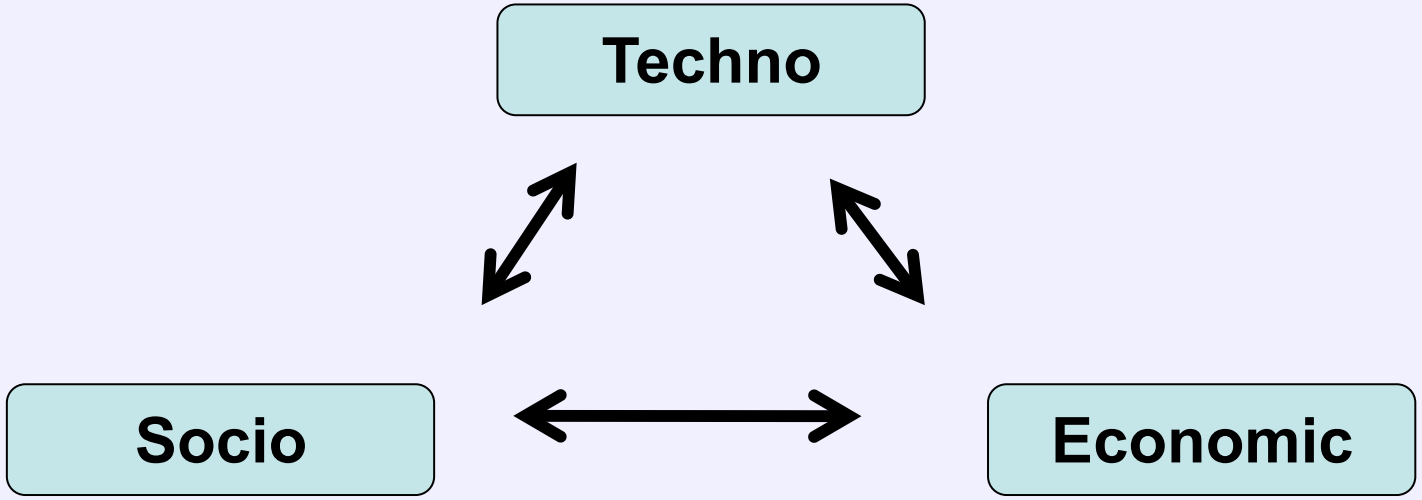
Prologue

The Techno-Socio-Economic Music Revolution

The Revolution

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



P2P Music Sharing
Social Networks

Self-Promotion
Add-Driven
Subscription Based
Expensive Concert Tickets

Proliferation

Supply

myspace.com a place for friends

8M Artists

last.fm the social music revolution

150M Songs

Audacity

1M Downloads/ Month

Demand

173M iPods

iTunes

6B Songs

NULLSOFT SHOUTcast

PANDORA

LIVE 365.COM

31% Americans

Music Search & Discovery Technology

Music Search

Search – retrieving of specific audio content

Common Paradigms:

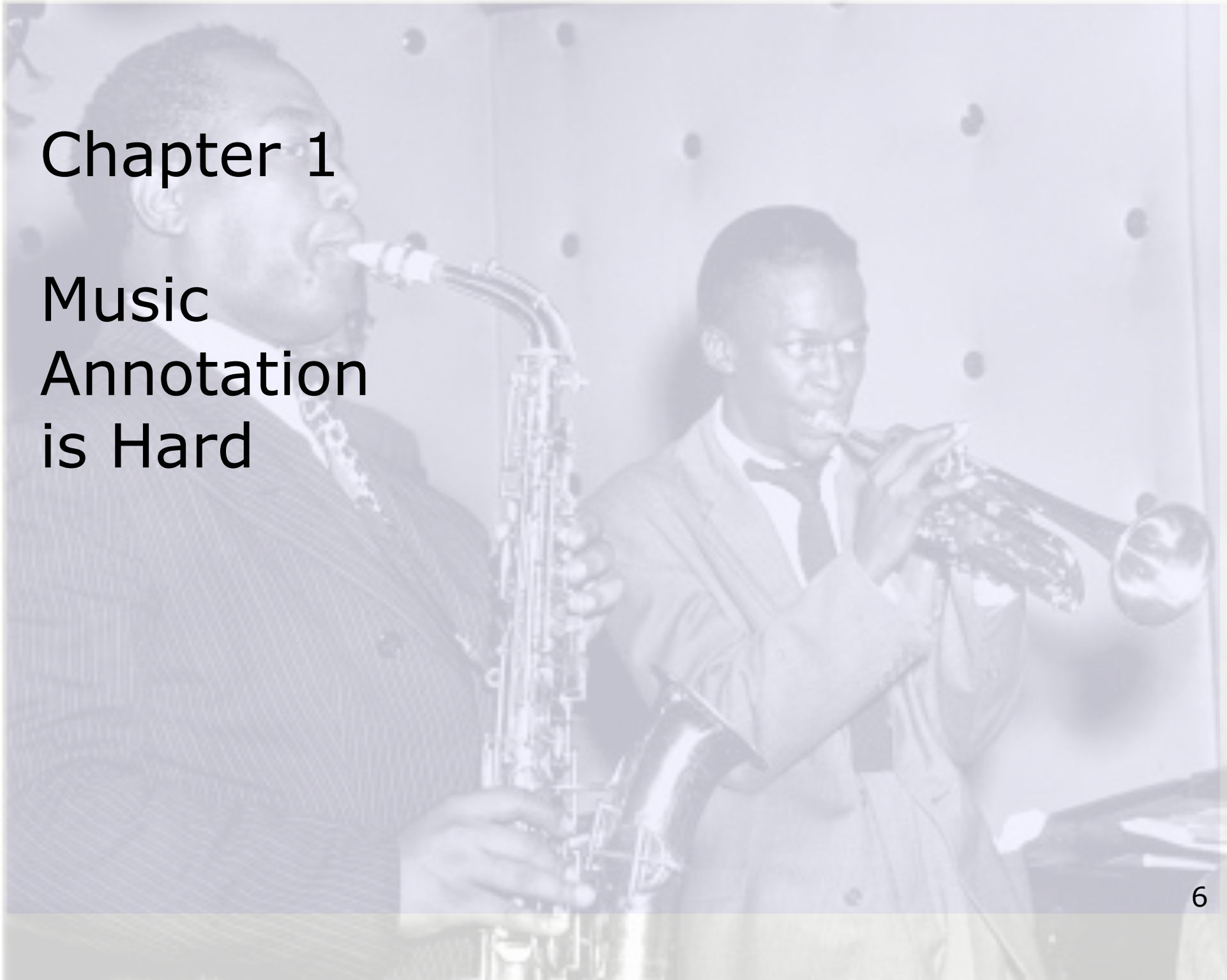
1. Query-by-**Metadata**
2. Query-by-**Performance**
3. Query-by-**Sample**

