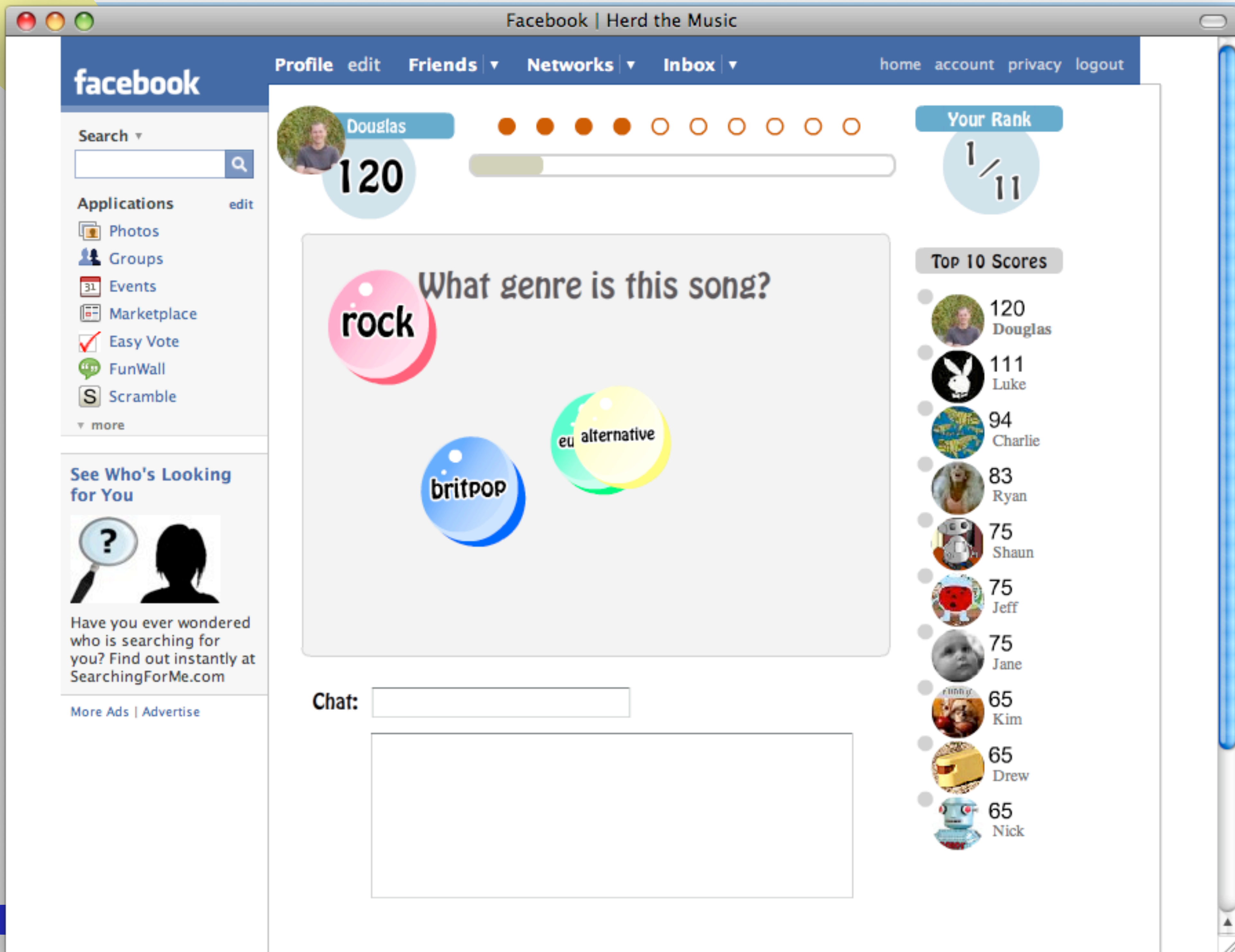
1. Text-mining web documents
2. Conducting a survey
3. Deploying a ‘Human-Computation’ game
    - Web-based, multi-player game with real-time interaction
    - ESPGame by Luis Von Ahn
    - Listen Game [ISMIR 07]


Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Herd It
Hi Douglas
Pick a Genre to start:
Everything
0 players
Electronic
0 players
Pop
0 players
Rock
0 players
Blues
0 players
Hip-Hop
0 players
Top Scoring Friends:
Wait - How Do I Play?
There are two types of short games:
Feel the Music
As you listen to music you will be asked what you feel about it along with everyone else. The more you agree with others the more points you will score .
Know the Music
You will also be asked trivia questions. But don't worry if you're not big on facts, only a small number of points are given for the trivia round.

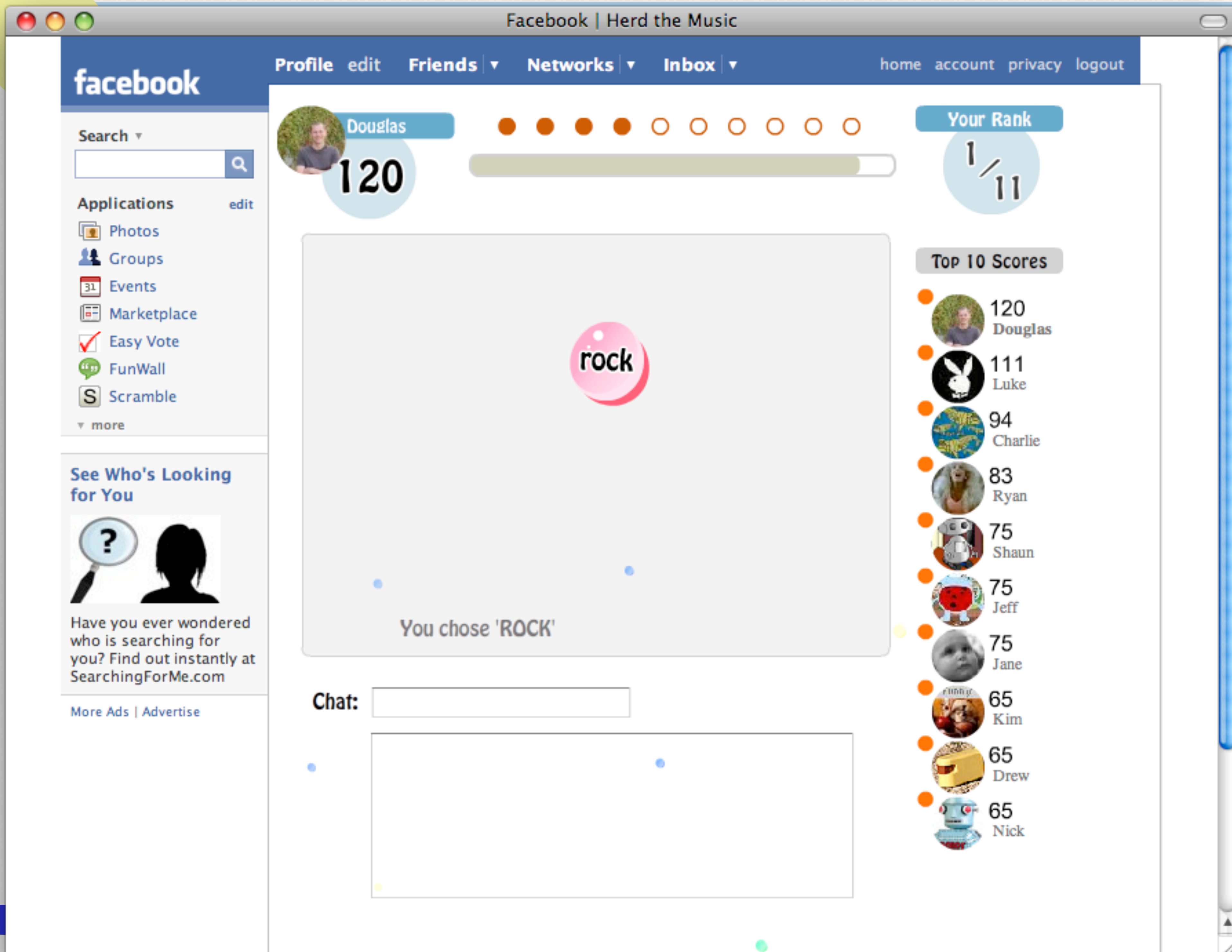

Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Douglas
0
Your Rank
11/11
ENERGETIC
Put this song on the Emoti-O-Meter
SAD
HAPPY
CALM
Chat:
Top 10 Scores
30 Scott
30 Charlie
20 Jeff
20 Jane
20 Shaun
20 Luke
10 Drew
10 Kim
10 Ryan
10 Nick

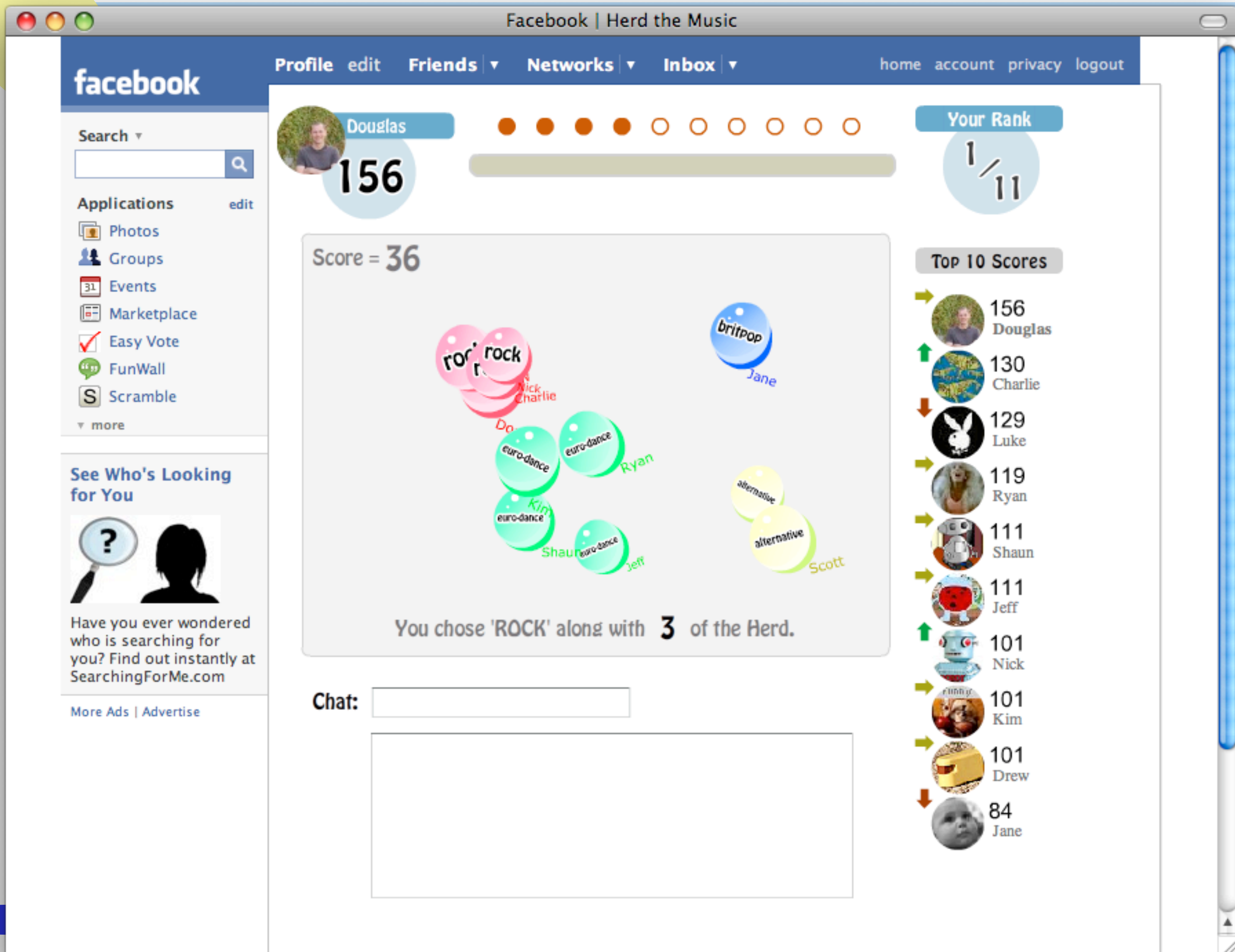

Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Douglas
0
Your Rank
11/11
ENERGETIC
SAD
HAPPY
CALM
Chat:
Top 10 Scores
30 Scott
30 Charlie
20 Jeff
20 Jane
20 Shaun
20 Luke
10 Drew
10 Kim
10 Ryan
10 Nick


Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Douglas
100
Your Rank
2/11
Score = 89
ENERGETIC
SAD
HAPPY
CALM
Ryan
Luke
Scott
Jane
Drew
Kim
8 of the Herd are near you
Chat:
Top 10 Scores
111 Luke
100 Douglas
94 Charlie
83 Ryan
75 Shaun
75 Jane
75 Jeff
65 Kim
65 Drew
65 Nick


Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Douglas
120
What genre is this song?
rock
britpop
alternative
Chat:
Your Rank
1/11
Top 10 Scores
120 Douglas
111 Luke
94 Charlie
83 Ryan
75 Shaun
75 Jeff
75 Jane
65 Kim
65 Drew
65 Nick


Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Douglas
120
britpop
rock
alternative
euro-dance
Chat:
Your Rank
1/11
Top 10 Scores
120 Douglas
111 Luke
94 Charlie
83 Ryan
75 Shaun
75 Jeff
75 Jane
65 Kim
65 Drew
65 Nick


Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Douglas
120
Your Rank
1/11
rock
You chose 'ROCK'
Chat:
Top 10 Scores
120 Douglas
111 Luke
94 Charlie
83 Ryan
75 Shaun
75 Jeff
75 Jane
65 Kim
65 Drew
65 Nick



Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Douglas
156
Your Rank
1/11
Score = 36
rock
britpop
euro-dance
alternative
Jane
Ryan
Scott
You chose 'ROCK' along with 3 of the Herd.
Chat:
Top 10 Scores
156 Douglas
130 Charlie
129 Luke
119 Ryan
111 Shaun
111 Jeff
101 Nick
101 Kim
101 Drew
84 Jane


Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Douglas
156
Your Rank
1/11
Trivia Round
What SONG have you been listening to?
Somewhere Only We Know
The Scientist
Stop Crying Your Heart Out
Creep
Chat:
Top 10 Scores
156 Douglas
130 Charlie
129 Luke
119 Ryan
111 Shaun
111 Jeff
101 Nick
101 Kim
101 Drew
84 Jane

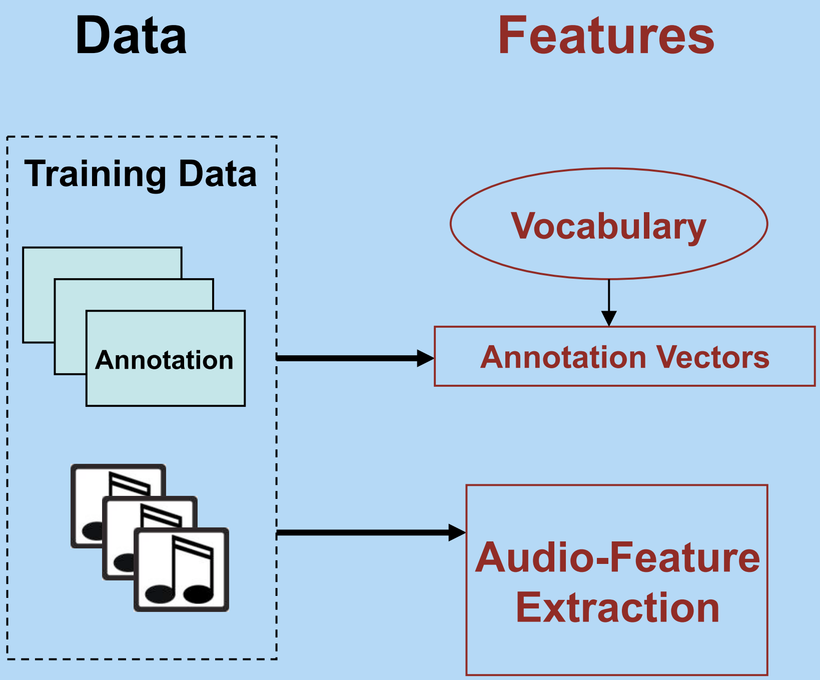
Facebook | Herd the Music
facebook
Profile edit Friends Networks Inbox
home account privacy logout
Search
Applications edit
Photos
Groups
Events
Marketplace
Easy Vote
FunWall
Scramble
more
See Who's Looking for You
Have you ever wondered who is searching for you? Find out instantly at SearchingForMe.com
More Ads | Advertise
Douglas
176
Your Rank
1/11
Trivia Round
What SONG have you been listening to?
Somewhere Only We Know
The Scientist
+20
Stop Crying Your Heart Out
Creep
Correct! You just heard:
Stop Crying Your Heart Out
by
Oasis
Chat:
Top 10 Scores
176 Douglas
130 Charlie
129 Luke
119 Ryan
111 Shaun
111 Jeff
101 Nick
101 Kim
101 Drew
84 Jane

Collecting an Annotated Music Corpus

We have explored three techniques

1. Text-mining web documents
2. Conducting a survey
3. Deploying a 'Human-Computation' game
   - ✓ Cheap, Scalable, Precise, Personalized
   - X Need to create a viral user experience

System Overview



Semantic Representation: y

Choose vocabulary of 'musically relevant' tags

- Instruments, Genre, Emotion, Rhythm, Energy, Vocal, Usages

Each annotation is converted to a real-valued vector

- 'Semantic association' between a tag and the song.

Example: Frank Sinatra's "Fly Me to the Moon"

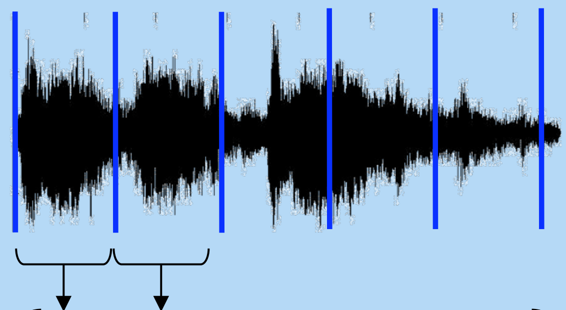
Vocab = {funk, jazz, guitar, female vocals, sad, passionate}

y = [0/4 , 3/4, 4/4 , 0/4 , 2/4, 1/4]

Acoustic Representation: X

Each song is represented as a bag-of-feature-vectors

- Pass a short time window over the audio signal
- Extract a feature vector for each short-time audio segment
- Ignore temporal relationships of time series



$$\mathbf{X} = \{ \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_t \}$$

Audio Features

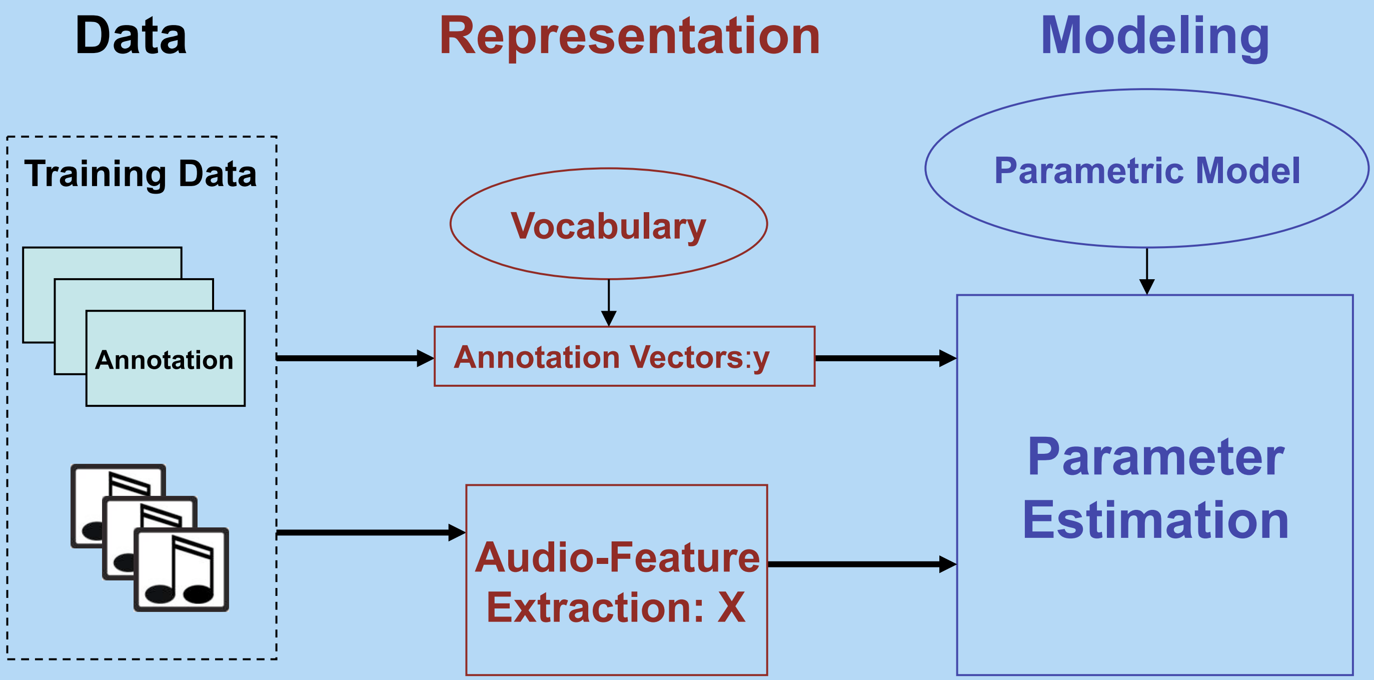
We calculate MFCC+Deltas feature vectors

- Mel-frequency Cepstral Coefficients (MFCC)
  - Low dimensional representation short-term spectrum
  - Popular for both representing speech, music, and sound effects
- Instantaneous derivatives (deltas) encode short-time temporal info
- 5,200 39-dimensional vectors per minute

Numerous other audio representations

- Spectral features, modulation spectra, chromagrams, …

System Overview



Statistical Model

Supervised Multi-class Labeling model

- Set of probability distributions over the audio feature space
- One Gaussian Mixture Model (GMM) per tag - $p(x|t)$
- **Key Idea:** Estimate parameters for GMM using the set of training songs that are positively associated with the tag

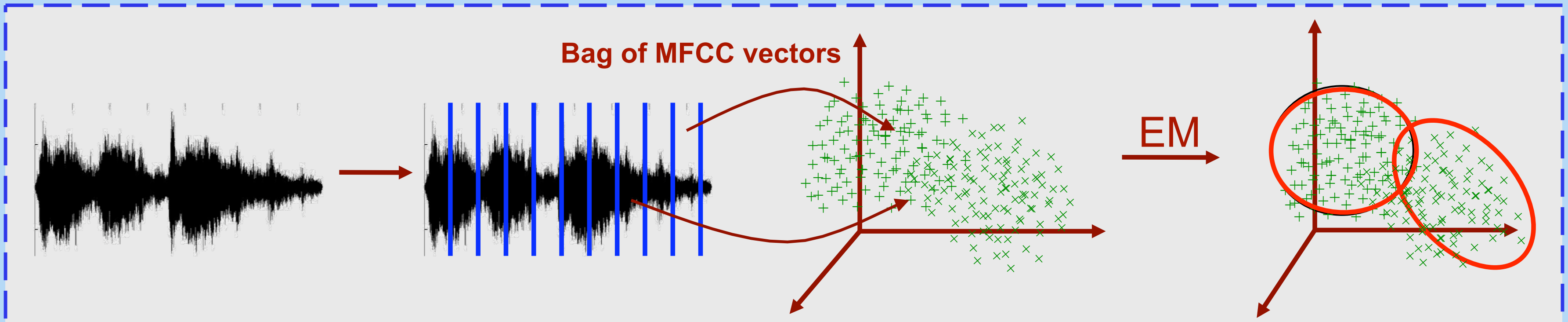
Notes:

- Developed for image annotation
- Scalable and Parallelizable
- Modified for real-value semantic weights rather than binary class labels
- Extended formulation to handle multi-tag queries

Modeling a Song

Algorithm

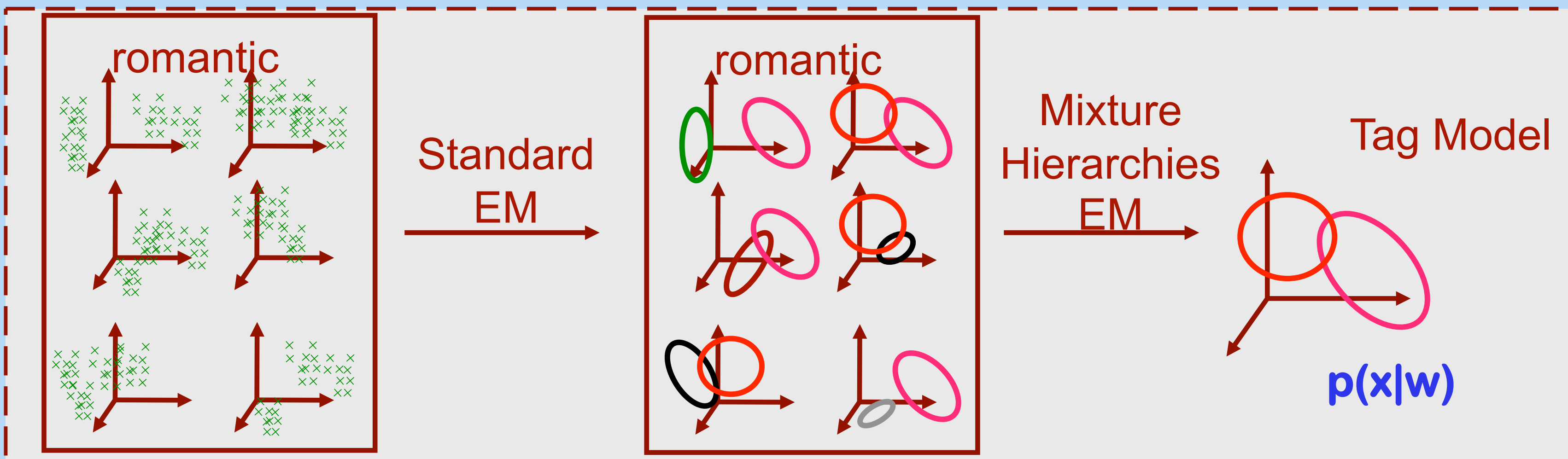
1. Segment audio signals
2. Extract short-time feature vectors
3. Estimate GMM
   - expectation maximization algorithm



