A faded, blue-tinted photograph of a man in a suit and hat playing an acoustic guitar. The man is looking towards the camera with a slight smile. The background is a textured, mottled grey.

Prologue

The  
Techno-Socio-Economic  
Music Revolution

# The Revolution

Compressed Digital Format  
Cheap Hard Disks  
Free Recoding Software  
High Bandwidth Networks

**Techno**



**Socio**

P2P Music Sharing  
Social Networks

**Economic**

Self-Promotion  
Add-Driven  
Subscription Based  
Expensive Concert Tickets

# Proliferation

Supply

Demand



**8M Artists**



**173M iPods**



**150M Songs**

**Music Search & Discovery Technology**



**6B Songs**



**1M Downloads/ Month**



**31% Americans**

# Music Search

**Search** – retrieving of specific audio content

Common Paradigms:

1. Query-by-**Metadata**
2. Query-by-**Performance**
3. Query-by-**Sample**

# Music Discovery

**Discovery** – serendipitous exploration

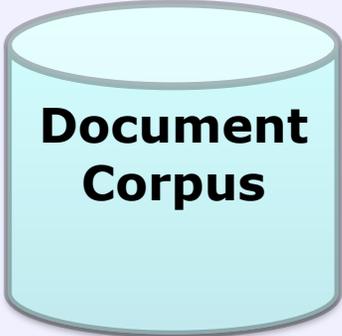
Common Paradigms:

1. Recommendation-by-**Popularity**
2. Browse-by-**Genre**
3. Query-by-**Similarity**
4. Query-by-**Description**

# Chapter 1

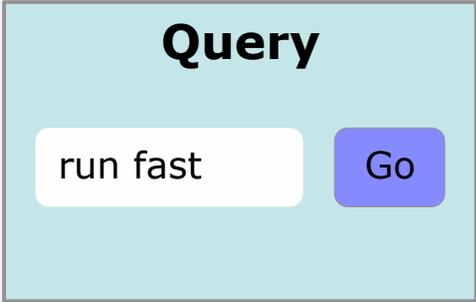
## Music Annotation is Hard

# IR Basics



**Index**

	D1	D2	D3	...
'run'	2		3	
'spot'	3			
'silly'		1		
'fast'		1	2	
...				



# “Apples and Oranges”



How do we annotate **music** with **words**?

## Tags

- text-based semantic token
- real-valued weight  $\propto$  strength of association

RHCP - *Give it away*

“an **aggressive punk-rock** song with a **funky bassline** and subtle use of **jew’s harp**”

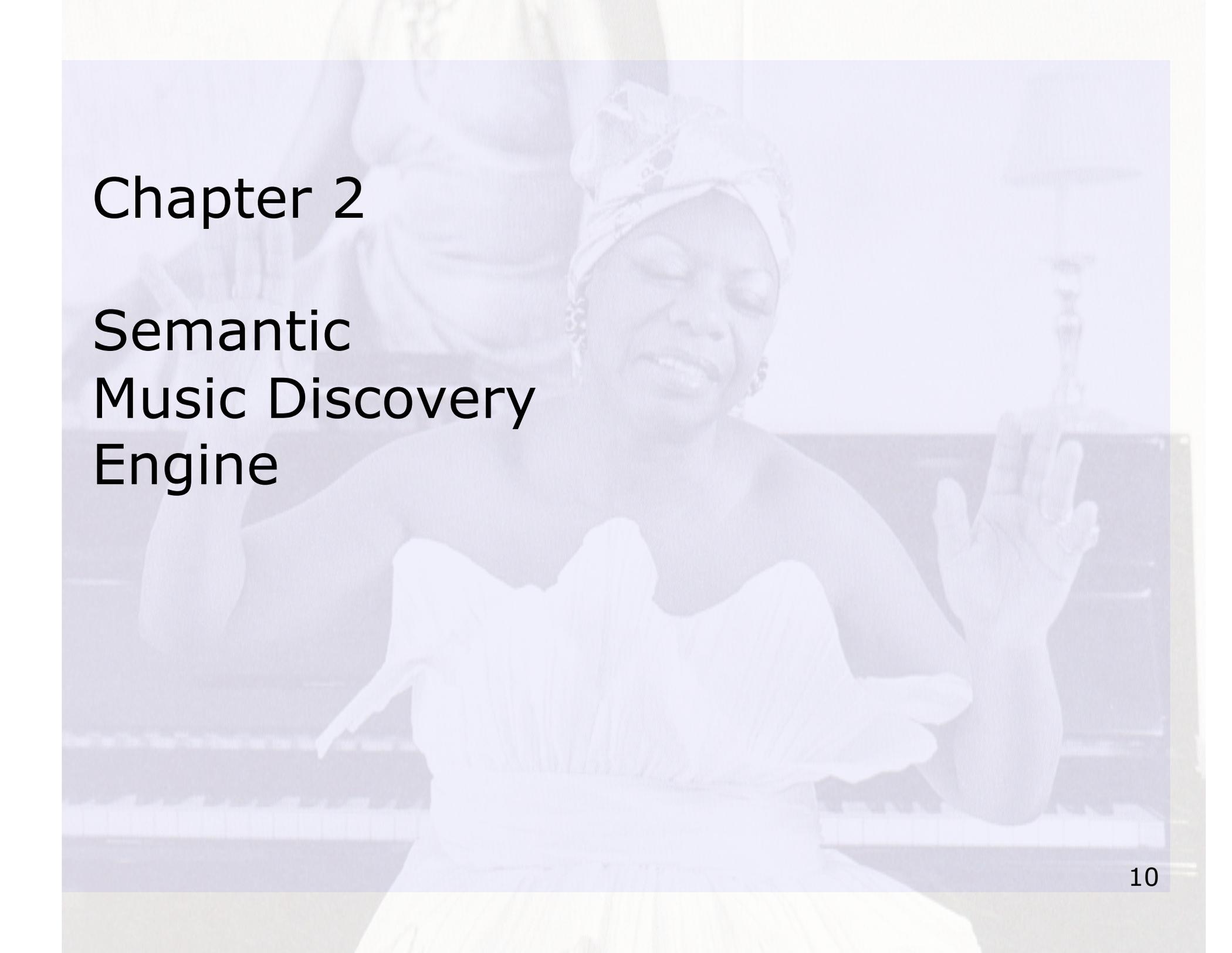
“There just ain’t no truth at all”

## **Music is subjective**

- Personal Experience
- Socially-Situated

## **Approach:**

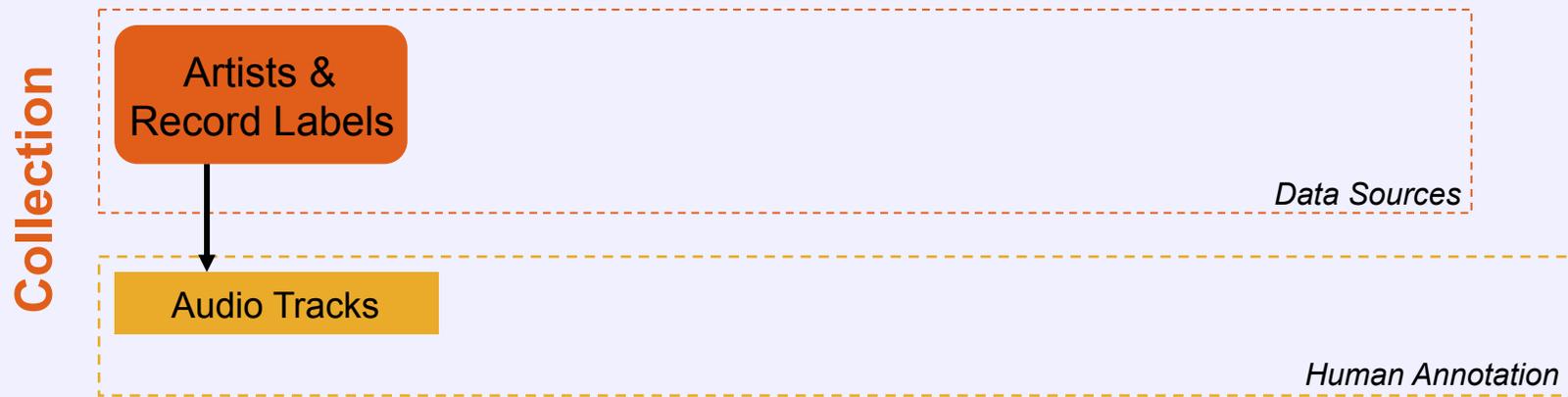
Use multiple sources of music information

A woman with her eyes closed, wearing a white headwrap and a white off-the-shoulder dress, is playing a piano. Her hands are raised in a gesture of passion or concentration. The background is slightly blurred, showing a lamp and other furniture.