Music Discovery

Discovery – serendipitous exploration

Common Paradigms:

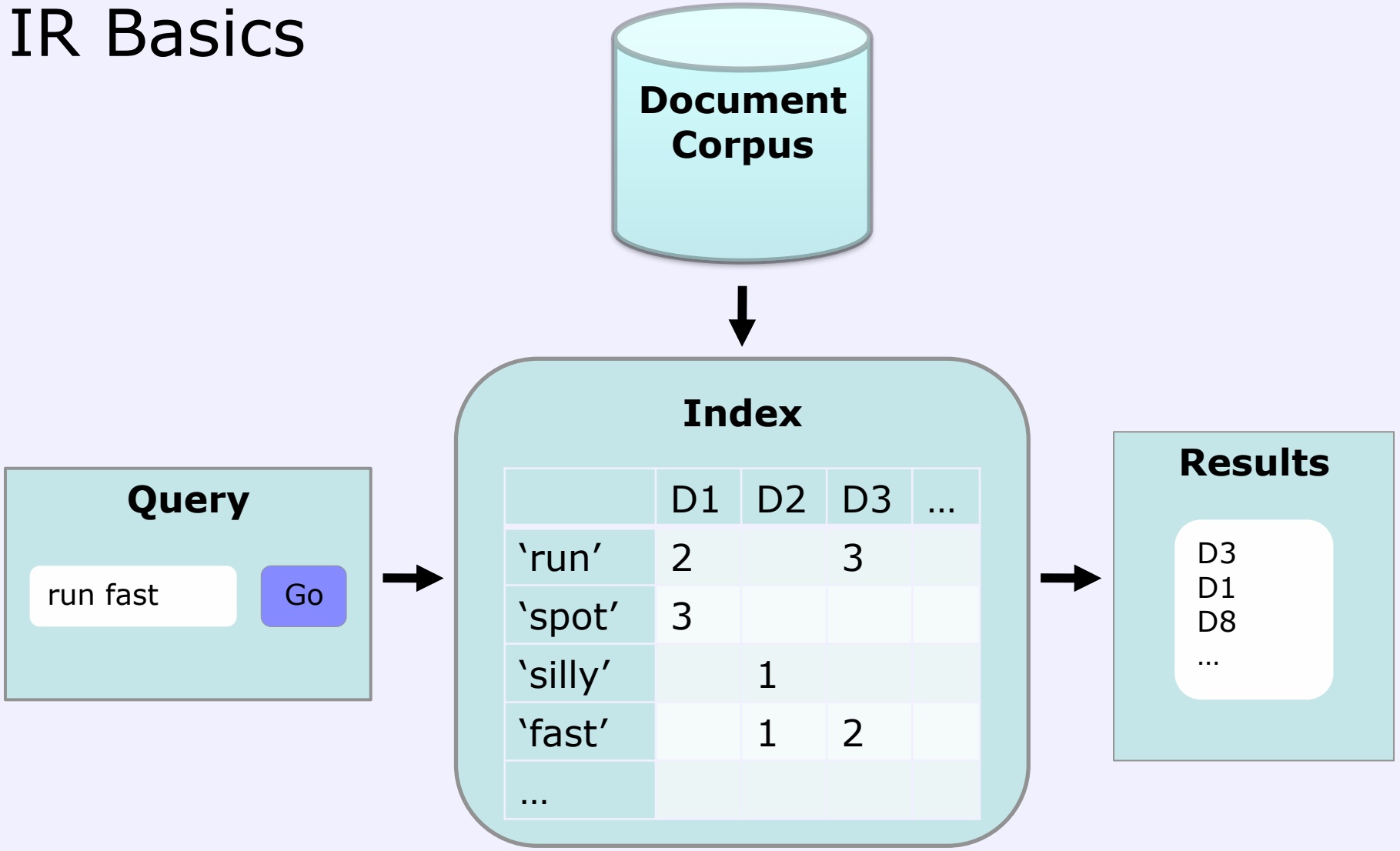
1. Recommendation-by-**Popularity**
2. Browse-by-**Genre**
3. Query-by-**Similarity**
4. Query-by-**Description**