Modeling a Tag

Algorithm:

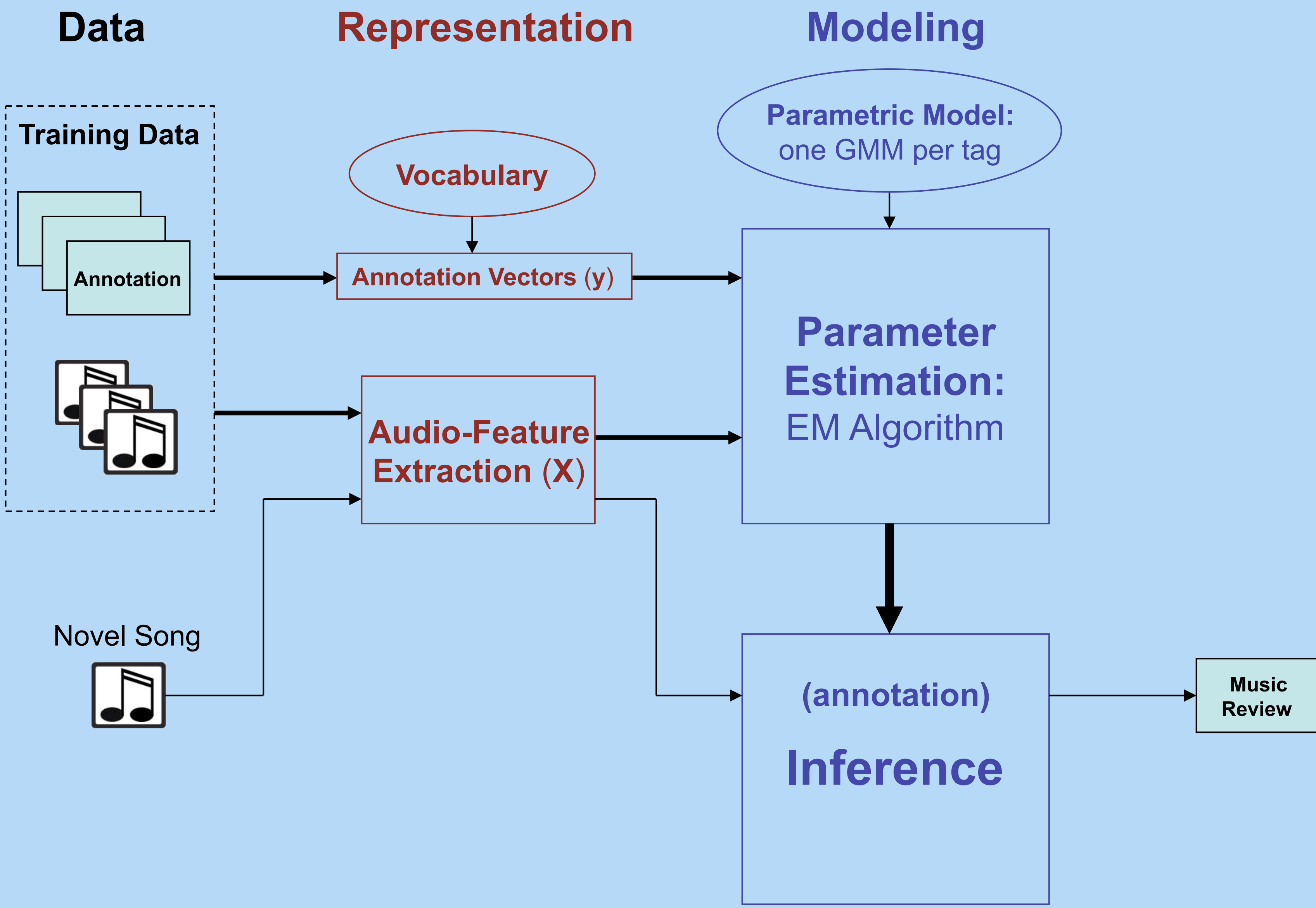
1. Identify songs associated with tag t
2. Estimate a 'song GMM' for each song - p(x|s)
3. Use the Mixture Hierarchies EM algorithm [Vasconcelos01]
   - Learn a 'mixture of mixture components'



Benefits

+ Computationally efficient for parameter estimation and inference
+ 'Smoothed' song representation → better density estimate

System Overview



Annotation

Given a novel song $X = \{x_1, \ldots, x_T\}$, calculate the probability of each tag given the song:

$$P(t|X) = \frac{P(X|t)P(t)}{P(X)}$$

Assuming

1. Uniform word prior P(t)
2. Vectors are conditionally independent given a tag
3. Geometric average of likelihoods

$$P(t|X) = \frac{\left(\prod_{i=1}^{T} P(\mathbf{x}_i|t)\right)^{\frac{1}{T}}}{\sum_{v \in V}\left(\prod_{i=1}^{T} P(\mathbf{x}_i|v)\right)^{\frac{1}{T}}}$$

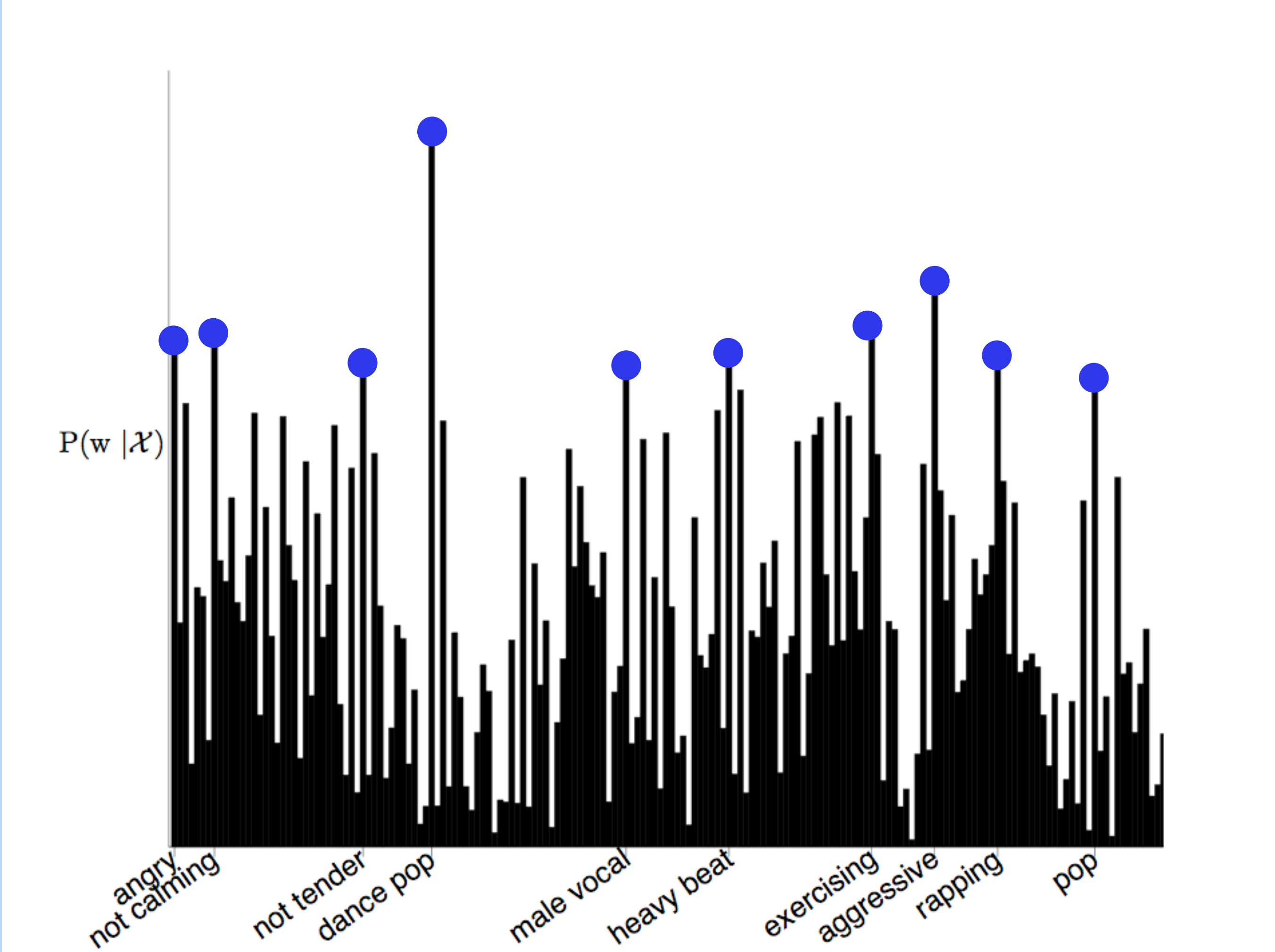
Semantic Multinomial:

- Conditional probabilities, P(t|X), defines multinomial over the vocabulary

Annotation: pick peaks of the semantic multinomial

Annotation

Semantic Multinomial for "Give it Away" by the Red Hot Chili Peppers



Annotation: Automatic Music Reviews

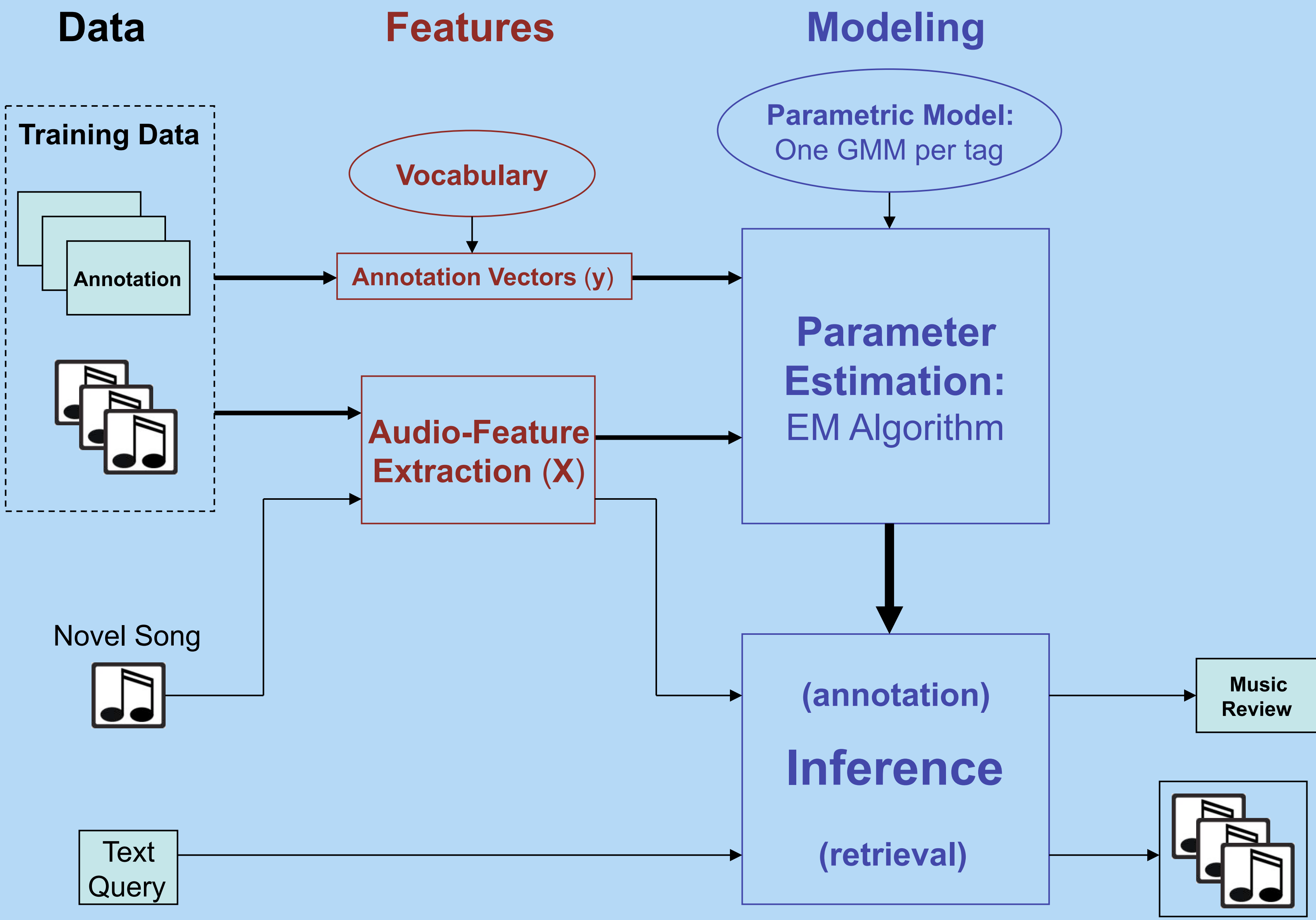
Dr. Dre (feat. Snoop Dogg) - Nuthin' but a 'G' thang

This is a dance poppy, hip-hop song that is arousing and exciting. It features drum machine, backing vocals, male vocal, a nice acoustic guitar solo, and rapping, strong vocals. It is a song that is very danceable and with a heavy beat that you might like listen to while at a party.