## Chapter 2

# Semantic Music Discovery Engine

# Semantic Music Discovery Engine



**Extraction**

**Discovery**

# Music Corpora

**Last.fm** - 150M songs by 16M artists

**CAL500** - 500 songs by 500 artist

**Long Tail Behavior**- Chris Anderson (2004)



## Cold Start Problem

- poorly annotated songs can not be discovered

## Popularity Bias

- less popular song tend to be poorly annotated

# Semantic Music Discovery Engine



Extraction

Discovery

# Metadata

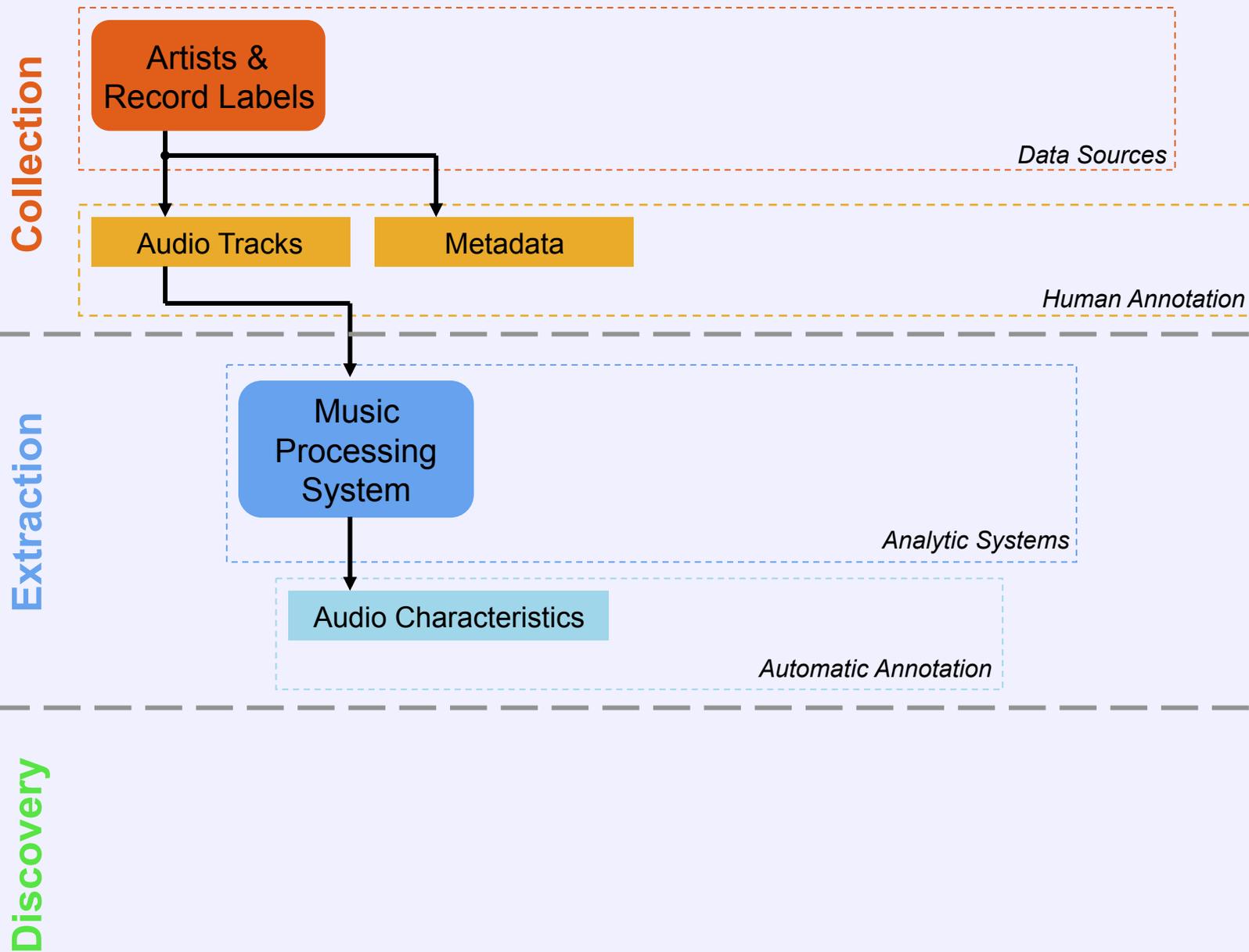
## Factual information about music

- song, album, artist, record label
- lyrics
- chronology, biography
- popularity, awards

## **Heterogeneous** data

- strings, numbers, images, graphs

# Semantic Music Discovery Engine



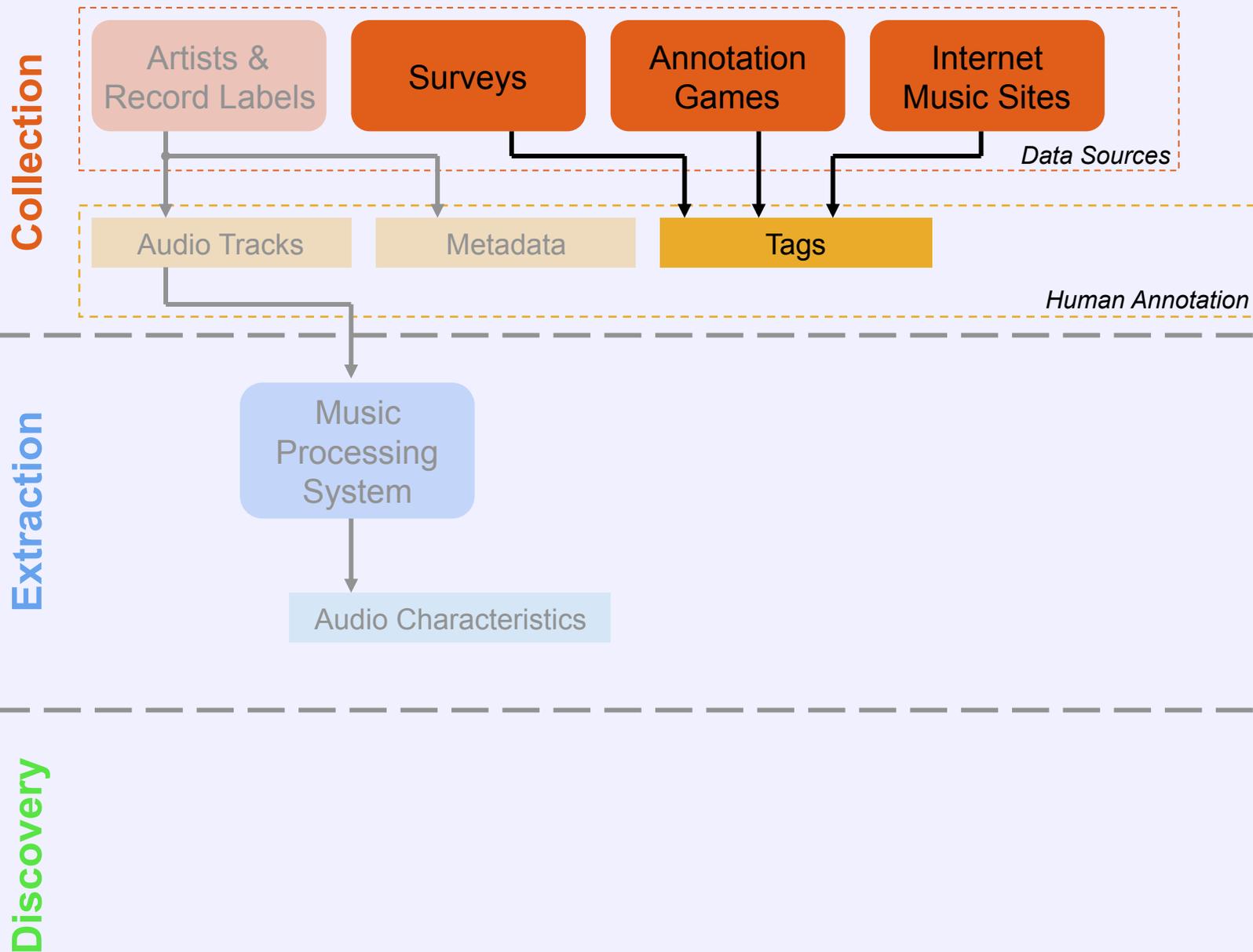
# Music Processing Systems

Information extracted from audio signal

- **Acoustic** - noise, roughness
- **Rhythmic** - tempo, patterns
- **Harmonic** - key, major/minor
- **Structural** - chorus locations

Useful for **Experts**

# Semantic Music Discovery Engine



# Surveys

## **Pandora** Music Genome Project

- 400 “Objective” Genes
- 50 trained music experts
- 750,000 songs annotated



# Surveys

## CAL500 Survey

- 174-tag vocab - genre, emotion, ...
- Paid 55 undergrads to annotate music for 120 hours
- 500 songs annotated by 3+ people

INSTRUMENTATION									
Which instruments are present, are prominent, or are featured in a solo.									
Instrument	None	Uncertain	Present	Prominent	Solo		Instrument	None	Uncertain
Voice							- String Ensemble	<input type="radio"/>	<input type="radio"/>
- Male Lead Vocals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		- Orchestra	<input type="radio"/>	<input type="radio"/>
- Female Lead Vocals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Wind Instruments		
- Backing vocals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		- Harmonica	<input type="radio"/>	<input type="radio"/>
- Choir	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		- Trumpet	<input type="radio"/>	<input type="radio"/>
Guitar Family							- Trombone	<input type="radio"/>	<input type="radio"/>
- Acoustic Guitar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		- Saxophone	<input type="radio"/>	<input type="radio"/>
- Electric Guitar (clean)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		- Horn Section	<input type="radio"/>	<input type="radio"/>
- Electric Guitar (distorted)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		Electronics		
- Slide Guitar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		- Samples	<input type="radio"/>	<input type="radio"/>
- Bass	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		- Ambient Sounds	<input type="radio"/>	<input type="radio"/>
- Banjo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="checkbox"/>		- Scratches	<input type="radio"/>	<input type="radio"/>

# Human Annotations

## Conducting a *survey*

- ✓ Reliable, Precise, Tailored to Application
- ✗ Expensive, Laborious, Not Scalable

# Annotation Games

## “Human-Computation”

- Web-based, multi-player game with real-time interaction
- Player contribute useful annotations through game play
- **ESPGame** for **images** [von Ahn]
- **Listen Game** for **songs**

# Listen Game

The screenshot shows the 'Listen Game' interface. At the top, there is a title 'listen\_game' with a gramophone icon. Below the title, a circular timer shows '1'. The main task is titled 'Musical Genre'. To the right, a score box shows 'Score: 0 + 60' and 'Round leaders: Superfreak 80, Lucy 80, Big A 60'. The central part of the screen displays a list of musical genres: Roots Rock, R&B, Alternative, Jazz, Bluegrass, and Funk. Each genre is represented by a horizontal bar with a green dot on the left and a black dot on the right. The 'best' and 'worst' labels are positioned above the bars. The 'Roots Rock' bar is highlighted in green and has a green bar with the number '2' to its left. The 'Funk' bar is highlighted in orange and has an orange bar with the number '3' to its right. The 'Bluegrass' bar has an orange bar with the number '1' to its right. The 'R&B' and 'Alternative' bars have no bars to their right. The 'Jazz' bar has a black dot to its right. The 'Home', 'High Scores', 'How do I play?', 'Troubleshooting', 'Spread the Love', and 'About Listen Game' links are located at the bottom of the interface.

**listen\_game** 🎧

Score: 0 + 60  
Round leaders:  
Superfreak 80  
Lucy 80  
Big A 60

1

'Musical Genre'

best worst

2 3

Roots Rock R&B Alternative Jazz Bluegrass Funk

1 3

Home High Scores How do I play? Troubleshooting Spread the Love About Listen Game

© 2007 UCSD Computer Audition Laboratory, all rights reserved. Patent Pending.

# Human Annotation

## Survey

- ✓ Reliable, Precise, Tailored to Application
- ✗ Expensive, Laborious, Not Scalable

## Annotation Game

- ✓ Cheap, Scalable, Precise, Personalized
- ✗ Need to create a viral user experience



# Music Web Sites

## 1. Social Tagging Site

- Users annotate music with tags
- Last.fm - 960K distinct tags



# Music Web Sites

## 2. Collecting Web Documents

- Song & Album Reviews
- Artist Biographies
- Music Blogs, Discussion Boards
- Allmusic, Rolling Stone, Amazon, Mog

# Web Document

## Genres:

Funk (3)  
Funk-metal  
Funk-rock  
Pop  
Rap

## Vocals:

Nasal  
Staccato Enunciation  
Distinctive vocals

## Instruments:

Guitar  
Bass  
Jew's-harp

## Adjective:

Hard-rocking (2)  
Noisy  
Scratchy

allmusic  
http://www.allmusic.com/cg/amg.dll?p=amg

## Give It Away

Red Hot Chili Peppers

**Composed By**  
Flea/John Frusciante/Anthony Kiedis/Chad Smith

**Other Lyrics**  
All Performers

### Song Review

The first single off the Red Hot Chili Peppers' quadruple-platinum "Give It Away" didn't achieve the massive pop success of its follow-up, "Under the Bridge" in 1992, but it did become one of the band's most instantly recognizable songs. The track features nonsensical raps, Flea's jumping, sliding, popping bass line, and Peppers' background; plus, MTV jumped all over the visually distinctive video of the bandmembers cavorting in the desert wearing silver body paint. Kiedis' lyrics are full of positive vibes, tributes to musical heroes, and free love, and they are easy to understand as Kiedis' nasal, staccato enunciation. But that distinct sound is also one of the most comprehensible lines even catchier and more memorable, greatly aided by Frusciante's guitar should not be underappreciated either, his noisy and textured sound adds and texture to the powerhouse rhythm section of Flea and Chad Smith. The song's most unpredictable change-ups: a sudden contrast to Kiedis' hyperactive delivery, a pre-recorded and dubbed backwards over the rhythm track, and a sample of a horn until the song's outro and bears a more than suspicious resemblance to the Peppers' and hard-rocking, horny and cheerfully loopy all at the same time. "Give It Away" is one of the Peppers' best singles, and a landmark single in relation to popular