Chapter 1

Music Annotation is Hard

IR Basics



	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 ∝ strength of association

RHCP - *Give it away*

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

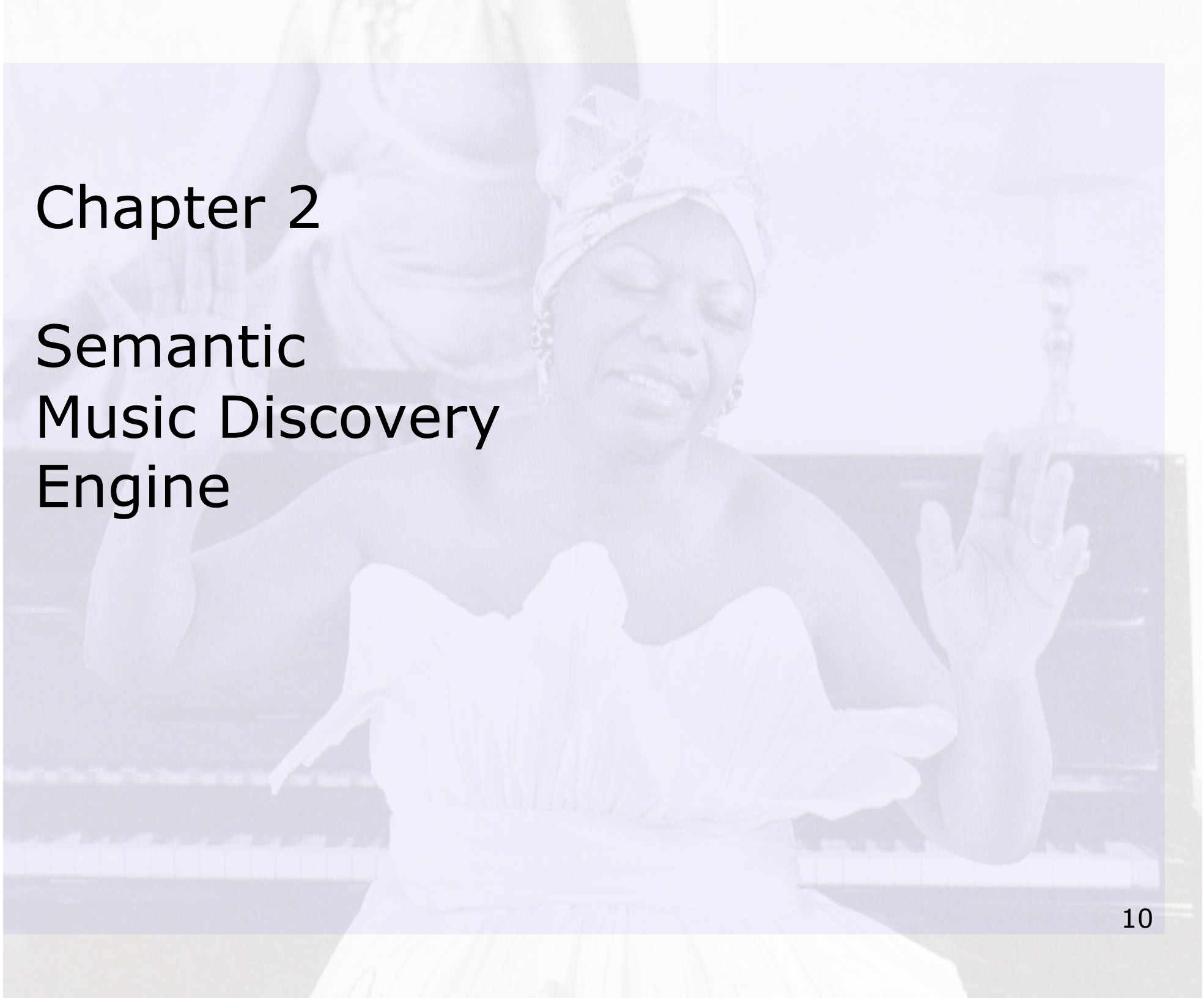
“There just ain’t no truth at all”

Music is **subjective**

- Personal Experience
- Socially-Situated

Approach:

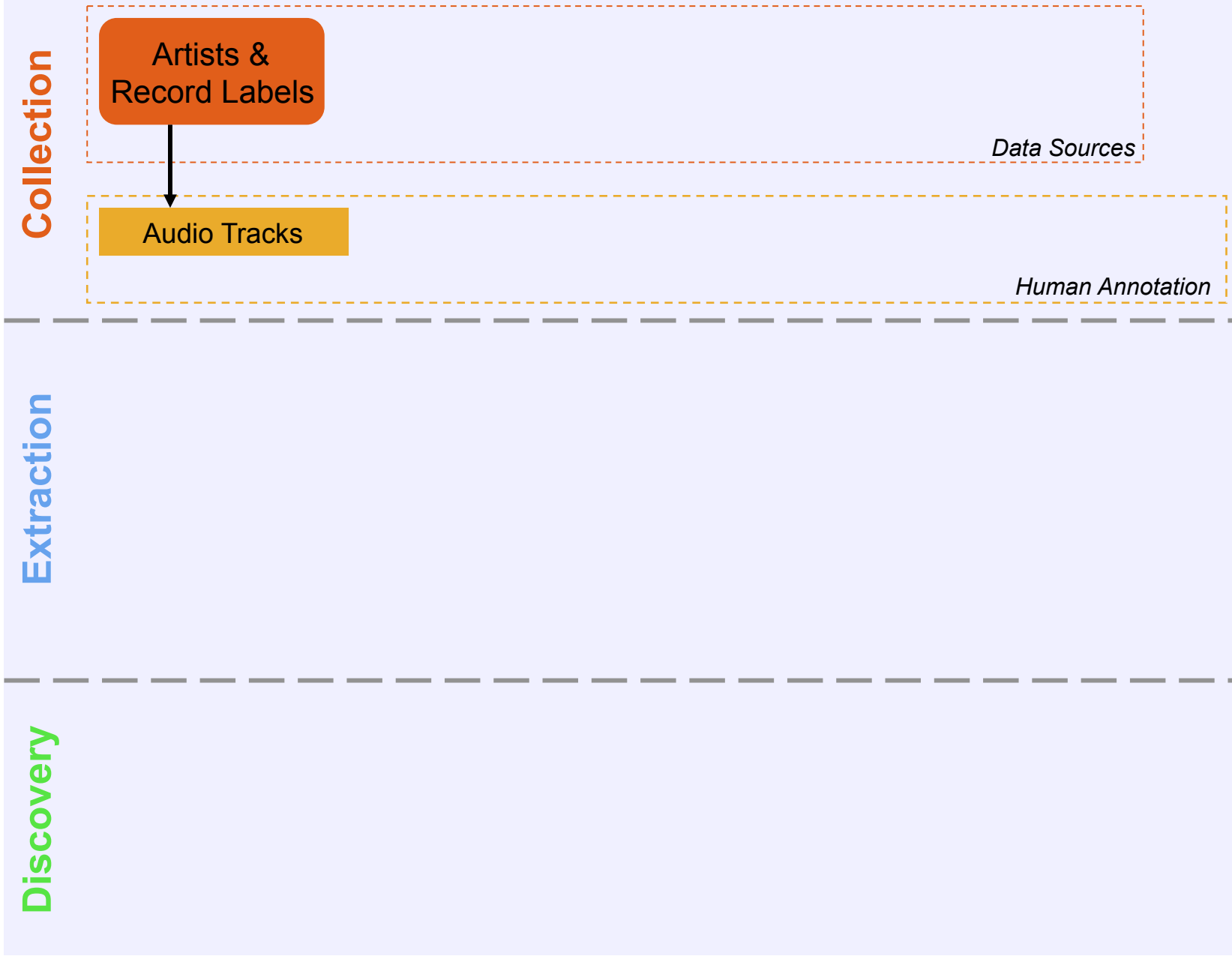
Use multiple sources of music information



Chapter 2

Semantic Music Discovery Engine

Semantic Music Discovery Engine

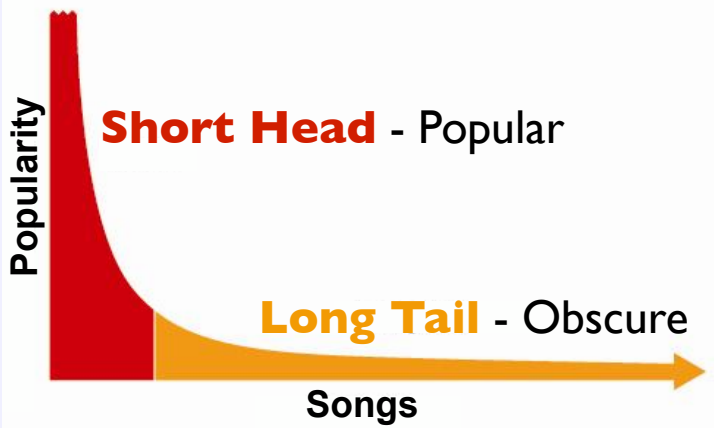


Music Corpora

Last.fm - 150M songs by 16M artists

CAL500 - 500 songs by 500 artist

Long Tail Behavior- Chris Anderson (2004)