Frank Sinatra - Fly me to the moon

This is a jazzy, singer / songwriter song that is calming and sad. It features acoustic guitar, piano, saxophone, a nice male vocal solo, and emotional, high-pitched vocals. It is a song with a light beat and a slow tempo that you might like listen to while hanging with friends.

System Overview



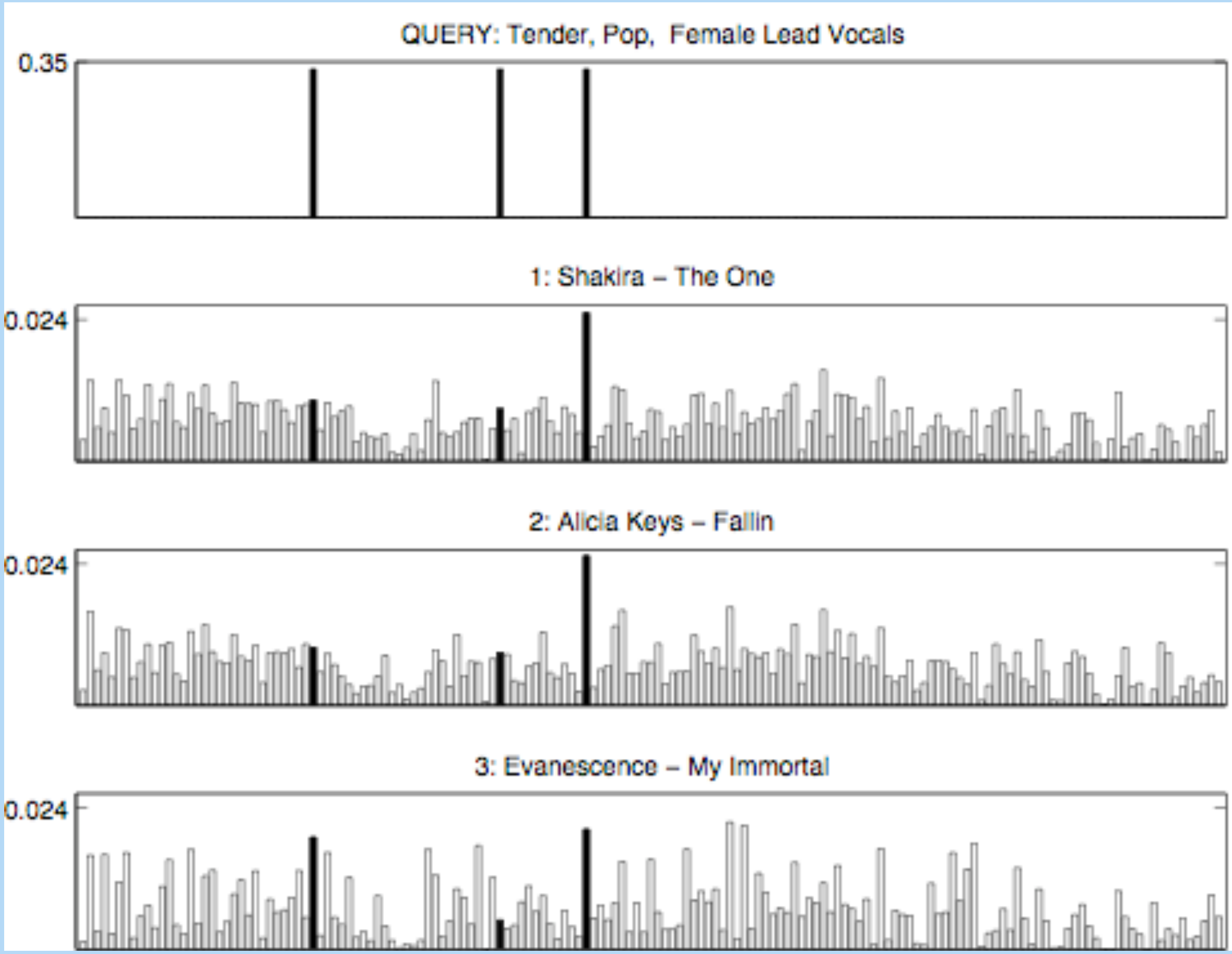
Retrieval

1. Annotate each song in corpus with a **semantic multinomial p**
   - $\mathbf{p} = \{P(t_1|X), \ldots, P(t_{|V|}|X)\}$
2. Given a text-based query, construct a **query multinomial q**
   - $q_i = 1/|t|$, if tag t appears in the query string
   - $q_i = 0$, otherwise
3. Rank all songs by the **Kullback-Leibler (KL) divergence**

$$KL(\mathbf{q}||\mathbf{p}) = \sum_{i=1}^{|\mathcal{V}|} q_i \log \frac{q_i}{p_i}$$

Retrieval

The top 3 semantic multinomials for the query "'pop', 'female lead vocals', 'tender'"



Retrieval: Query-by-Semantic-Description

Query	Retrieved Songs
'Tender'	**Crosby, Stills and Nash** - Guinevere **Jewel** - Enter from the East **Art Tatum** - Willow Weep for Me **John Lennon** - Imagine **Tom Waits** - Time
'Female Vocals'	**Alicia Keys** - Fallin' **Shakira** - The One **Christina Aguilera** - Genie in a Bottle **Junior Murvin** - Police and Thieves **Britney Spears** - I'm a Slave 4 U
'Tender' AND 'Female Vocals'	**Jewel** - Enter from the East **Evanescence** - My Immortal **Cowboy Junkies** - Postcard Blues **Everly Brothers** - Take a Message to Mary **Sheryl Crow** - I Shall Believe

Digression: Music Similarity

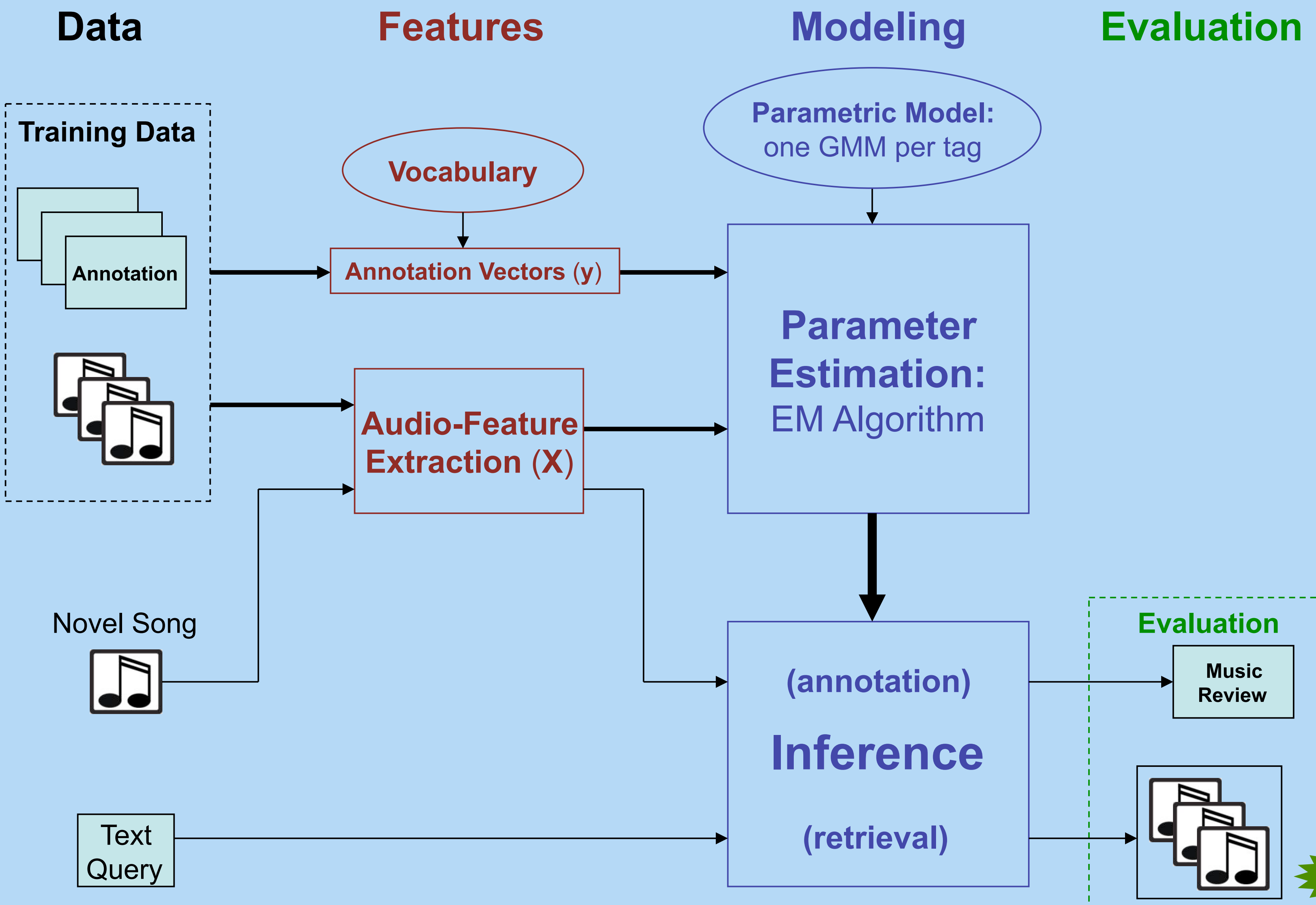
Query-by-semantic-similarity [ICASSP 07]

- KL divergence between 2 semantic multinomials
- 3rd Place in 2007 MIREX Similarity Task
  - No statistical difference between top 4 teams

Advantages:

1. Semantically Interpretable Comparisons
   - What makes two songs similar?
2. Heterogeneous queries
   - “Find me ‘sad’ songs that are like ‘Hey Jude’ ”

System Overview



Quantifying Annotation

Our system annotates the Cal-500 songs with 10 tags from our 174-tag vocabulary.

- 'Consensus Annotation' Ground Truth

Metric: 'Tag' Precision & Recall

$$\text{Precision} = \frac{\#\ \text{songs correctly annotated with}\ t}{\#\ \text{songs annotated with}\ t}$$

$$\text{Recall} = \frac{\#\ \text{songs correctly annotated with}\ t}{\#\ \text{songs that should have been annotated}\ t}$$

Mean Tag Recall and Tag Precision are the averages over all tags in our vocabulary.

Quantifying Annotation

Our system annotates the Cal-500 songs with 10 tags from our 174-tag vocabulary.

Method	Precision	Recall
Random	0.14	0.06
Upper Bound	0.71	0.38
Our System	0.27	0.16
Human	0.30	0.15

Compared with a human, our model is

- worse on objective categories - instrumentation, genre
- about the same on subjective categories - emotion, usage

Quantifying Retrieval

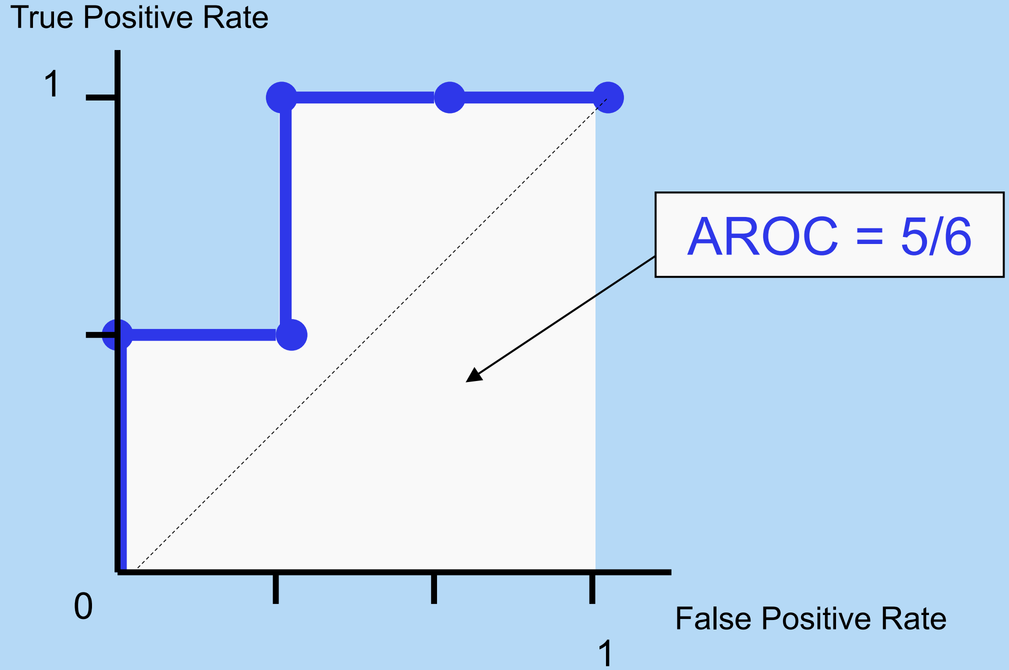
Rank order test set songs

- KL between a query multinomial and semantic multinomials
- 1-, 2-, 3-word queries with 5 or more examples