# Collecting an Annotated Music Corpus

## Survey

- ✓ Reliable, Precise, Tailored to Application
- ✗ Expensive, Laborious, Not Scalable

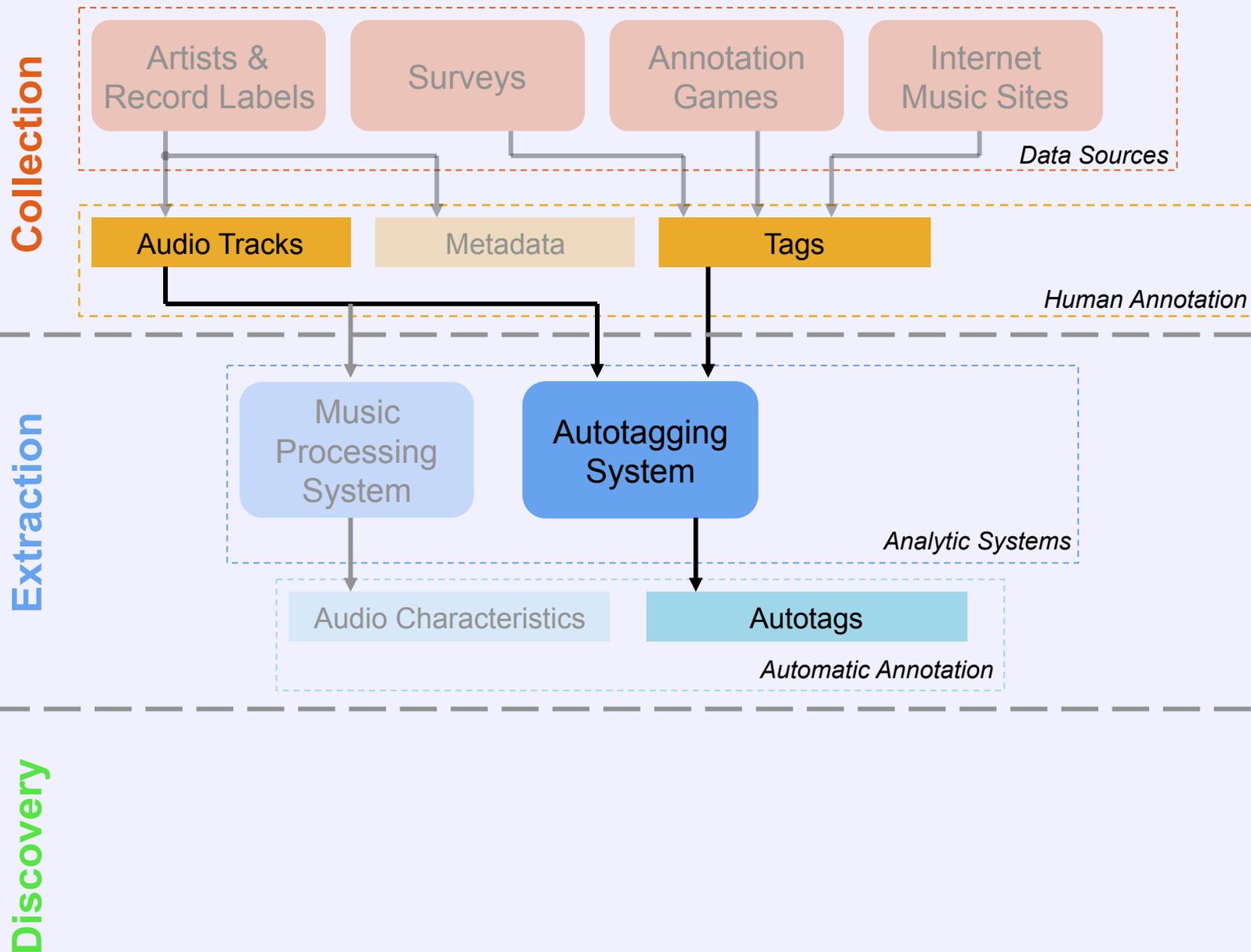
## Annotation Game

- ✓ Cheap, Scalable, Precise, Personalized
- ✗ Need to create a viral user experience

## Music Web Sites

- ✓ Cheap, Annotations for short-head
- ✗ Noisy, long-tail is poorly represented

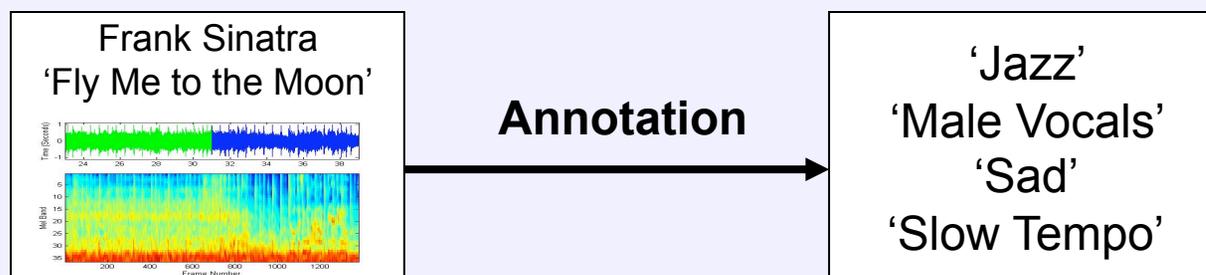
# Semantic Music Discovery Engine



# Autotagging



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



# Statistical Model

## **Supervised Multi-class Labeling** model

- One Gaussian Mixture Model (GMM) per tag -  $p(x|t)$
- Mixture Hierarchies EM Algorithm

### **Notes:**

- Developed for image annotation [Carneiro 06]
- Scalable and Parallelizable
- Top system on 2008 MIREX Autotagging Task

# Automatic Music Reviews

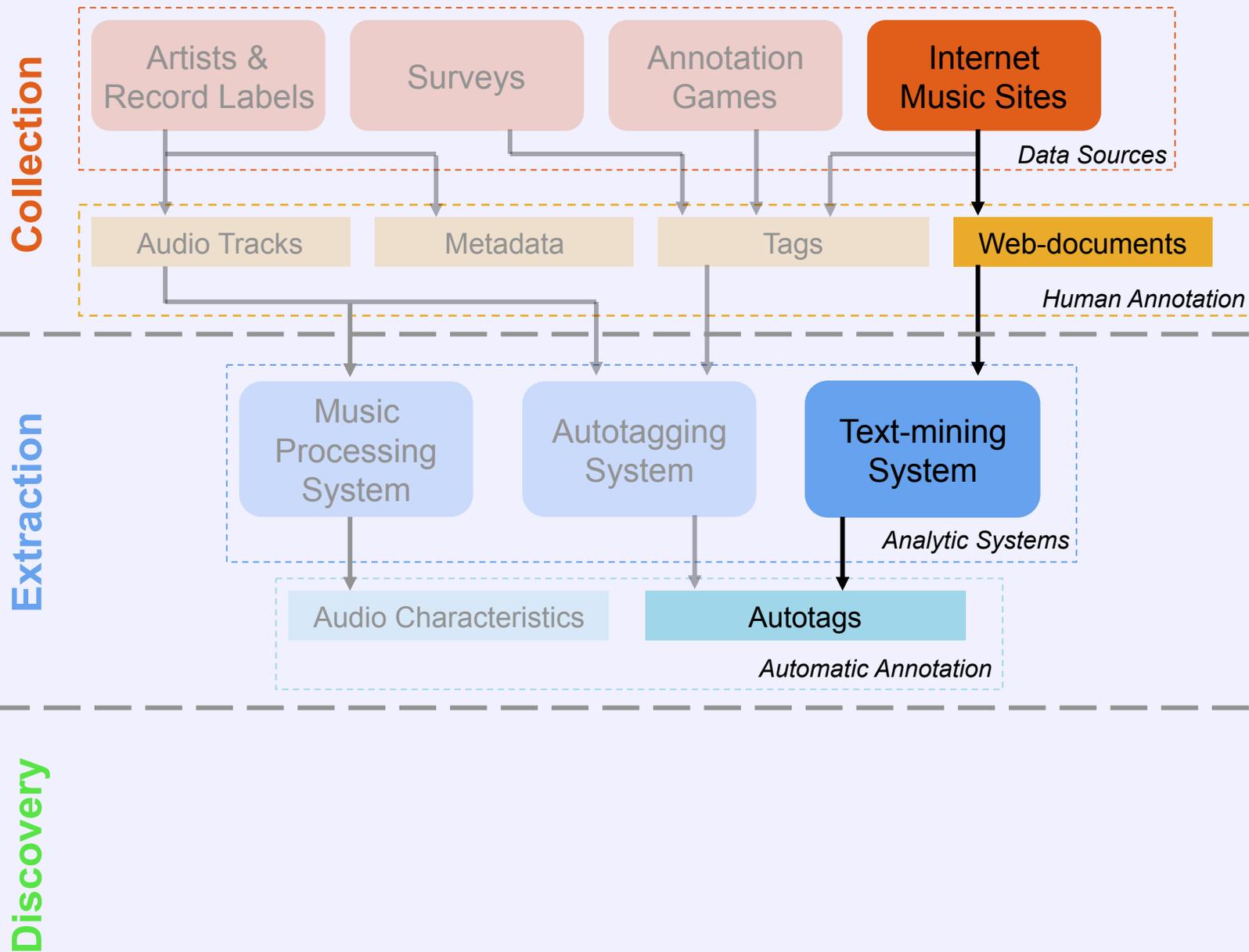
**Dr. Dre** (feat. Snoop Dogg) - Nuthin' but a 'G' thang 

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**.