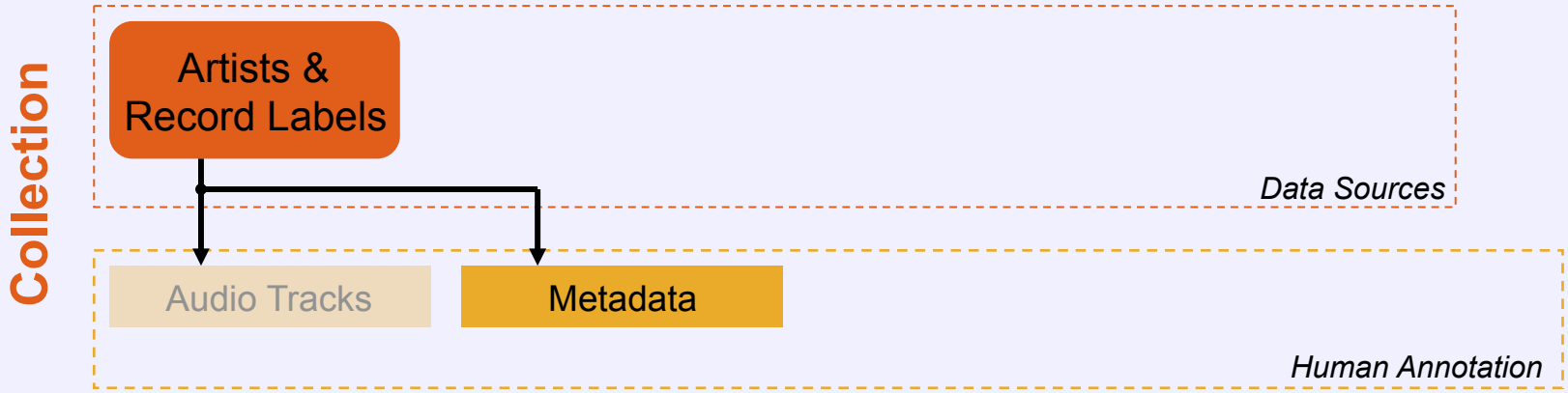
Cold Start Problem

- poorly annotated songs can not be discovered

Popularity Bias

- less popular song tend to be poorly annotated

Semantic Music Discovery Engine



Extraction

Discovery

Metadata

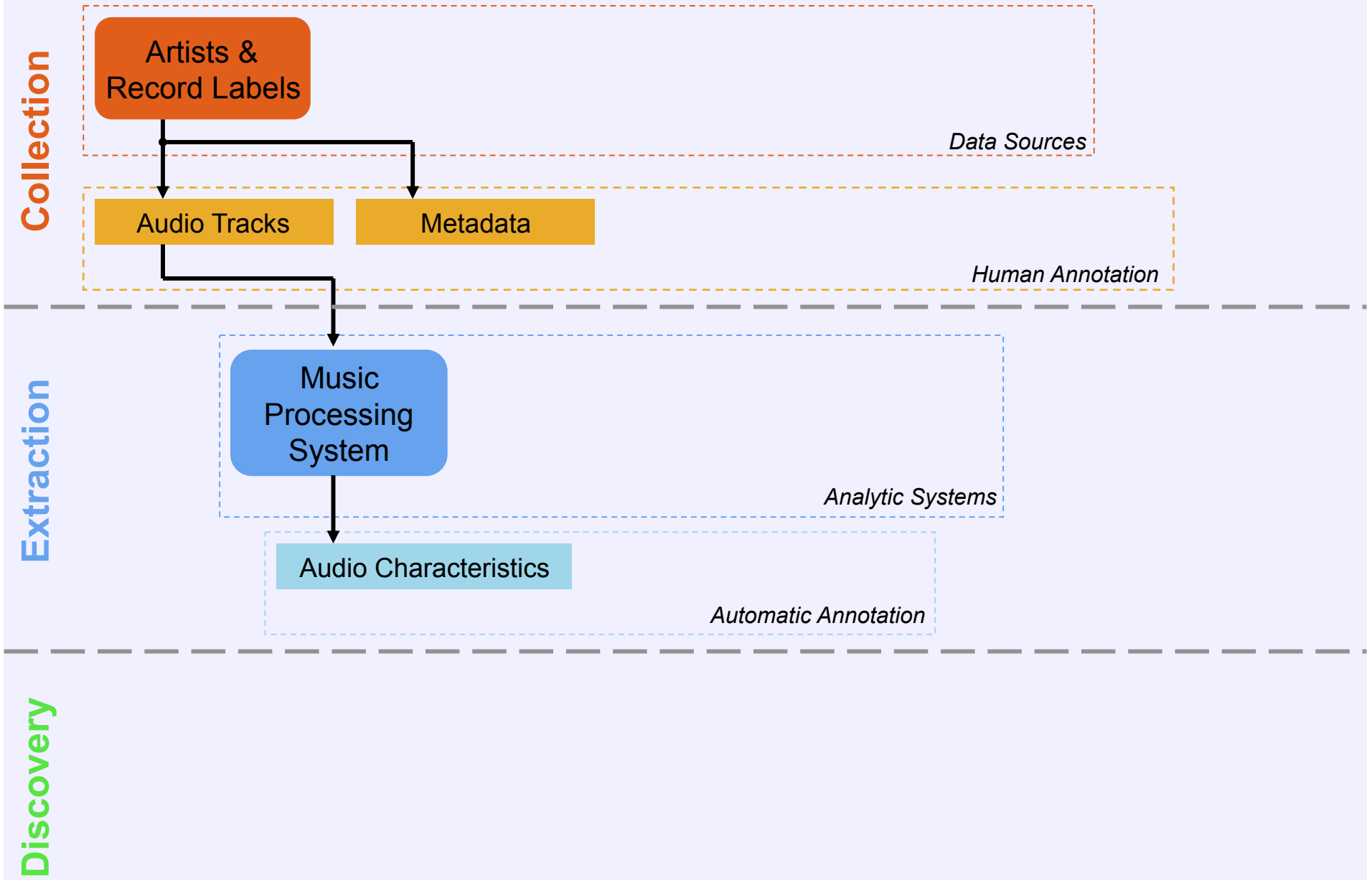
Factual information about music