Metric: Area under the ROC Curve (AROC)

Rank by 'Romantic'

Rank	Label	TP	FP
1	R	1/2	0
2	-	1/2	1/3
3	R	1	1/3
4	-	1	2/3
5	-	1	1



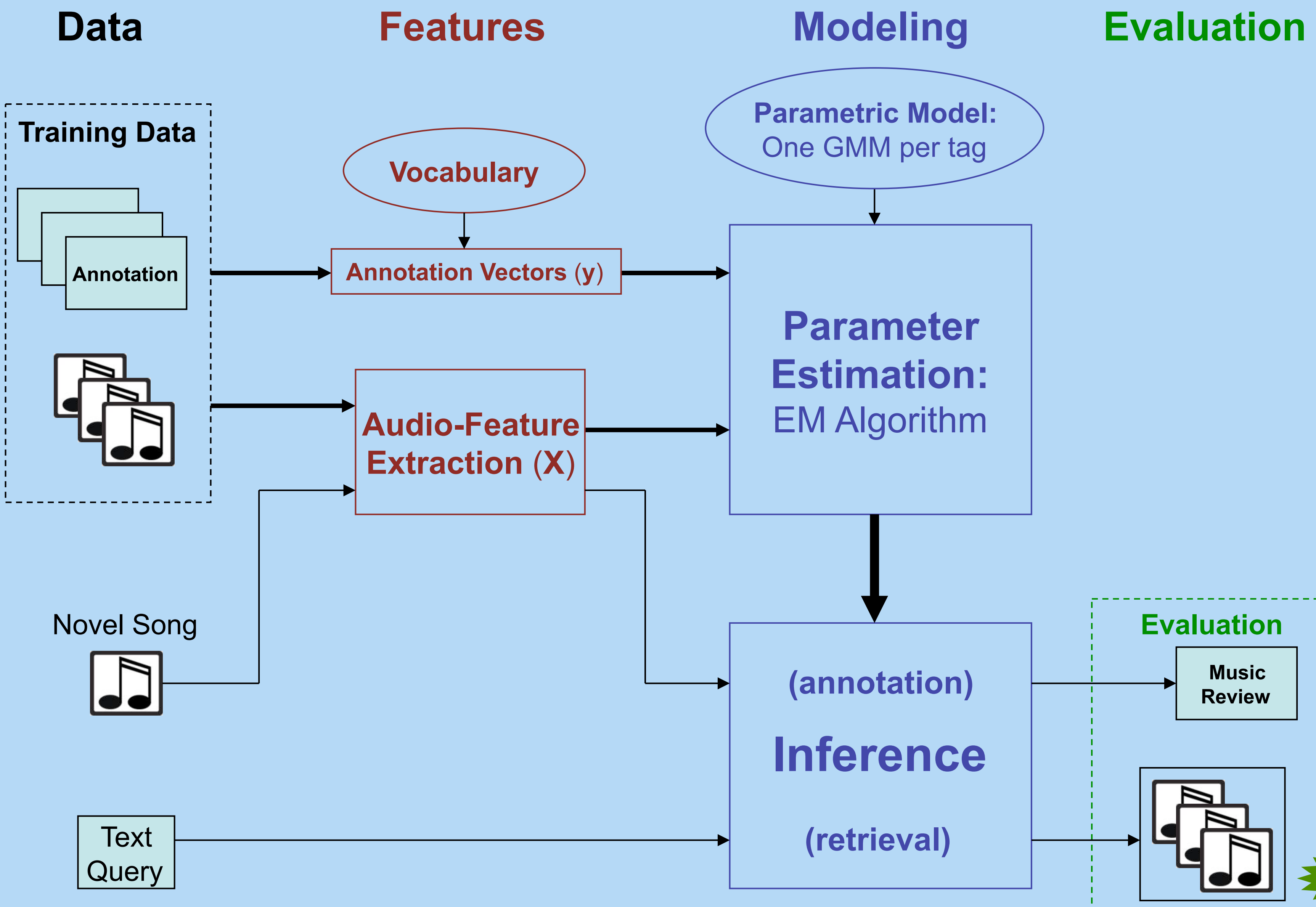
Mean AROC is the average AROC over a large number of queries.

Quantifying Retrieval

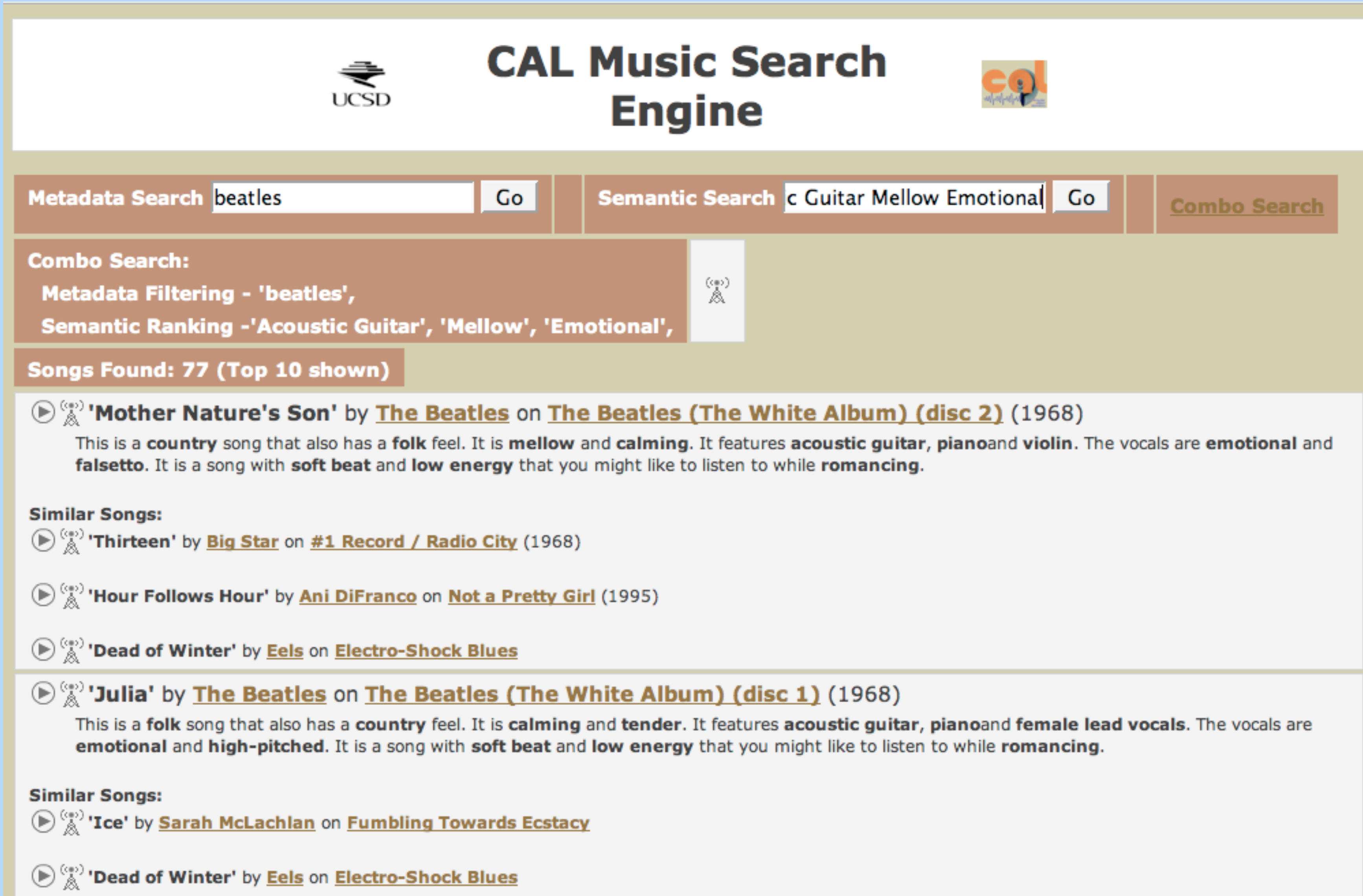
We rank order song according to songs once for each query.

Model	AROC
Random	0.50
Upper Bound	1.00
Our System - 1 Tag	0.71
Our System - 2 Tags	0.72
Our System - 3 Tags	0.73

System Overview



CAL Music Search Engine



CAL Music Search Engine

Metadata Search beatles Go

Semantic Search c Guitar Mellow Emotional Go

Combo Search

Combo Search:
Metadata Filtering - 'beatles',
Semantic Ranking -'Acoustic Guitar', 'Mellow', 'Emotional',

Songs Found: 77 (Top 10 shown)

'Mother Nature's Son' by The Beatles on The Beatles (The White Album) (disc 2) (1968)

This is a **country** song that also has a **folk** feel. It is **mellow** and **calming**. It features **acoustic guitar**, **piano**and **violin**. The vocals are **emotional** and **falsetto**. It is a song with **soft beat** and **low energy** that you might like to listen to while **romancing**.

Similar Songs:

'Thirteen' by Big Star on #1 Record / Radio City (1968)

'Hour Follows Hour' by Ani DiFranco on Not a Pretty Girl (1995)

'Dead of Winter' by Eels on Electro-Shock Blues

'Julia' by The Beatles on The Beatles (The White Album) (disc 1) (1968)

This is a **folk** song that also has a **country** feel. It is **calming** and **tender**. It features **acoustic guitar**, **piano**and **female lead vocals**. The vocals are **emotional** and **high-pitched**. It is a song with **soft beat** and **low energy** that you might like to listen to while **romancing**.

Similar Songs:

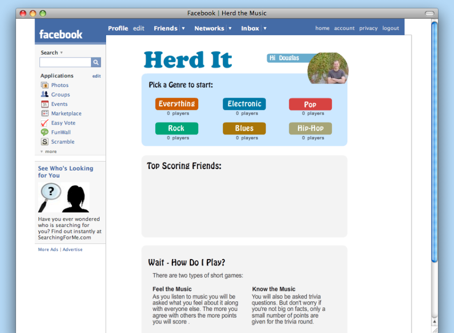
'Ice' by Sarah McLachlan on Fumbling Towards Ecstacy

'Dead of Winter' by Eels on Electro-Shock Blues

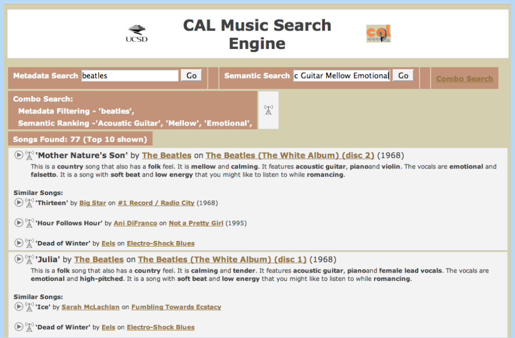
The BIG picture

Music Fans

GAMES



search influences game design



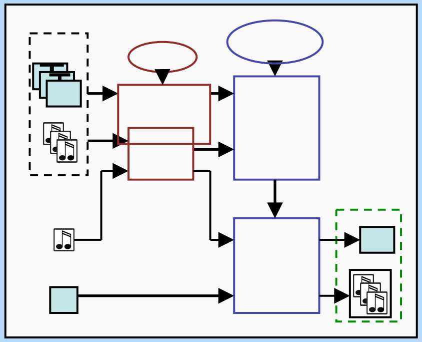
SEARCH & DISCOVERY

annotation data

annotations power search

DATABASE

CA system learns to annotate new songs



COMPUTER AUDITION

What's on tap…

Research Challenges

1. Explore song similarity
   - Query-by-semantic-example - ICASSP 07, MIREX 07
2. Model correlation between tags
3. Explore discriminative approaches
4. Combine heterogeneous data sources
   - Game Data, Semantic Tags, Web Documents, Popularity Info
5. Focus on individuals / groups rather than population
   - Emotional state of listener

"Talking about music is like dancing about architecture"

- origins unknown

Douglas Turnbull

Computer Audition Lab

UC San Diego

dturnbul@cs.ucsd.edu

cs.ucsd.edu/~dturnbul

References

Semantic Annotation and Retrieval [SIGIR 07, IEEE TASLP 08]

Music Annotation Games [ISMIR 07]

Query-by-Semantic-Similarity [ICASSP 07, MIREX 07]

Tag Vocabulary Selection [ISMIR 07]

- Sparse Canonical Correlation Analysis

Work-in-Progress:

1. (More) Social Music Annotation Games
2. Combining Tags from Multiple Sources
3. Music Similarity with Semantics