**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**.

# Semantic Music Discovery Engine



# Text-mining System

## Relevance Scoring [Knees 08]

### Step 1: Collect Corpus

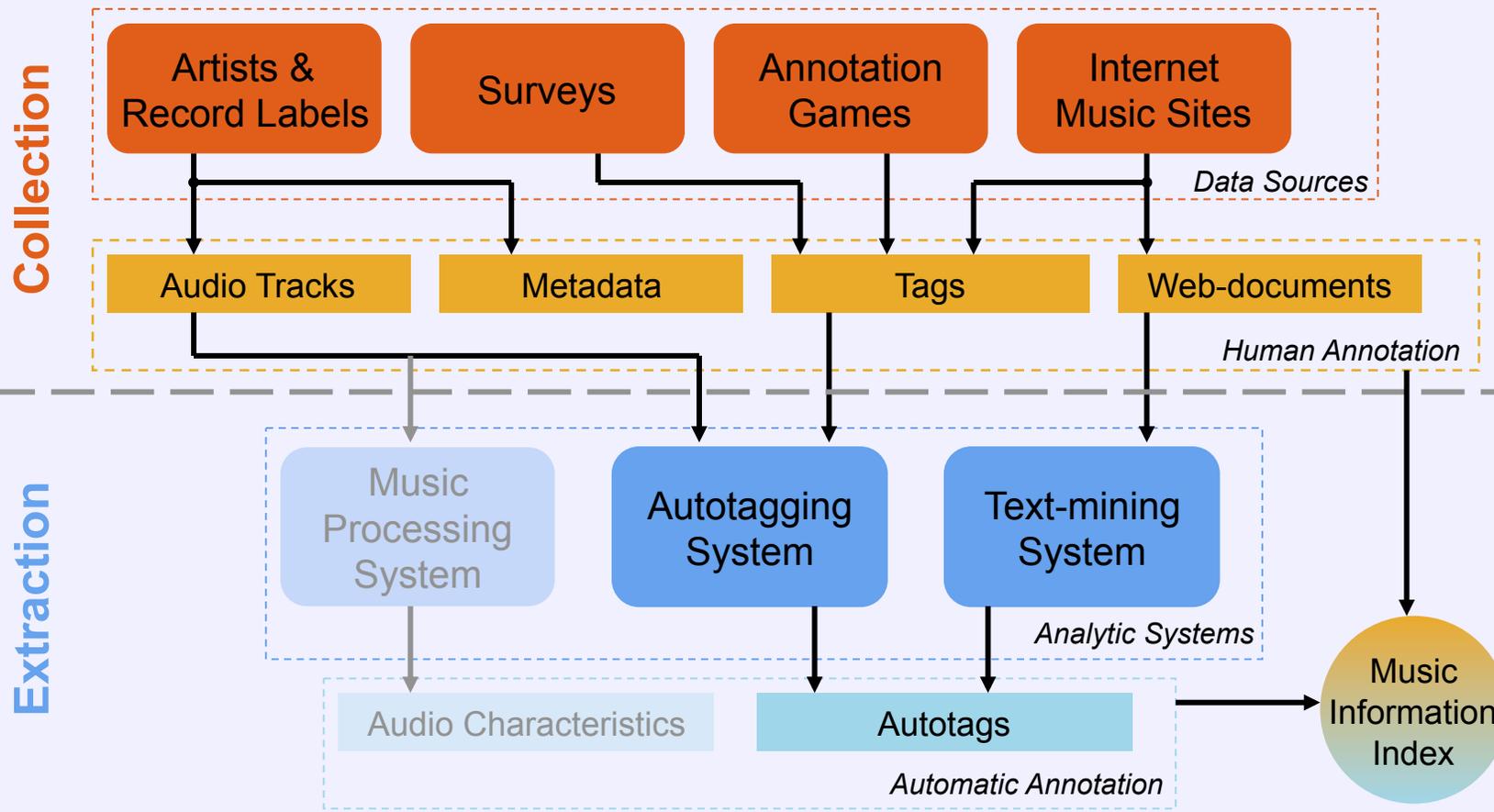
For each song, use a **search engine** to **retrieve web pages**:

- site:<website> "<artist>" music
- site:<website> "<artist>" "<album>" music review
- site:<website> "<artist>" "<song>" music review

### Step 2: Generate Tags

1. Query corpus with tag to find relevant pages
2. Map relevant pages back to songs

# Semantic Music Discovery Engine



Collection

Extraction

Discovery

# Comparing Data Sources

## Groundtruth

- **CAL500** - binary labeling of song-tag pairs
- **Long Tail** - subset of 87 obscure songs

## Approaches

1. **Social Tags** - Last.fm
2. **Annotation Game** - Listen Game
3. **Web Autotags** - Site-specific relevance scoring
4. **Audio Autotags** - SML model w/ MFCCs

# Comparing Data Sources

For each approach:

For each tag:

1. Rank songs
2. Calculate **Area under the ROC curve** (AROC)
  - 0.5 random ranking (Bad)
  - 1.0 perfect ranking (Good)

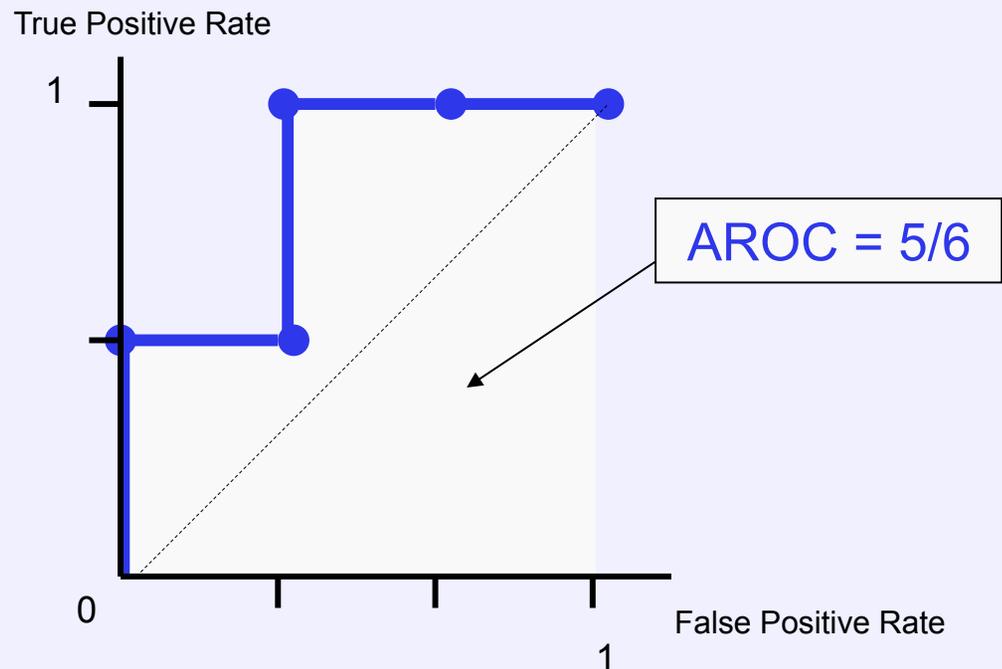
Calculate mean AROC

# Comparing Data Sources

**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



# Comparing Tags

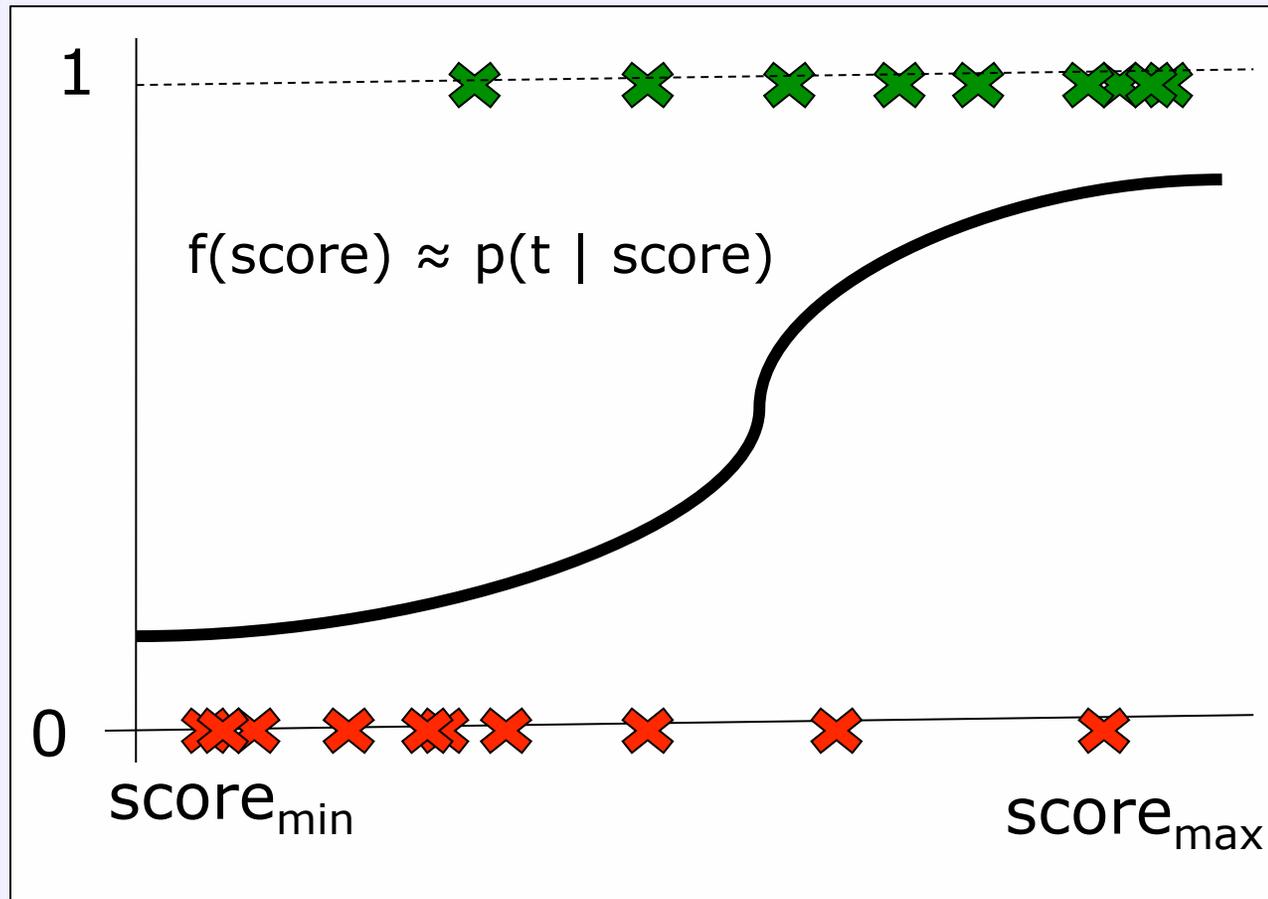
Approach	Songs	AROC
<b>Social Tags</b>	<b>CAL500</b>	<b>0.62</b>
	Long Tail	0.54
<b>Game</b>	<b>CAL500</b>	<b>0.65</b>
	Long Tail	*
<b>Web Autotags</b>	<b>CAL500</b>	<b>0.66</b>
	Long Tail	0.56
<b>Audio Autotags</b>	<b>CAL500</b>	<b>0.69</b>
	Long Tail	0.70

# Combining Data Sources

## Approaches

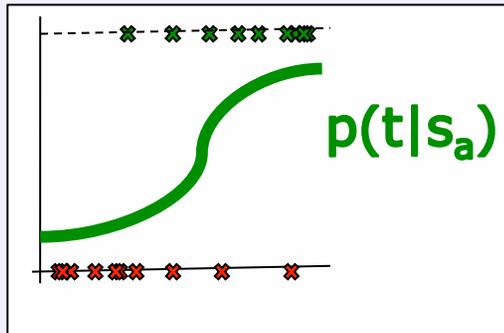
- 1. Calibrated Score Averaging** - [Zadrozny 02]
- 2. RankBoost** - [Freund 03]
- 3. Kernel Combination SVM**- [Lanckriet 04]