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

Heterogeneous data

- strings, numbers, images, graphs

Semantic Music Discovery Engine



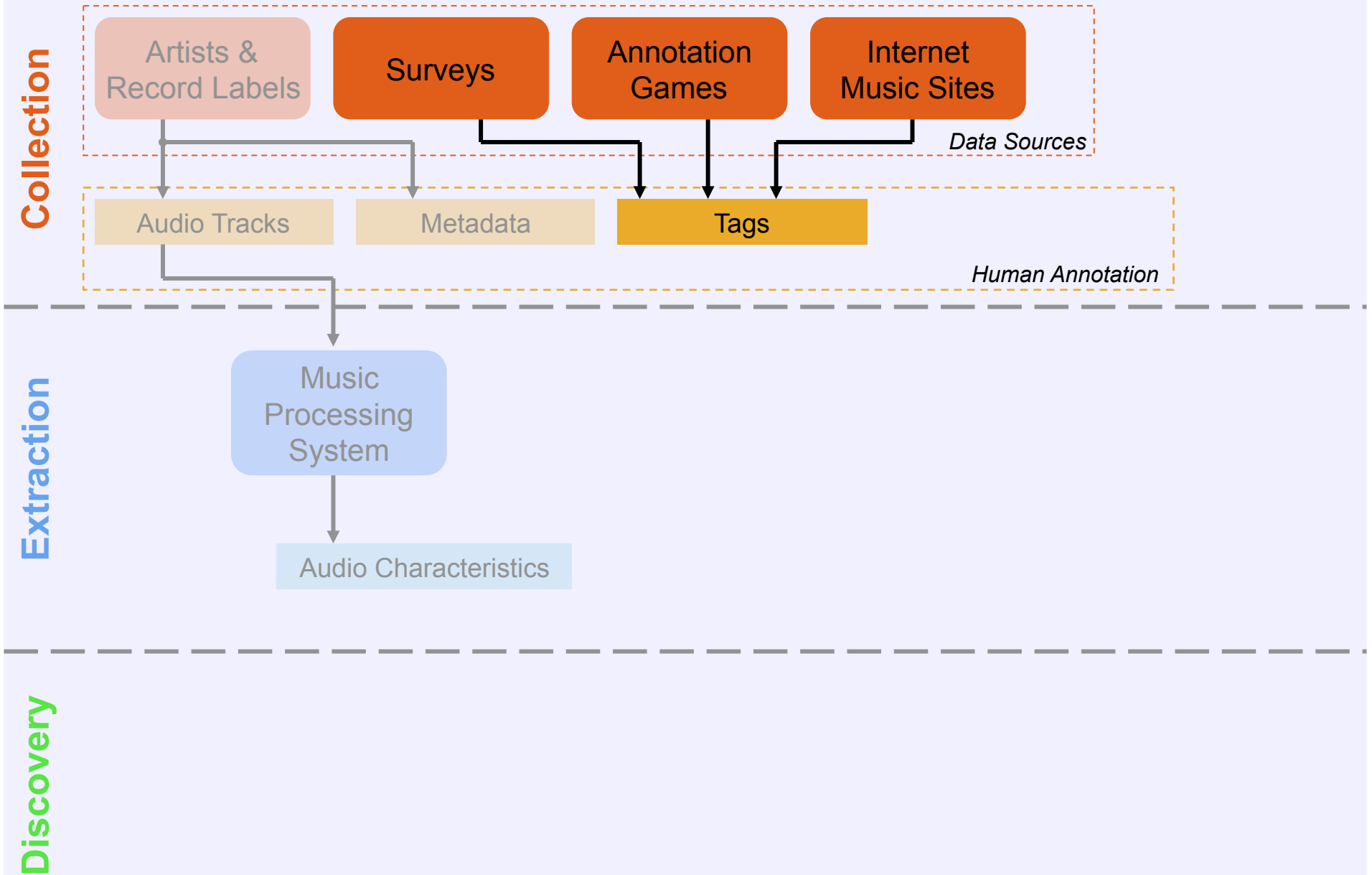
Music Processing Systems

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	Uncerta
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	○	○