What's up next…

Building 'Commercial Grade' system

1. Collecting data
   - 'Legally' collecting music
   - Herd It Game - [ISMIR 07]
2. Vocabulary expansion
   - LastFM - 25,000 tags
     - Vocab selection using Sparse CCA - [ISMIR 07]
   - Web Documents - All words
3. User interface design
   - Natural language music search engine
   - Customizable radio player
4. Automated 'Large Scale' System

Gaussian Mixture Model (GMM)

A GMM is used to model probability distributions over high dimensional spaces:

$$P(\mathbf{x}|w) = \sum_{r=1}^{R} \pi_r \mathcal{N}(\mathbf{x}|\mu_r, \Sigma_r)$$

A GMM is a weighted combo of R Gaussian distributions

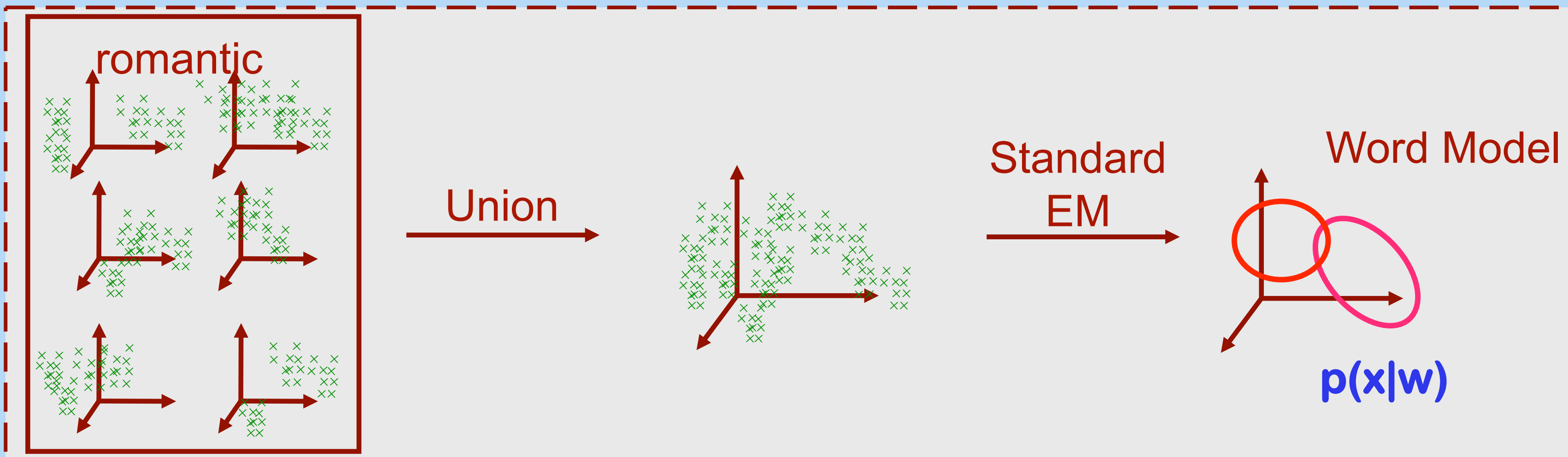
- π_r is the r-th mixing weight
- μ_r is the r-th mean
- Σ_r is the r-th covariance matrix

These parameters are usually estimated using a ‘standard’ **Expectation Maximization** (EM) algorithm.

Three approaches for estimating p(x|w)

1. Direct Estimation

1. Identify songs associated with w
2. Union of feature vectors for these songs
3. Estimate GMM using ‘standard’ EM

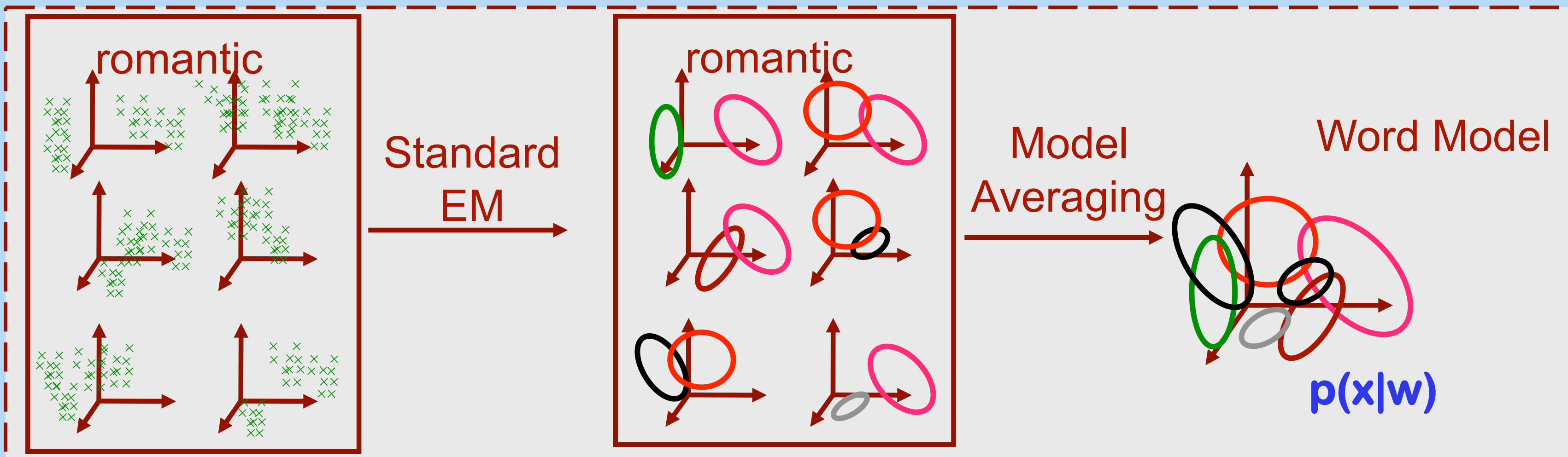


Problem: Direct Estimation is computationally difficult and empirically converges to poor local optima.

Three approaches for estimating p(x|w)

2. Model Averaging Estimation

1. Identify songs associated with w
2. Estimate a 'song GMM' for each song - p(x|s)
3. Use all mixture components from 'song GMMs'



Problem: As the training set size grows, evaluating this distribution becomes prohibitively expensive.

A biased view of Music Classification

2000-03: **Music classification** (by genre, emotion, instrumentation) becomes a popular MIR task

- Undergrad Thesis on Genre Classification with G. Tzanetakis

2003-04: MIR community starts to **criticize** music classification problems

- ill-posed problem due to subjectivity
- not an end in itself
- performance ‘glass ceiling’

2004-06: Focus turns to **Music Similarity** research

- Recommendation
- Playlist generation

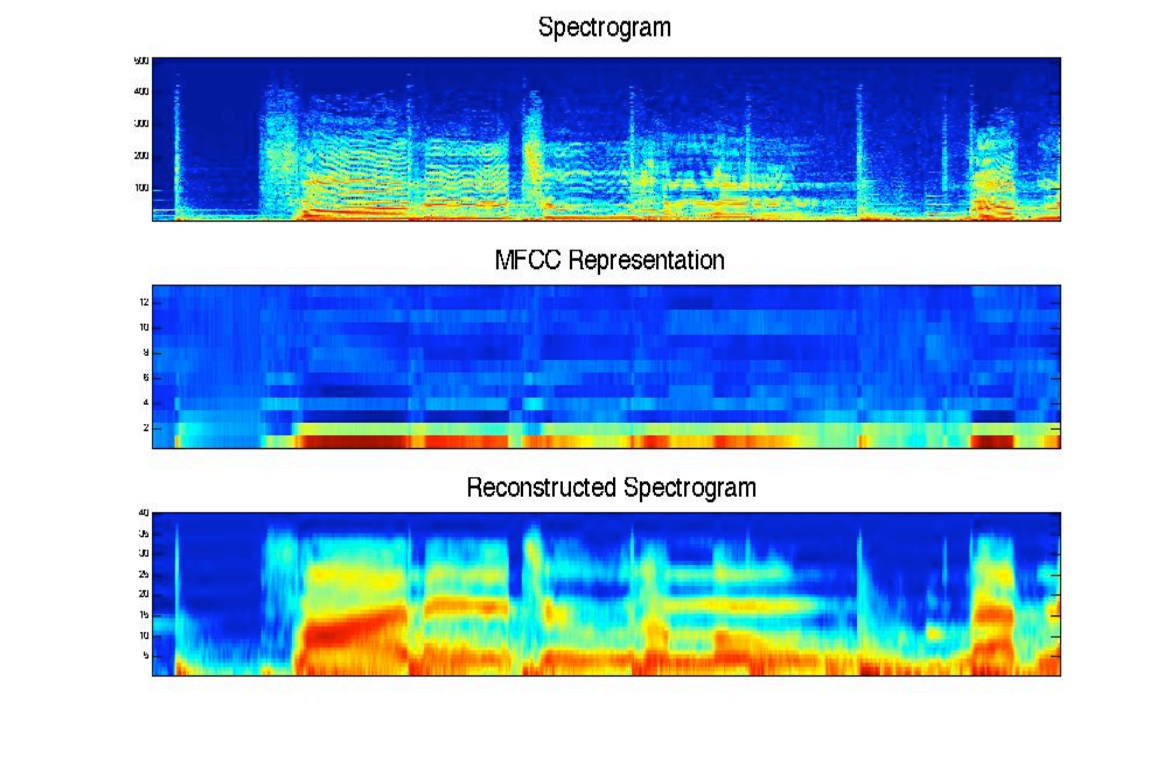
2006-07: We view **Music Annotation** as a supervised multi-class labeling problem

- Like classification but with large, less-restrictive vocabulary

Acoustic Representation

Calculating Delta MFCC feature vectors

- Calculate a time-series for 13 MFCCs
- Append 1st and 2nd instantaneous derivatives
- 5,200 39-dimensional feature vectors per minute of audio content
- Denoted by $X = \{x_1, \ldots, x_T\}$ where T depends on the length of the song



Short-Time Fourier Transform

Time Series of MFCCs

Reconstructed based on MFCCs (log frequency)

Quantifying Retrieval

We rank order test set songs according to KL divergence between a query multinomial and the semantic multinomials.

– 1-, 2-, 3-word queries with 5 or more examples

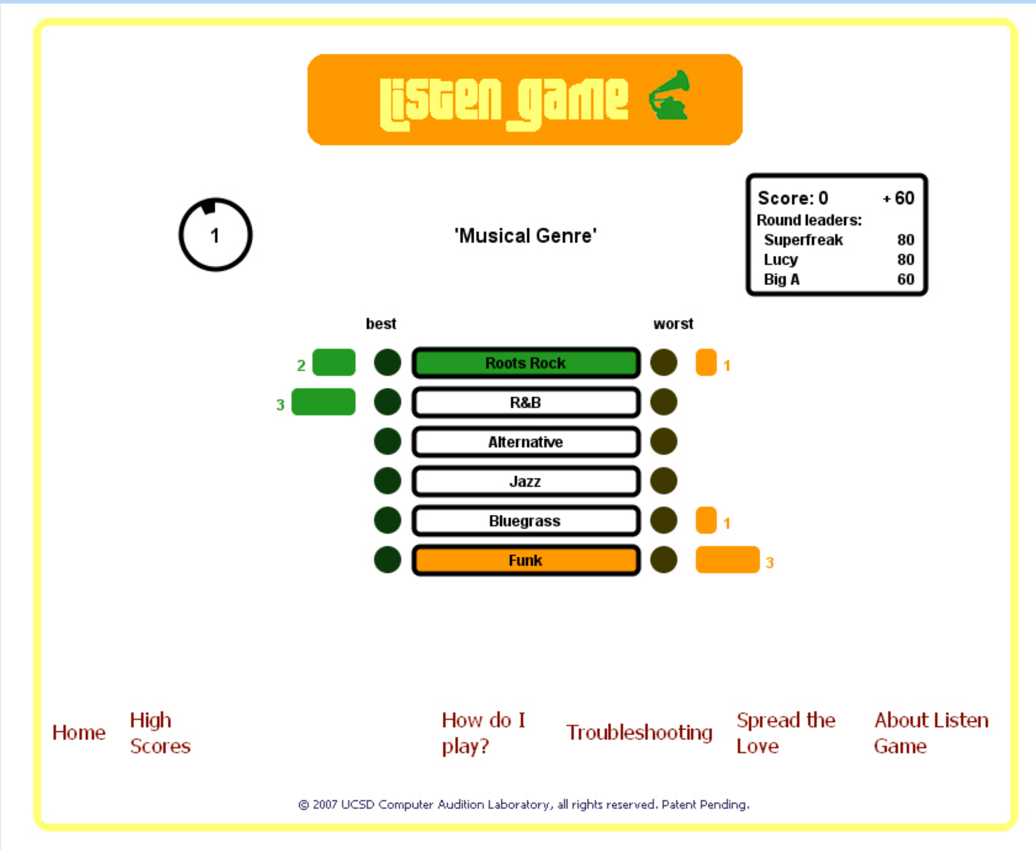
Metric: Area under the ROC Curve (AROC)