# 1. Calibrated Score Averaging

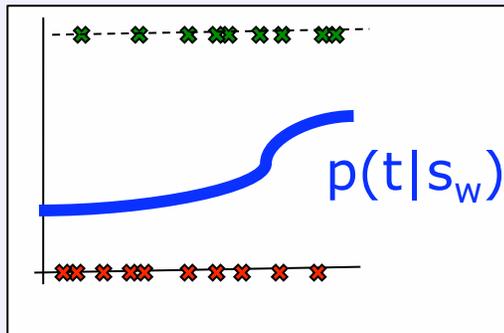


# 1. Calibrated Score Averaging

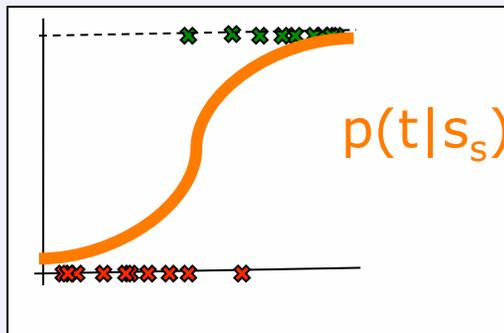
Autotag  
Score



Web  
Document  
Score



Social Tag  
Score



$$p(t|\mathbf{s}) = \prod_i p(t|s_i)$$

## 2. RankBoost

### **Weak Ranker**

- Data Source, Score Threshold & Orientation

### **Strong Ranker**

- Linear Combination of Weak Rankers

### **Rankboost**

- Iterative greedy algorithm
- picks best weak ranker & assigns weight based on loss function in each iteration

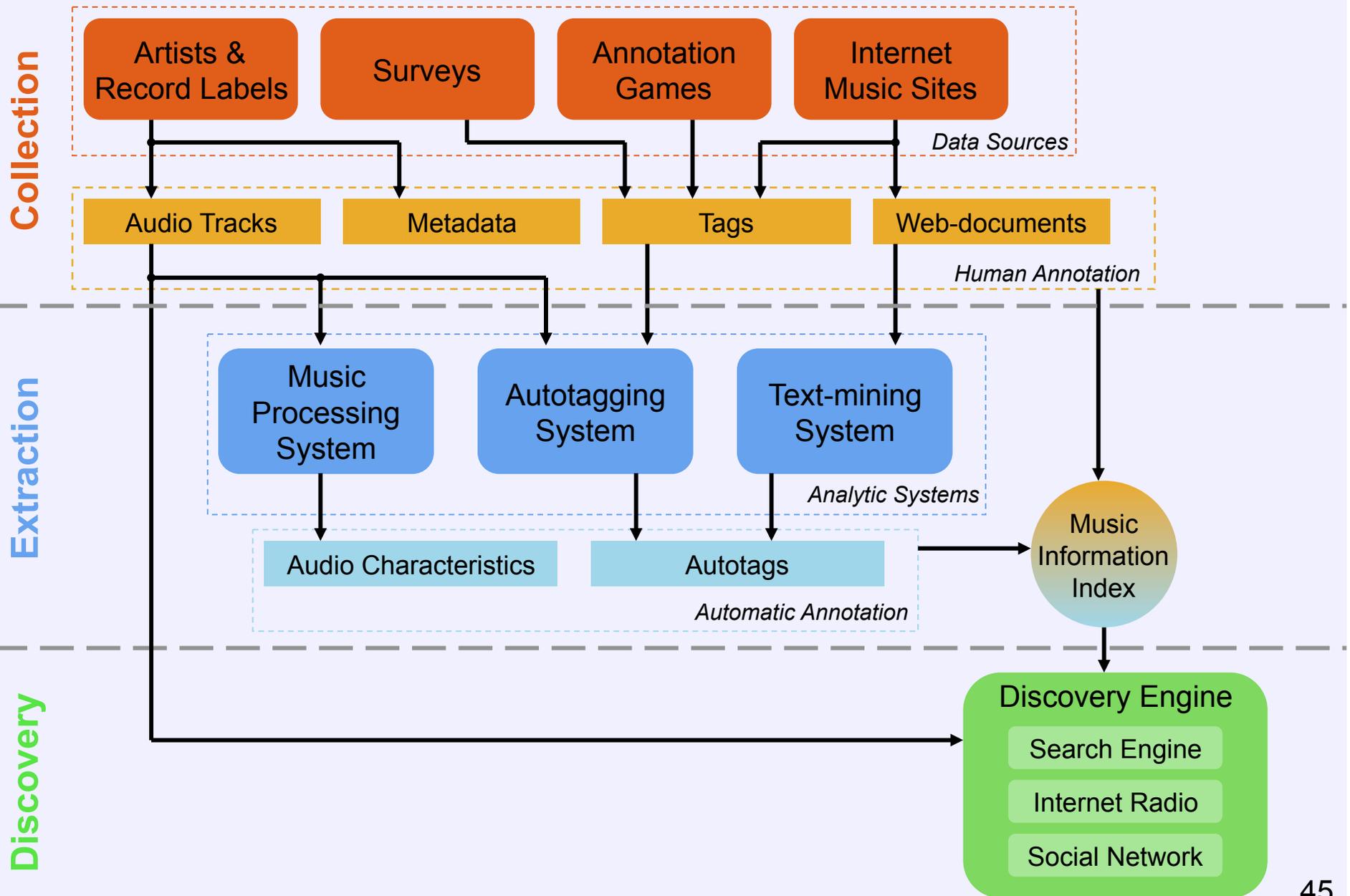
## 3. Kernel Combo SVM

1. Compute **kernel matrix** for each data source.
  - Song X Song similarity matrix
2. Learn an optimal **linear combination of the kernel matrices** using convex optimization
  - Produces single kernel matrix for SVM
3. Rank songs based on score from SVM
  - **positive distance** from **separating hyperplane**.

# Combining Tags

Approach	AROC
<b>Audio Autotags</b>	<b>0.69</b>
<b>Calibrated Score Averaging</b>	<b>0.75</b>
<b>Rankboost</b>	<b>0.75</b>
<b>Kernel Combo SVM</b>	<b>0.74</b>

# Semantic Music Discovery Engine



# Chapter 3

## The User Experience



# CA Lab Search

The screenshot shows a web browser window titled "Music Search" with the URL `http://theremin.ucsd.edu/~a5huynh/search/#`. The page features a search bar with a "search" button and example searches: "punk, calming, party". A sidebar on the left, titled "narrow your search", lists categories: bands, emotion, characteristics, genre, instrumentation, use of music, and vocal. The main content area displays a list of search results, including "Symphony of Destruction by Megadeth", "All in the Family by Korn", "The Frayed Ends of Sanity by Metallica", "Birth Defect by Helmet", "Hyperactive Child by Dead Kennedys", "The Thing That Should Not Be by Metallica", and "I Don't Want to Go Down to the Basement by Ramones". The "Symphony of Destruction by Megadeth" result is highlighted. At the bottom, a blue player bar shows playback controls and the text "Symphony of Destruction by Megadeth".

CA Lab Search

example searches: punk, calming, party

narrow your search

- bands
- emotion
- characteristics
- genre
- instrumentation
- use of music
- vocal

aggressive × electric guitar ×

Symphony of Destruction by Megadeth

All in the Family by Korn

The Frayed Ends of Sanity by Metallica

Birth Defect by Helmet

Hyperactive Child by Dead Kennedys

The Thing That Should Not Be by Metallica

I Don't Want to Go Down to the Basement by Ramones

⏪ ⏩ | Symphony of Destruction by Megadeth o hai there

# SMIR – Swarthmore Music IR

SMIR Music Discovery Experience

http://mugwort.cs.swarthmore.edu/smir/playerPage.html

Google

## SMIR

About SMIR

A new kind of music discovery experience

**Now Playing**

**Santeria**

by **Sublime**

Enter a tag or artist

I like music with these features...

ska

alternative

This song has these features...

reggae	favorites
90s	rock
ska punk	alternative
punk	sublime
ska	chill



# Herd-it

The screenshot shows a web browser window titled "Herd It on Facebook" with the URL <http://apps.facebook.com/herd-it/?ref=ts>. The Facebook navigation bar is visible at the top, showing the user's name "Douglas Turnbull" and options for "Settings" and "Logout".

The main content area features a "Herd It" logo in a red speech bubble. Below it, a blue bar prompts the user to "Start by choosing a musical style". Six buttons represent different styles with the number of players online:

Musical Style	Players Online
Everything	8
Electronic	8
Pop	3
Rock	8
Blues	5
Hip-Hop	7

Below this is a blue bar with the text "Are you ready to rock out with the world?" and a "Challenge Your Friends!" button.

On the left, a section titled "How do I play Herd It?" offers two options: "Play an interactive demo:" with a "play demo" button, and "View quick instructions:" with an "instructions" button. The word "or" is centered between these two options.

On the right, a "Top Scoring Friends" section is displayed with three tabs: "Top Scoring Friends", "Leaderboard", and "Musical Profile". The "Top Scoring Friends" tab is active, showing a list of four friends with their profile pictures, names, scores, and preferred musical styles:

Friend Name	Score	Style
Antoni	466	Rock
Luke	459	Rock
Oscar	447	Rock
Damien	439	Rock

A link "Challenge Your Friends to Play Herd It!" is located below the list.

The bottom of the browser window shows a taskbar with various application icons, including "Applications", "Bookmark Herd It", and "Online Friends (12)".

# Herd-it

facebook Home Profile Friends Inbox Douglas Turnbull Settings Logout

10-Question Survey Report a bug

Your Rank **5** / 9

GAME ROUNDS ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

TIMER

Douglas **0**

Top 10 Scorers

- 0 Fred
- 0 John
- 0 Alex
- 0 John
- 0 Douglas
- 0 Jane
- 0 Chris
- 0 Andy
- 0

What adjective describes the ...?

electronic irreverent

brash thuggish

sardonic theatrical

100% 50% 0% agree-O-meter

CHAT

Applications Bookmark Herd It Online Friends (12)

# Herd-it

facebook Home Profile Friends Inbox Douglas Turnbull Settings Logout

10-Question Survey Report a bug

Your Rank 1/9

GAME ROUNDS ●●●●●○○○○

TIMER

Douglas 209

Top 10 Scorers

- 209 Douglas
- 203 Jane
- 191 John
- 181 Alex
- 180 Fred
- 180 Andy
- 169 Chris
- 147 John
- 124