Human Annotations

Conducting a survey

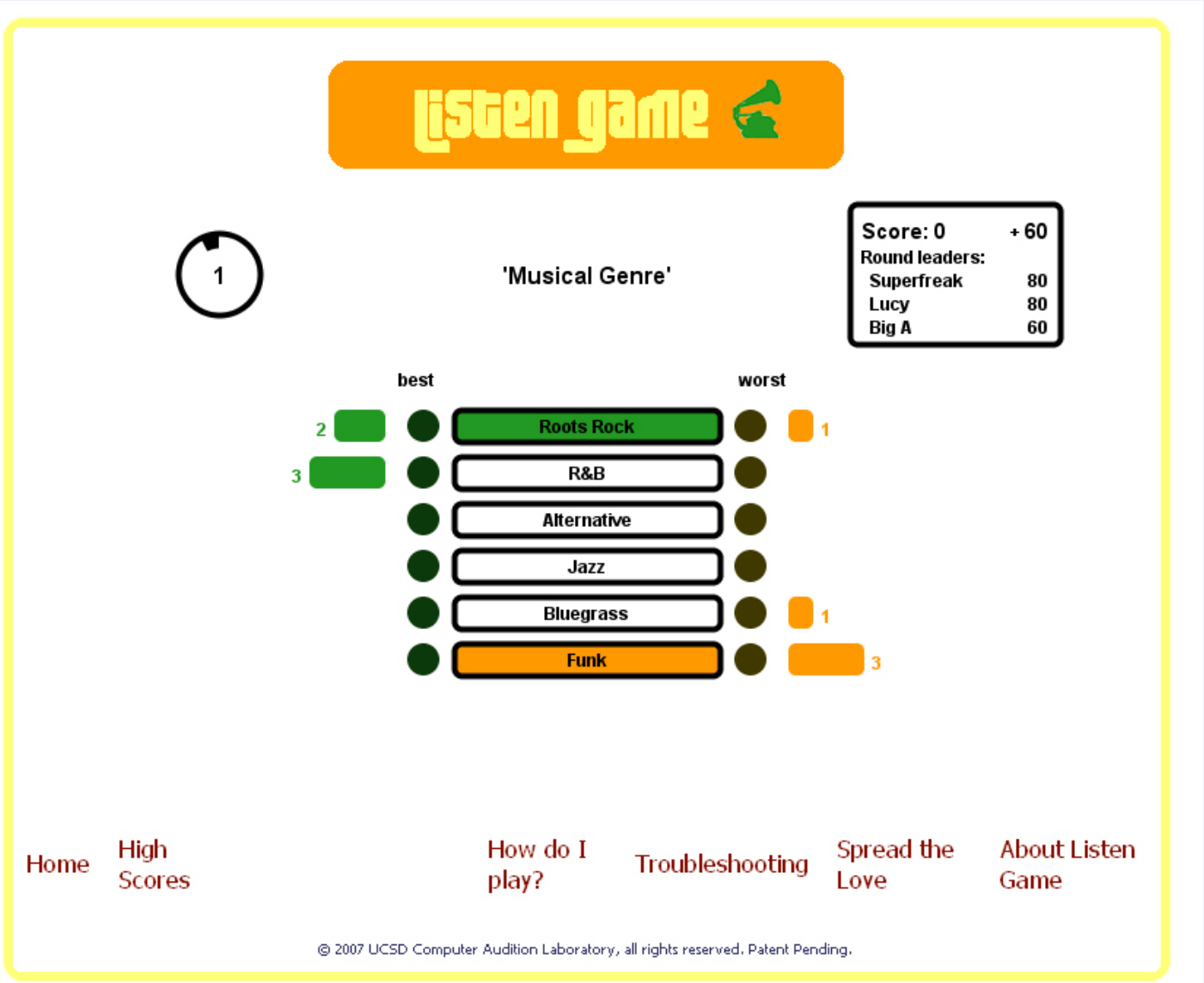
- ✓ Reliable, Precise, Tailored to Application
- X Expensive, Laborious, Not Scalable

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


listen game
1
'Musical Genre'
Score: 0 + 60
Round leaders:
Superfreak 80
Lucy 80
Big A 60
best
worst
2
Roots Rock
1
3
R&B
Alternative
Jazz
Bluegrass
1
Funk
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

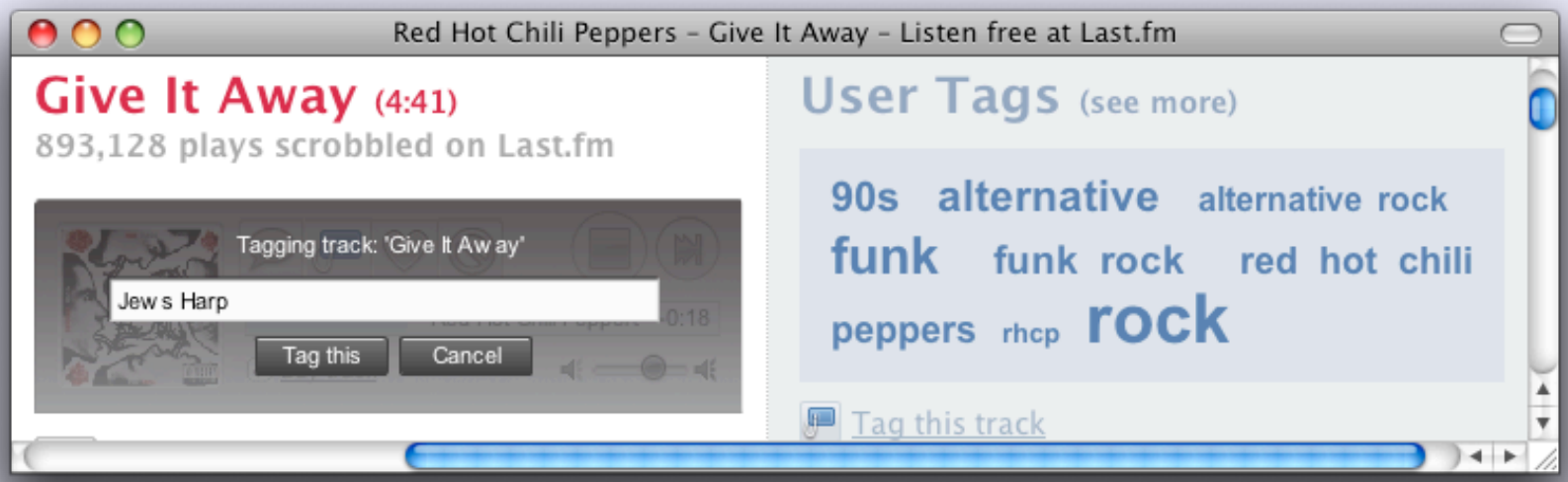
- ✓ Reliable, Precise, Tailored to Application
- X Expensive, Laborious, Not Scalable