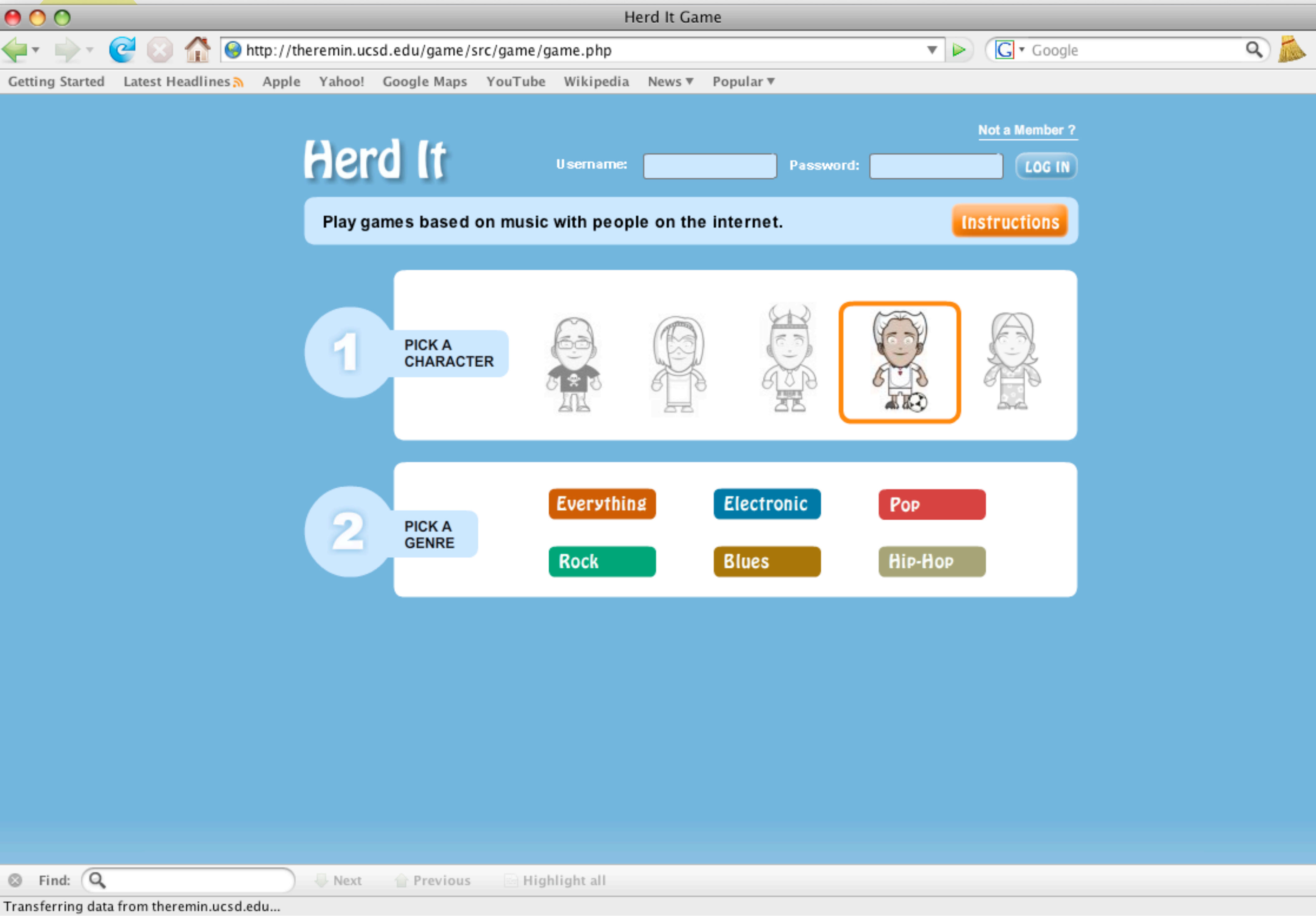
– An ROC curve is a plot of the true positive rate as a function of the false positive rate as we move down this ranked list of songs.

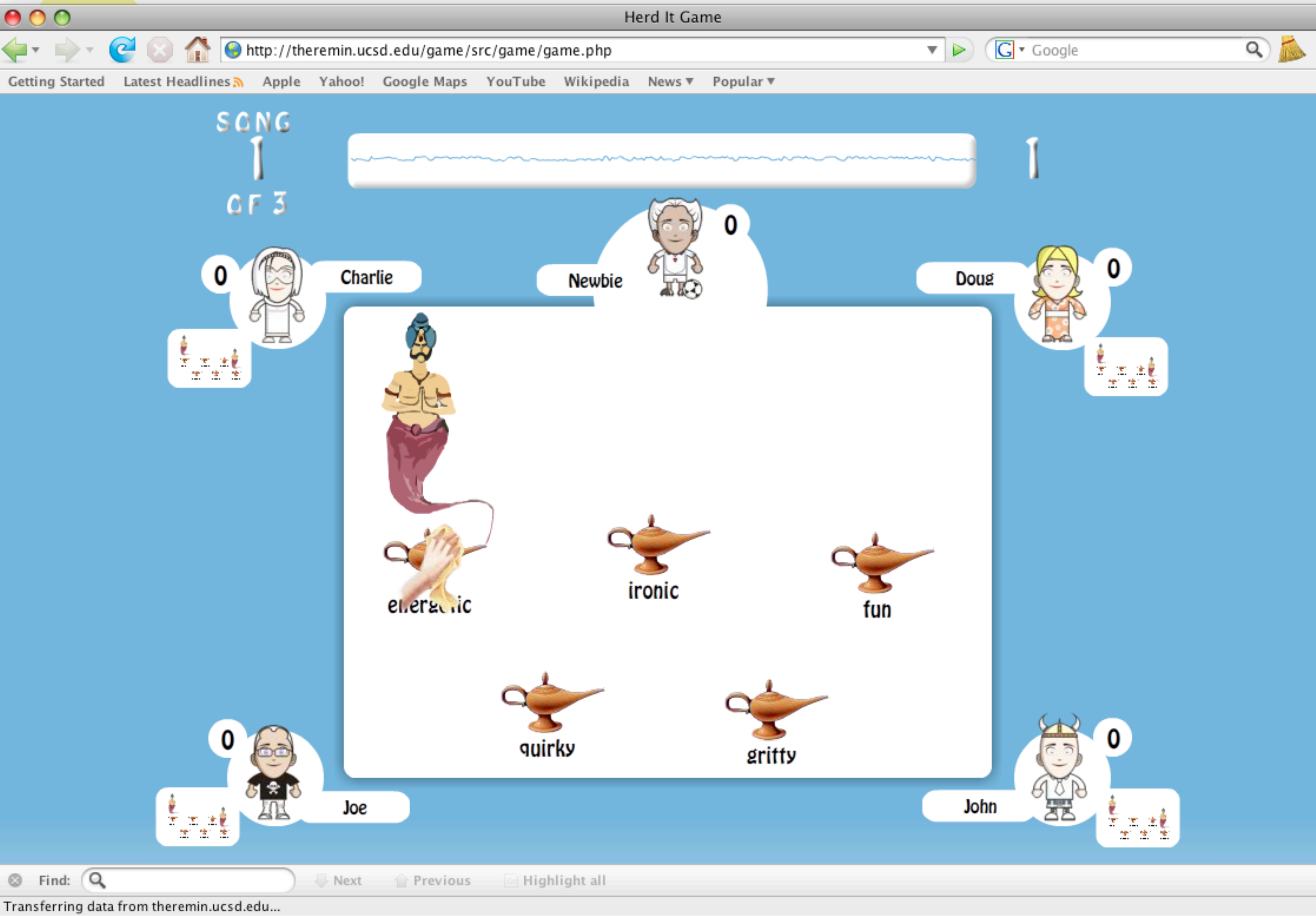
– Integrating the curve gives us a scalar between 0 and 1 where 0.5 is the expected value when randomly guessing.

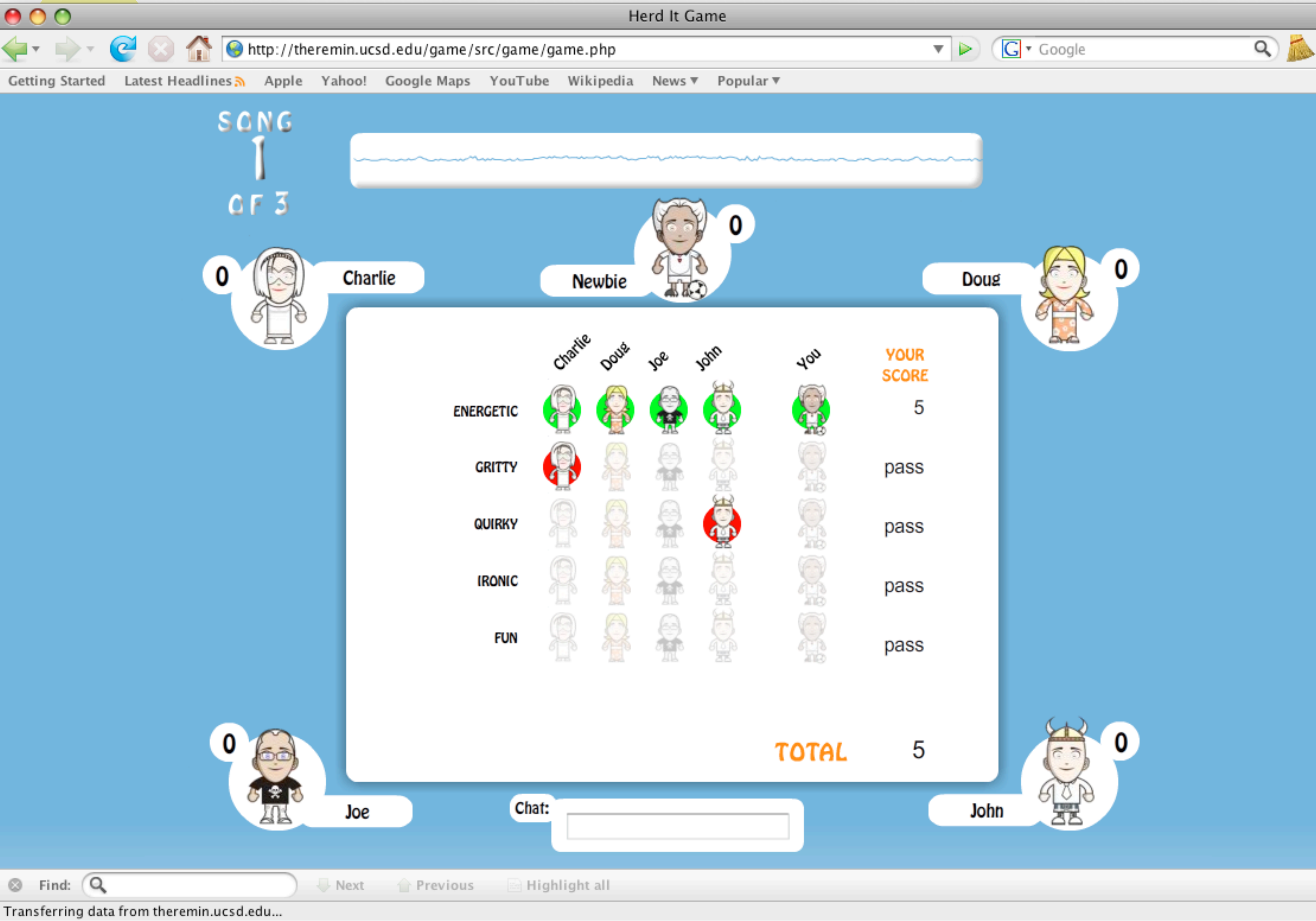
Mean AROC is the average AROC over a large number of queries.