DANCEABLE

UNDANCEABLE

SLOW ← → FAST

Alex John Jane John Fred Nick Andy

100% 50% 0%

44

agree-o-meter

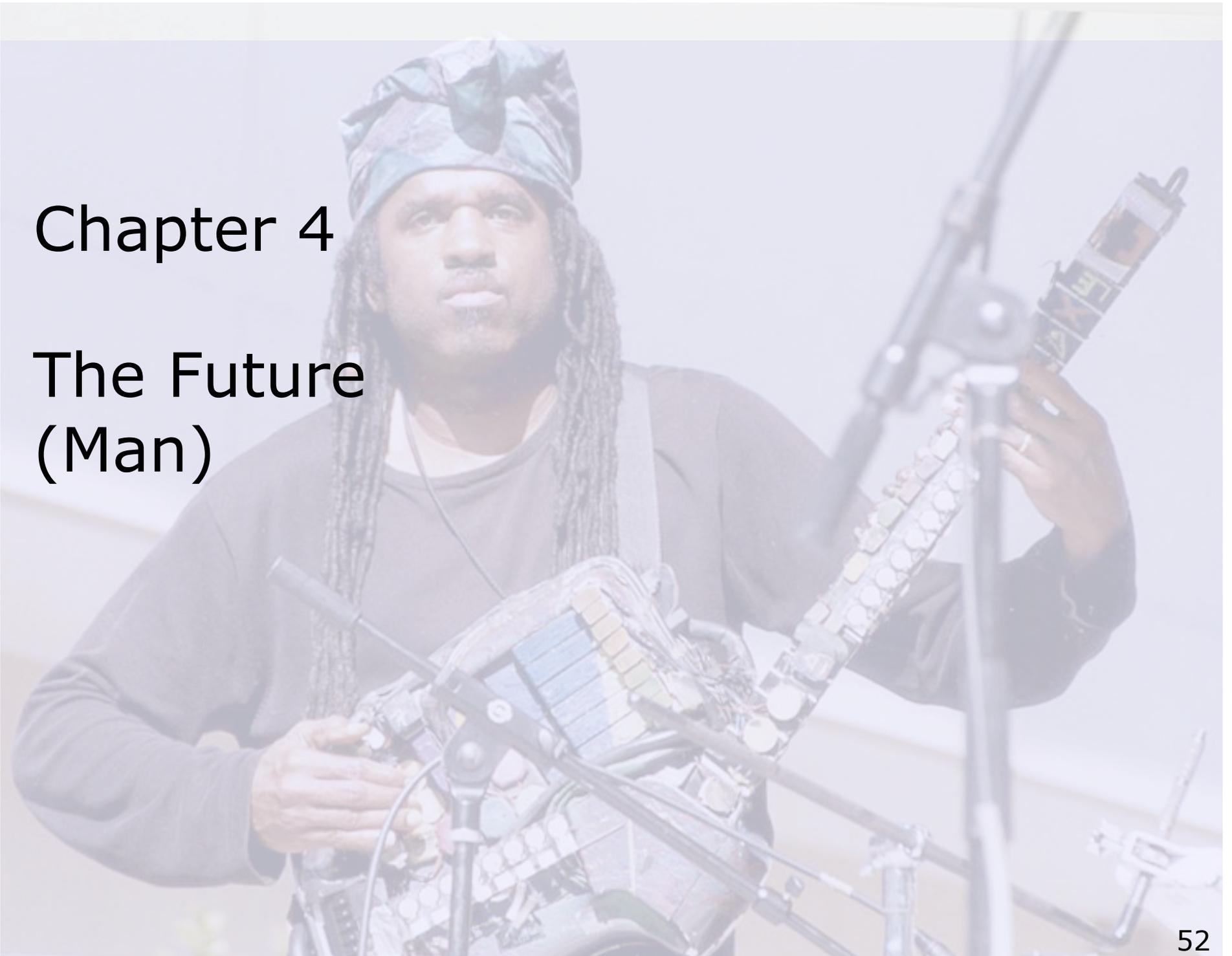
CHAT

Douglas: Hello World  
Douglas: What is going on  
Last Song: 'I'm Back' by Eminem  
Last Song: 'Heat' by 50 Cent  
Douglas: West Coast!!!  
Last Song: 'Lowrider' by Cypress Hill

Applications Bookmark Herd It Online Friends (12)

# Chapter 4

## The Future (Man)



# Research Challenges

## 1. Combine Music Information Sources

- Games, Social Networks, Web Documents, Popularity Info

## 2. Improving Autotagging

- Discriminative Approaches [Mandel 08, Eck 07]

## 3. Recommendation with Semantics

## 4. Personalization

- Demographic Groups
- Psychographic Groups
- Individual Preference
- Emotional states of Individual

# What's on tap

## 1. Big new data set

- 10,000 songs
- Acoustic, Genre, Social Tags

## 2. Herd-it

- Facebook launch
- Analyzing data

## 3. New Everything

- Autotagging Approaches
- Content-Context Approach
- User Interfaces

# References

Social Context-Audio Content [SIGIR 09, ISMIR 08]

Autotagging [IEEE TASLP 08, SIGIR 07]

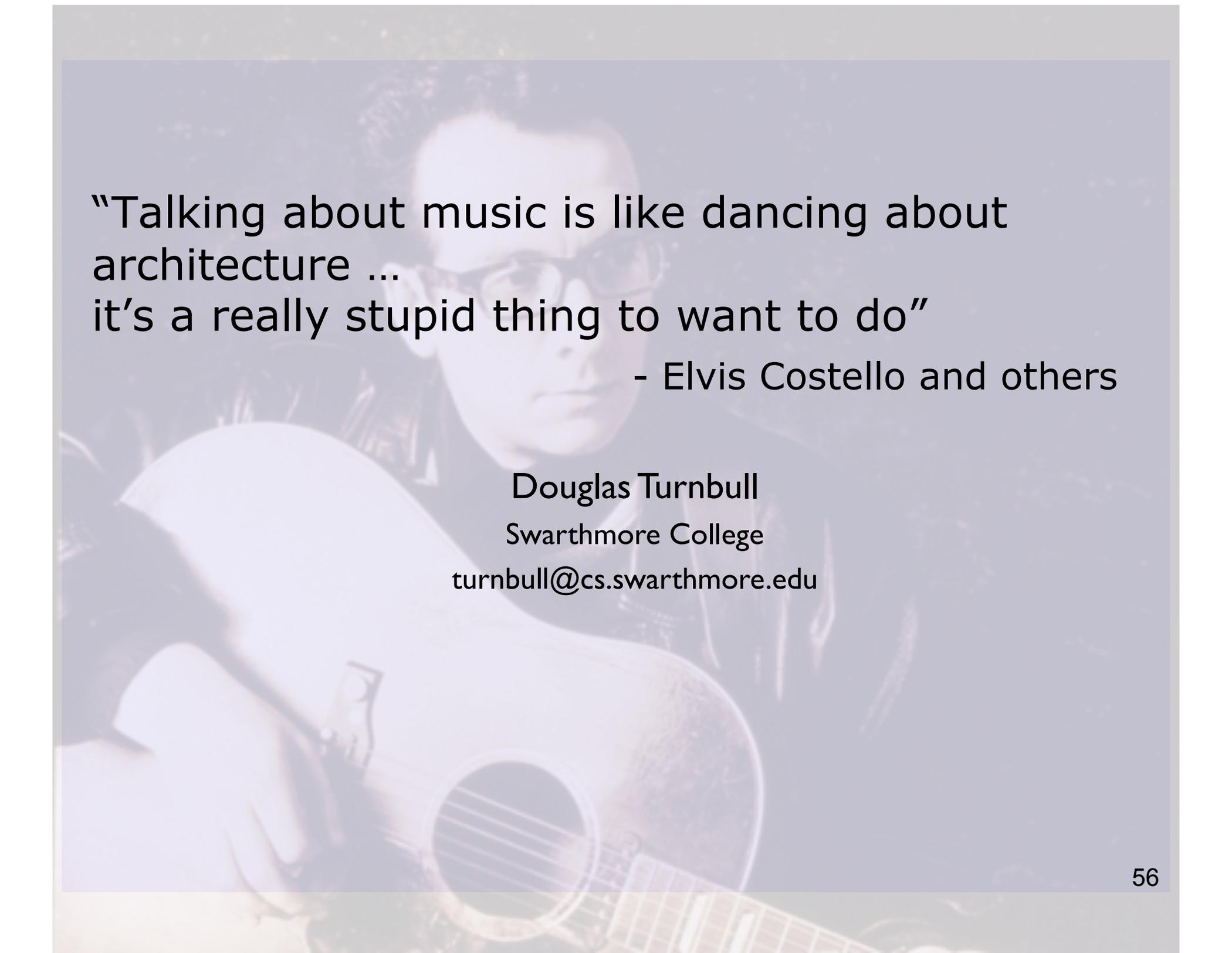
Music Annotation Games [ISMIR 07a]

## **Related:**

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

Tag Vocabulary Selection with Sparse CCA [ISMIR 07]

Supervised Music Boundary Detection [ISMIR 07]



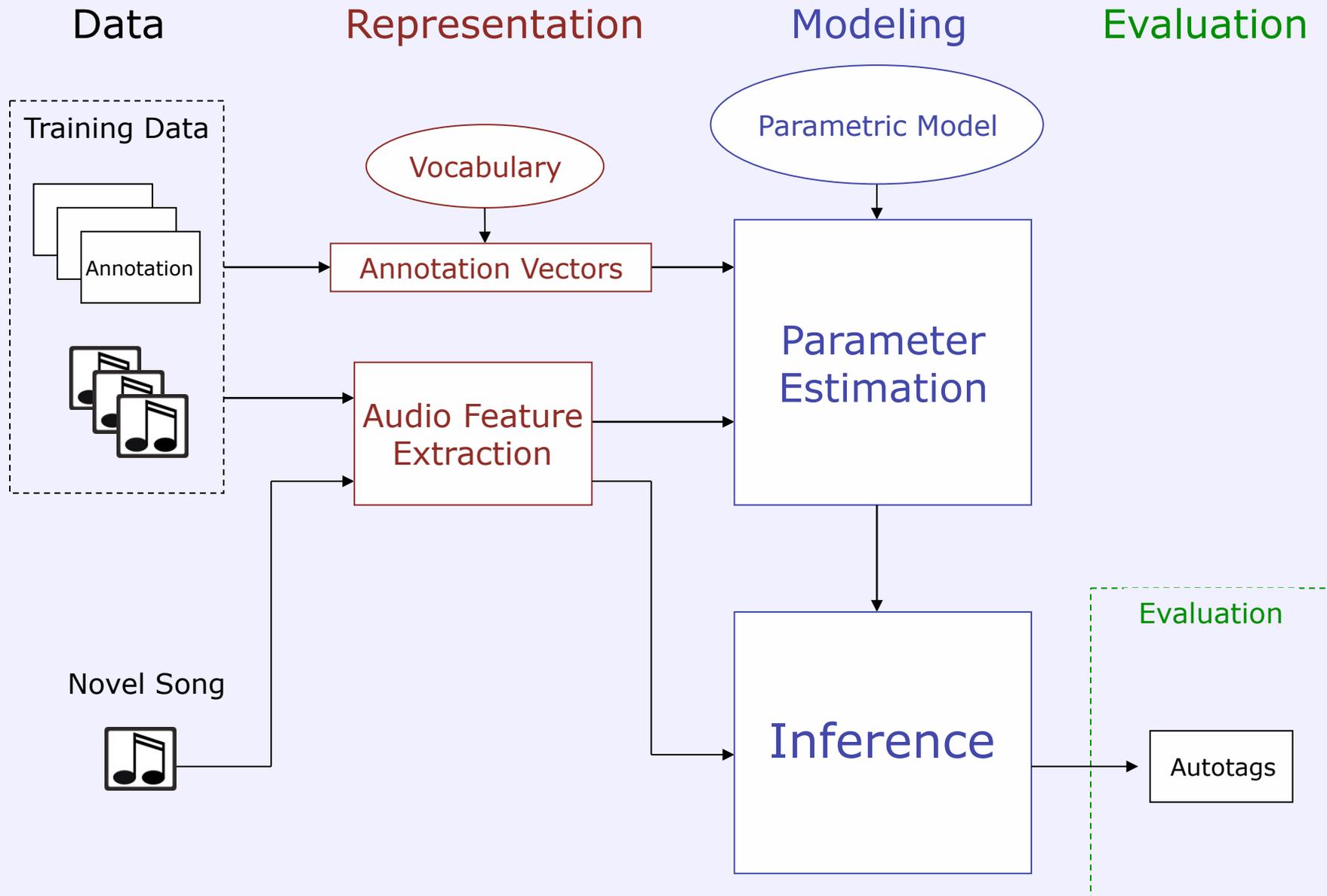
“Talking about music is like dancing about  
architecture ...  
it’s a really stupid thing to want to do”

- Elvis Costello and others

Douglas Turnbull  
Swarthmore College  
[turnbull@cs.swarthmore.edu](mailto:turnbull@cs.swarthmore.edu)

Questions?