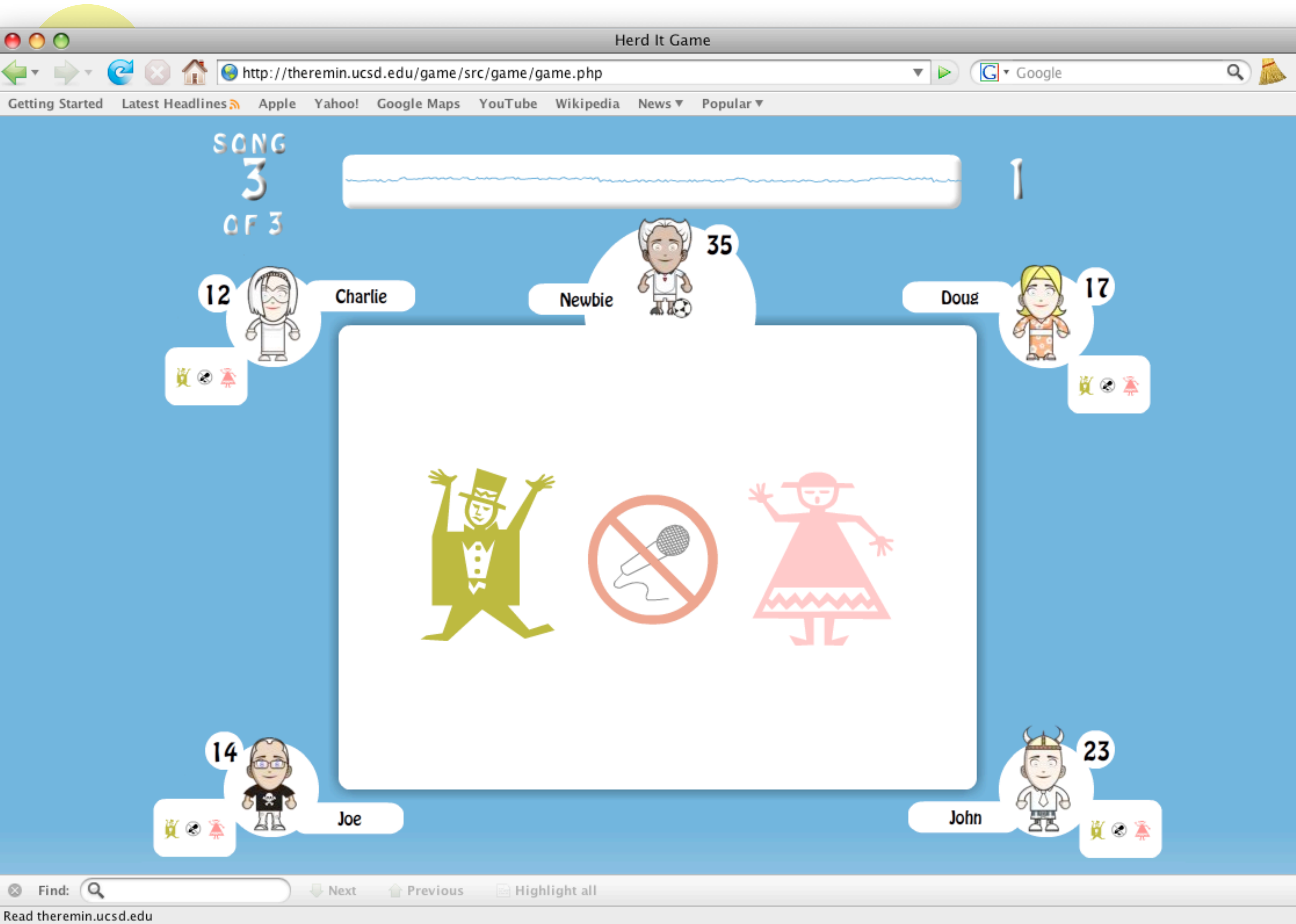
Listen Game Demo

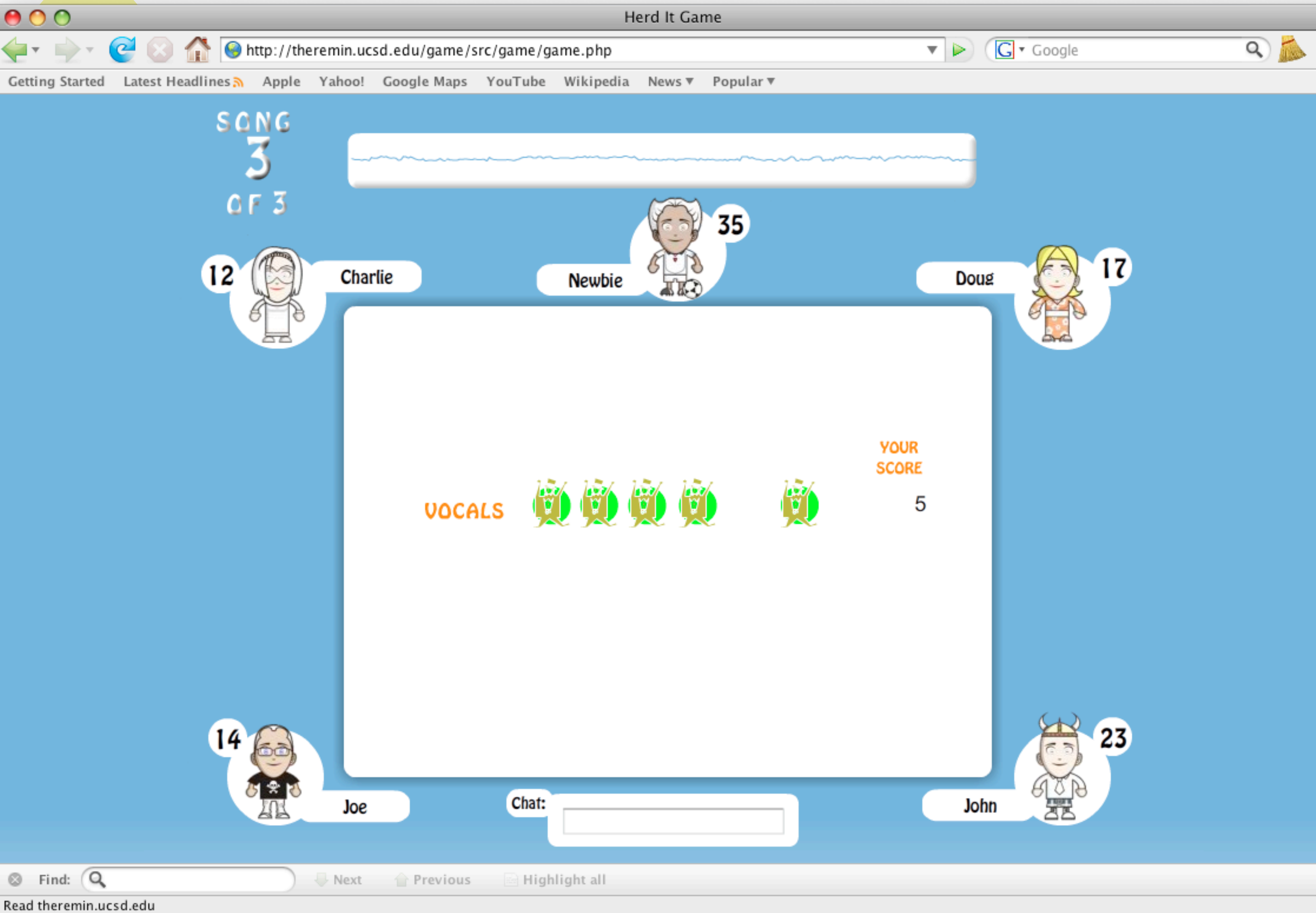



Herd It Game
http://theremin.ucsd.edu/game/src/game/game.php
Google
Getting Started
Latest Headlines
Apple
Yahoo!
Google Maps
YouTube
Wikipedia
News
Popular
Not a Member ?
Herd It
Username:
Password:
LOG IN
Play games based on music with people on the internet.
Instructions
1
PICK A CHARACTER
2
PICK A GENRE
Everything
Electronic
Pop
Rock
Blues
Hip-Hop
Find:
Next
Previous
Highlight all
Transferring data from theremin.ucsd.edu...

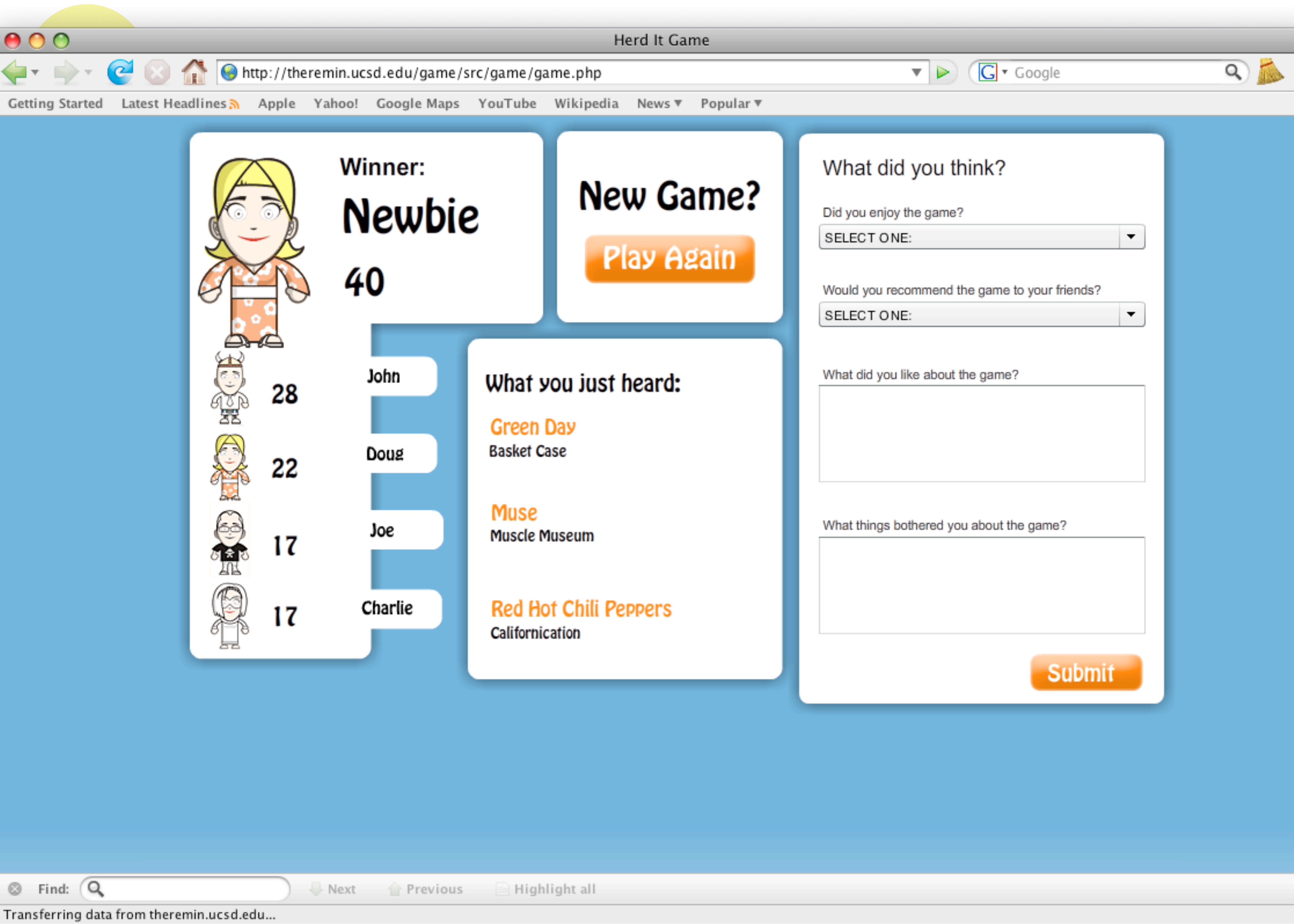

Herd It Game
http://theremin.ucsd.edu/game/src/game/game.php
Google
Getting Started
Latest Headlines
Apple
Yahoo!
Google Maps
YouTube
Wikipedia
News
Popular
SONG
1
OF 3
1
0
Newbie
0
Charlie
0
Doug
ironic
fun
quirky
gritty
0
Joe
0
John
Find:
Next
Previous
Highlight all
Transferring data from theremin.ucsd.edu...


Herd It Game
http://theremin.ucsd.edu/game/src/game/game.php
Google
Getting Started
Latest Headlines
Apple
Yahoo!
Google Maps
YouTube
Wikipedia
News
Popular
SONG 1 OF 3
0
Charlie
Newbie
0
Doug
0
Charlie
Doug
Joe
John
You
YOUR SCORE
ENERGETIC
5
GRITTY
pass
QUIRKY
pass
IRONIC
pass
FUN
pass
TOTAL
5
0
Joe
Chat:
John
0
Find:
Next
Previous
Highlight all
Transferring data from theremin.ucsd.edu...


Herd It Game
http://theremin.ucsd.edu/game/src/game/game.php
Google
Getting Started
Latest Headlines
Apple
Yahoo!
Google Maps
YouTube
Wikipedia
News
Popular
SONG 3 OF 3
1
35
Newbie
12
Charlie
Doug
17
14
Joe
John
23
Find:
Next
Previous
Highlight all
Read theremin.ucsd.edu



Herd It Game
http://theremin.ucsd.edu/game/src/game/game.php
Google
Getting Started
Latest Headlines
Apple
Yahoo!
Google Maps
YouTube
Wikipedia
News
Popular
SONG 3 OF 3
35
Newbie
12
Charlie
17
Doug
YOUR SCORE
VOCALS
5
14
Joe
Chat:
23
John
Find:
Next
Previous
Highlight all
Read theremin.ucsd.edu


Herd It Game
http://theremin.ucsd.edu/game/src/game/game.php
Google
Getting Started
Latest Headlines
Apple
Yahoo!
Google Maps
YouTube
Wikipedia
News
Popular
Winner:
Newbie
40
28
John
22
Doug
17
Joe
17
Charlie
New Game?
Play Again
What you just heard:
Green Day
Basket Case
Muse
Muscle Museum
Red Hot Chili Peppers
Californication
What did you think?
Did you enjoy the game?
SELECT ONE:
Would you recommend the game to your friends?
SELECT ONE:
What did you like about the game?
What things bothered you about the game?
Submit
Find:
Next
Previous
Highlight all
Transferring data from theremin.ucsd.edu...