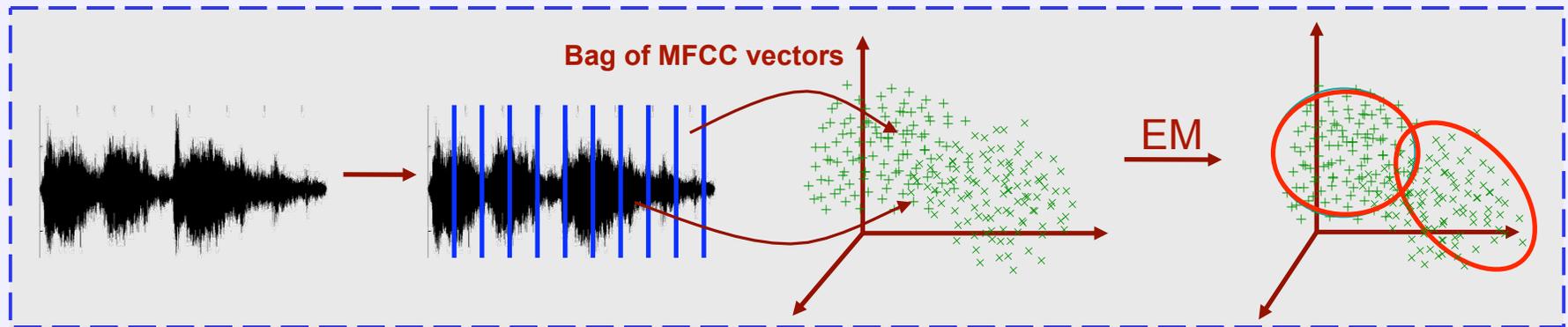
# System Overview



# Modeling a Song

## Algorithm

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

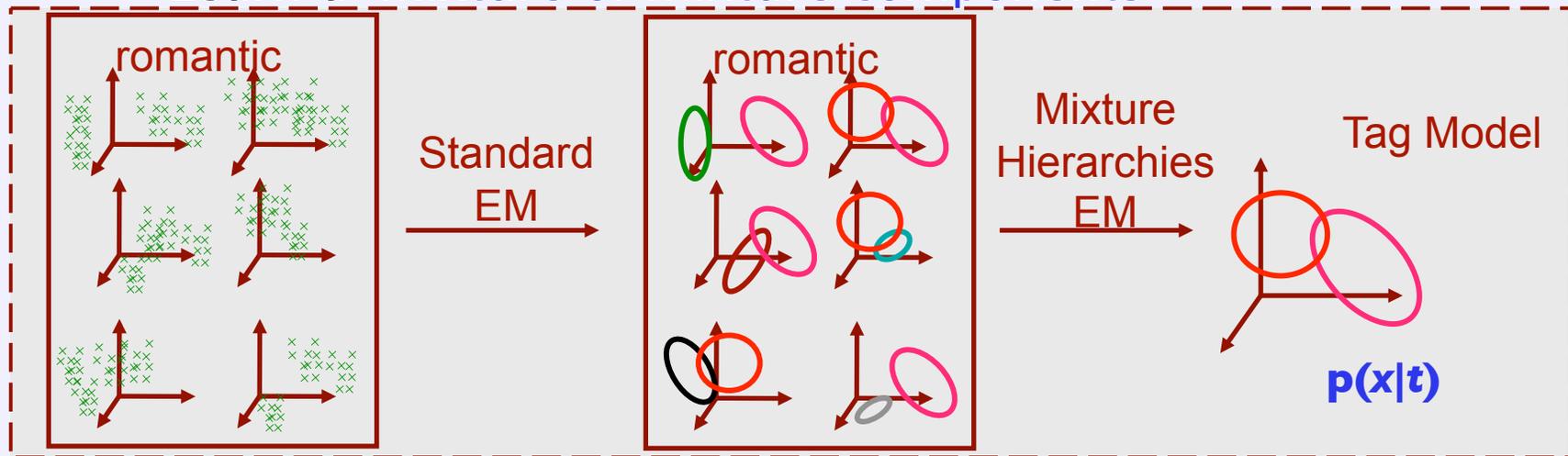


# Modeling a Tag

## 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  $\rightarrow$  better density estimate

# Annotation

Given a novel song  $X = \{x_1, \dots, x_T\}$ , calculate

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

1. Uniform tag prior
2. Vectors are conditionally independent given a tag
3. Geometric average of likelihoods
4. Tags are mutually exclusive and exhaustive

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

## Semantic Multinomial:

- $P(t|X)$ 's  $\rightarrow$  multinomial distribution over the tag vocabulary

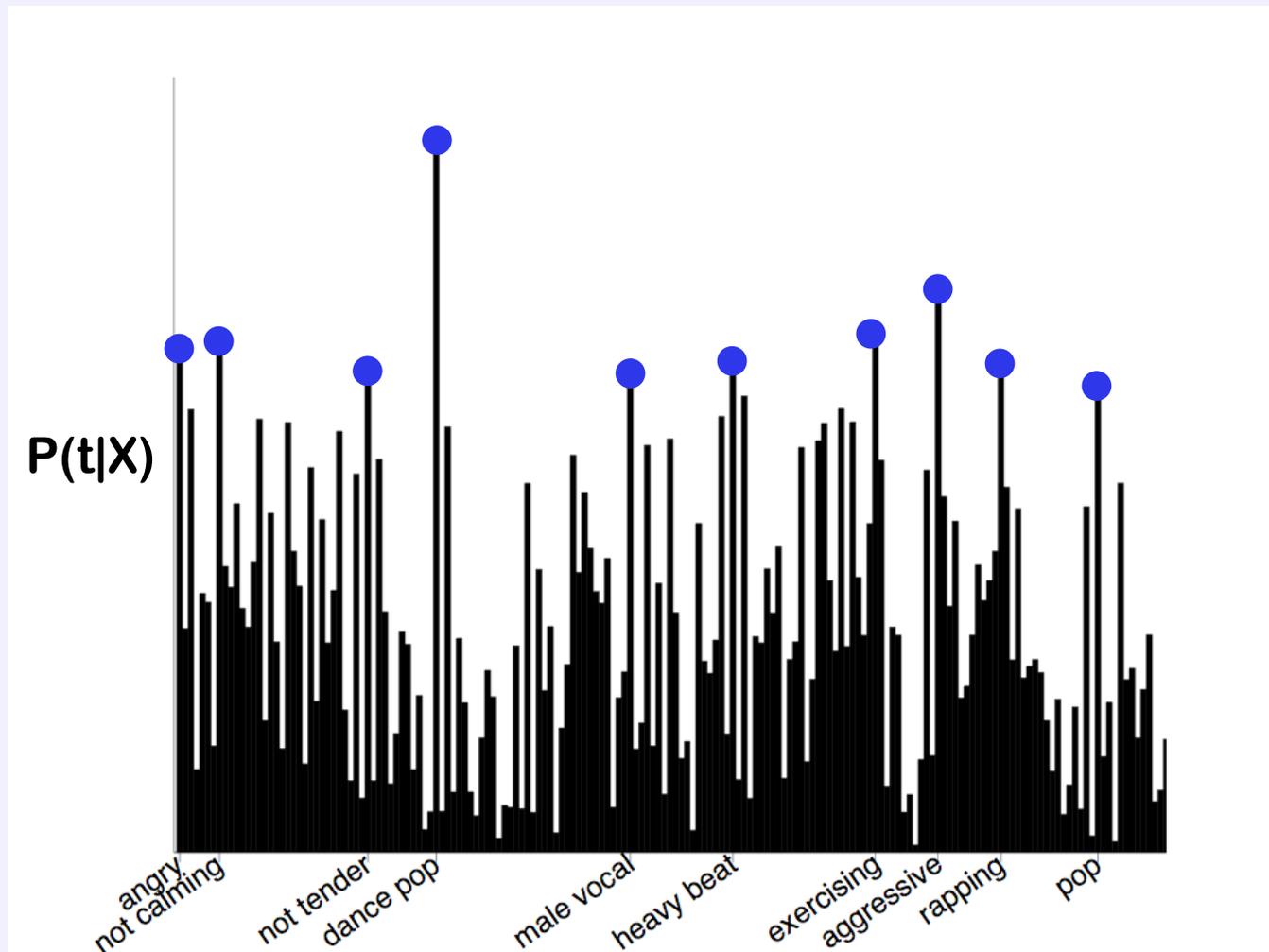
**Annotation:** peaks of multinomial



# Annotation

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

Chili



# Retrieval

1. Annotate each song in corpus with a **semantic multinomial  $\mathbf{p}$**

- $\mathbf{p} = \{P(t_1|X), \dots, P(t_{|\mathcal{V}|}|X)\}$

2. Given a text-based query, construct a **query multinomial  $\mathbf{q}$**

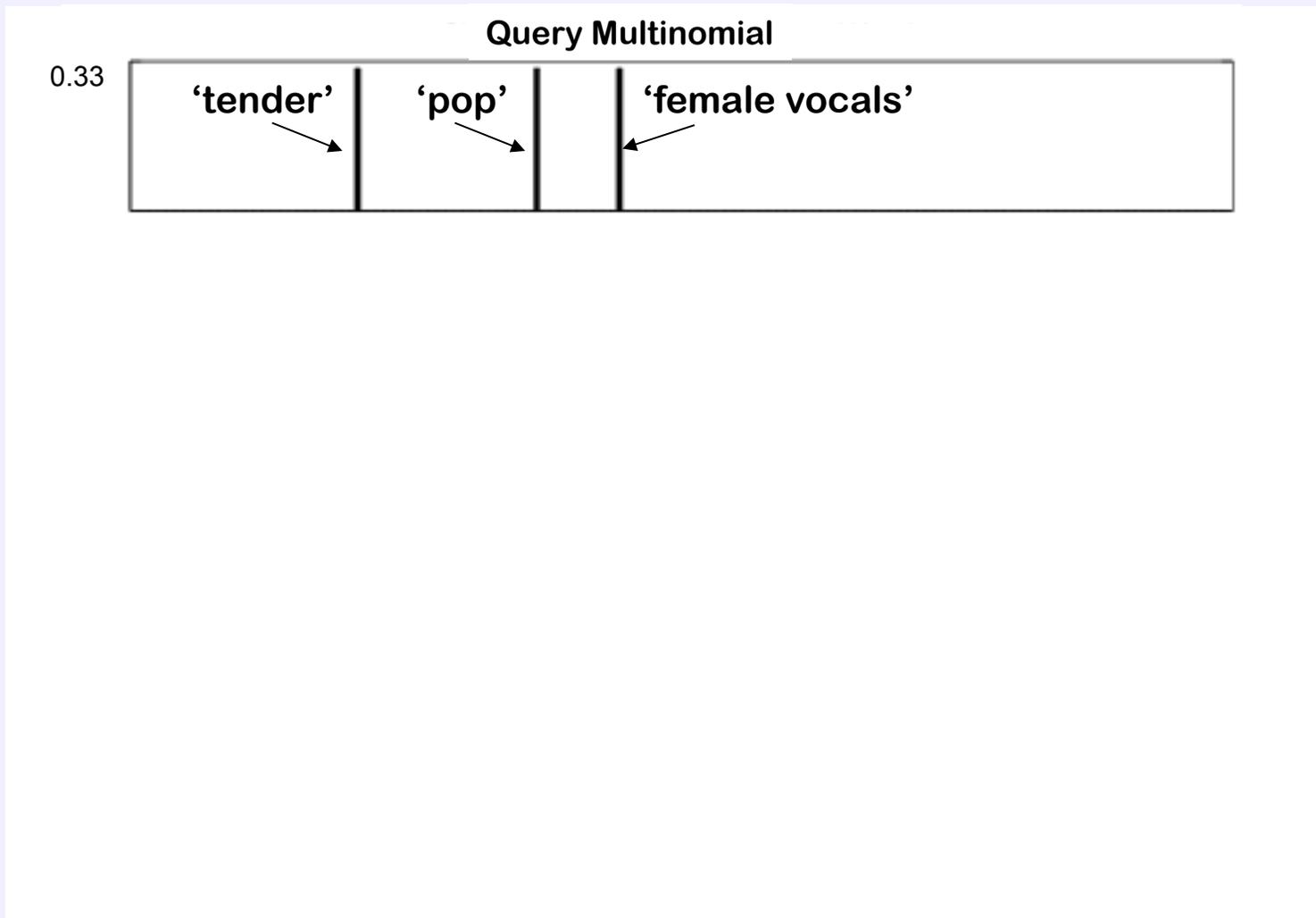
- $q_i = 1/|t|$  , if tag  $t$  appears in the query string
- $q_i = 0$ , otherwise

3. Rank all songs

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

# Retrieval

Query: 'a tender pop song with female vocals'



# Retrieval

Query	Retrieved Songs
'Tender'	<b>Crosby, Stills and Nash</b> - Guinevere  <b>Jewel</b> - Enter from the East  <b>Art Tatum</b> - Willow Weep for Me
'Female Vocals'	<b>Alicia Keys</b> - Fallin'  <b>Shakira</b> - The One  <b>Junior Murvin</b> - Police and Thieves 
'Tender' AND 'Female Vocals'	<b>Jewel</b> - Enter from the East  <b>Evanescence</b> - My Immortal  <b>Cowboy Junkies</b> - Postcard Blues 

# The Age of Music Proliferation

## Production:

- 5M artist pages  a place for friends
- 150M distinct songs  the social music revolution

## Distribution

- 1.5M simultaneous P2P users (Fe  1) -
- 27K record labels 
- 4B songs to 50M customers  -

## Consumption

- 11M Internet radio users  created by the Music Genome Project™
- 110M iPods sold 

# Comparing Tags

Approach	Songs	Density	AROC
<b>Ground Truth</b> CAL500	All	0.15	1.00
	Long-Tail	0.15	1.00
<b>Social Tags</b> Last.fm	All	0.23	0.62
	Long-Tail	0.03	0.54
<b>Game</b> Listen Game	All	0.37	0.65
	Long-Tail	*	*
<b>Web Autotags</b>	All	0.67	0.66
	Long-Tail	0.25	0.56
<b>Audio Autotags</b>	All	1.00	0.69
	Long-Tail	1.00	0.70

# Text-mining System

## Relevance Scoring [Knees 08]

### Step 1: Collect Corpus

For each song, use a **search engine** to **retrieve web pages**:

- site:<website> "<artist>" music
- site:<website> "<artist>" "<album>" music review
- site:<website> "<artist>" "<song>" music review

Maintain  $M_{s,d}$  = mapping of **songs** to **documents**

# Text-mining System

## Step 2: Autotag songs

For each tag  $t$ :

1. Query corpus with tag  $t$  to find relevant documents
  - $w_{t,d}$   $\rightarrow$  relevance score for document  $d$
2. For each song  $s$ , sum relevance scores for documents that are related to song  $s$

$$w_{s,t} = \sum_d M_{s,d} w_{t,d}$$

CAL Music Discovery Engine

UCSD **CAL Music Discovery Engine** eol

**Metadata Search**   **Semantic Search**   [Combo Search](#)

**Metadata Filter:** pick songs by song title, artist name, and album title

<b>Song Title</b> contains	<input type="text"/>	does <b>not</b> contain	<input type="text"/>
<b>Artist Name</b> contains	beatles	does <b>not</b> contain	<input type="text"/>
<b>Album Title</b> contains	<input type="text"/>	does <b>not</b> contain	<input type="text"/>
<b>Song, Artist, or Album</b> contains	<input type="text"/>	does <b>not</b> contain	<input type="text"/>

**Semantic Ranking:** order songs by musical characteristics

Musical Genre	Instrumentation	Emotional Content
Alternative <input type="checkbox"/> yes	Acoustic Guitar <input checked="" type="checkbox"/> yes	Aggressive <input type="checkbox"/> yes
Bebop <input type="checkbox"/> yes	Ambient Sounds <input type="checkbox"/> yes	Annoying <input type="checkbox"/> yes
Bluegrass <input type="checkbox"/> yes	Backing vocals <input type="checkbox"/> yes	Arousing <input type="checkbox"/> yes
Blues <input type="checkbox"/> yes	Bass <input type="checkbox"/> yes	Bizarre <input type="checkbox"/> yes
Brit Pop <input type="checkbox"/> yes	Distorted Electric Guitar <input type="checkbox"/> yes	Boring <input type="checkbox"/> yes
Classic Rock <input type="checkbox"/> yes	Electric Guitar <input type="checkbox"/> yes	Calming <input checked="" type="checkbox"/> yes
Cool Jazz <input type="checkbox"/> yes	Female Lead Vocals <input type="checkbox"/> yes	Carefree <input type="checkbox"/> yes
Country <input type="checkbox"/> yes	Hand Drums <input type="checkbox"/> yes	Cheerful <input type="checkbox"/> yes
Dance Pop <input type="checkbox"/> yes	Harmonica <input type="checkbox"/> yes	Emotionless <input type="checkbox"/> yes
Electronica <input type="checkbox"/> yes	Horn Section <input type="checkbox"/> yes	Gloomy <input type="checkbox"/> yes
Folk <input checked="" type="checkbox"/> yes	Male Lead Vocals <input type="checkbox"/> yes	Happy <input type="checkbox"/> yes

CAL Music Discovery Engine

UCSD **CAL Music Discovery Engine** eol

Metadata Search  Go Semantic Search  Go Combo Search

Combo Search:  
 Metadata Filtering - 'beatles',  
 Semantic Ranking - 'Folk', 'Acoustic Guitar', 'Calming',

Songs Found: 77 (Top 10 shown)

▶ '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 Ecstasy](#)

▶ 'Dead of Winter' by [Eels](#) on [Electro-Shock Blues](#)

▶ 'Ribbons Undone' by [Tori Amos](#) on [The Beekeeper](#) (2005)

▶ 'Yesterday' by [The Beatles](#) on [Help!](#)  
 This is a **singer/songwriter** song that also has a **country** feel. It is **calming** and **boring**. It features **acoustic guitar**, **saxophone** and **female lead vocals**. The vocals are **emotional** and **high-pitched**. It is a song with **low energy** and **soft beat** that you might like to listen to while **romancing**.

Similar Songs:

▶ 'Rose of Aberdeen' by [Simon & Garfunkel](#) on [Sounds of Silence](#)

▶ 'Moonshiner' by [Uncle Tupelo](#) on [89/93: An Anthology](#) (2002)

▶ 'Where Is the Highway Tonight%' by [Neil Young](#) on [Lucky Thirteen](#)

# Music & Technology

**Technology** is changing how **music** is **produced, distributed, promoted** and **consumed**.

# Thanks

Gert, Charles, Lawrence, Shlomo, Serge,  
Sanjoy

- Advice and perspective

Gary Cottrell, Virginia de Sa, IGERT

- Enabling creative and interdisciplinary pursuits

Damien O'malley, Aron Tremble, VLC

- Thinking beyond the walls of academia

Luke Barrington, Antoni Chan, David  
Torres

- Friends and collaborators

# Red Hot Chili Peppers

Send to Friend



Photo by Gus Van Sant

**Picture Browser**  
 < Previous    Next >

**Formed**  
 1983

**Years Active**  
 1910 20 30 40 50 60 70 80 90 2000

- | Genre | Styles       |
|-------|--------------|
| Rock  | Alternative  |
|       | Pop/ Rock    |
|       | Funk Metal   |
|       | College Rock |
|       | Dance-Rock   |
|       | Adult        |
|       | Alternative  |
|       | Pop/ Rock    |
|       | Rap-Rock     |

- Moods**
- Carefree
  - Confident
  - Provocative
  - Brash
  - Freewheeling
  - Organic

## Biography

by Greg Prato

Few rock groups of the '80s broke down as many musical barriers and were as original as the Red Hot Chili Peppers. Creating an intoxicating new musical style by combining funk and punk rock together (with an explosive stage show, to boot), the Chili Peppers spawned a slew of imitators in their wake, but still managed to be the leaders of the pack by the dawn of the 21st century. The roots of the band lay in a friendship forged by three school chums, [Anthony Kiedis](#), [Michael Balzary](#), and [Hillel Slovak](#), while they attended Fairfax High School in California back in the late '70s/early '80s. While [Balzary](#) and [Slovak](#) showed great musical promise (on trumpet and guitar, respectively), [Kiedis](#)... [Read More...](#)

Only Capture... Can... from... For me

verizon wireless



Watch music videos by this artist!

## Other Entries

- [Movie Entry](#)

## Group Members

- [Flea](#)
- [John Frusciante](#)
- [Jack Irons](#)
- [Anthony Kiedis](#)
- [Arik Marshall](#)

## Influenced By

- [Sylvester "Sly Stone" Stewart](#)
- [George Clinton](#)
- [Bootsy Collins](#)
- [Led Zeppelin](#)
- [Gang of Four](#)
- [Funkadelic](#)
- [Bad Brains](#)