Annotation Game

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

Music Web Sites

1. Social Tagging Site
   - Users annotate music with tags
   - Last.fm - 960K distinct tags



http://www.last.fm/music/Red+Hot+Chili+Peppers/_/Give+It+Away

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

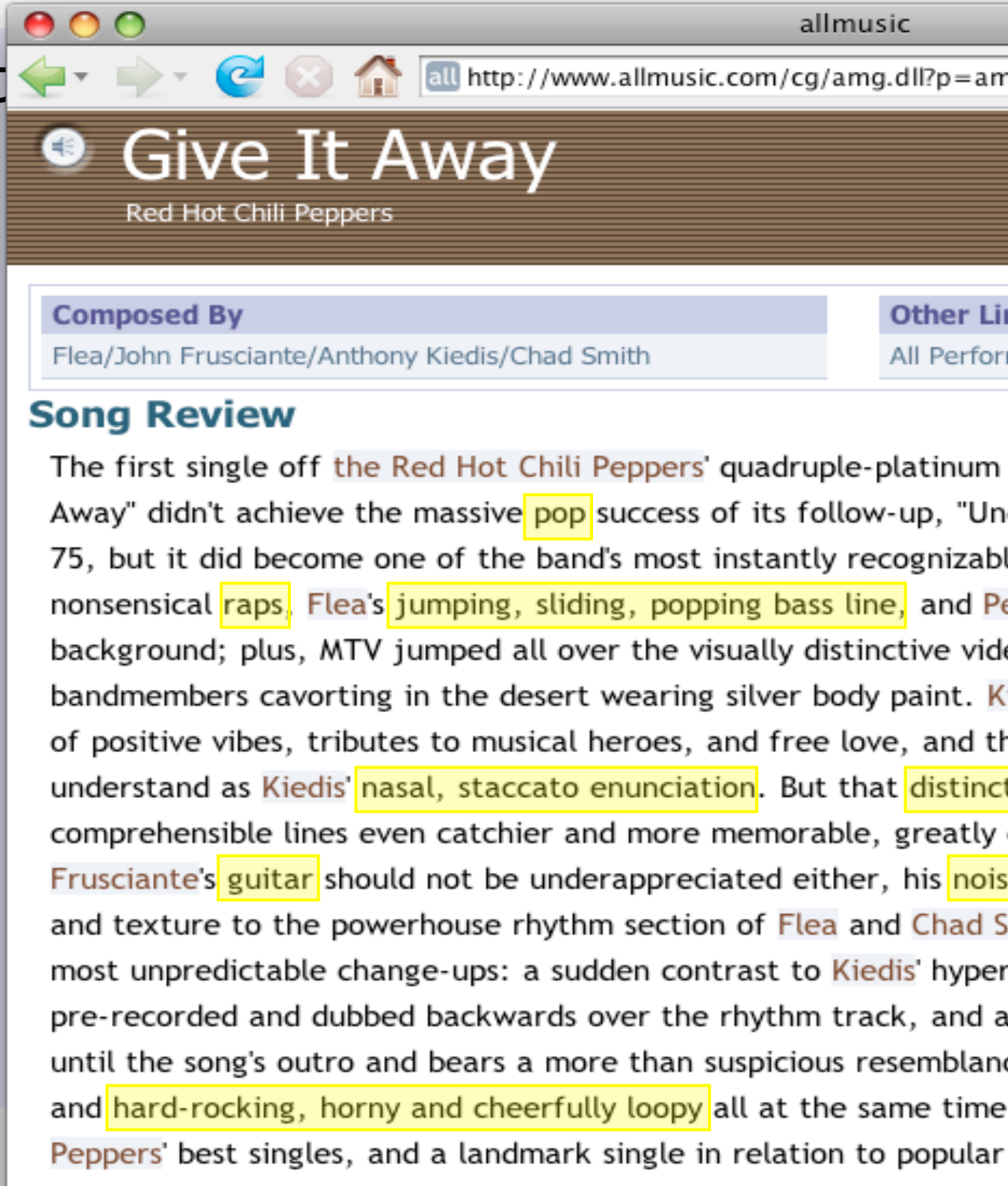
Jew's-harp

Adjective:

Hard-rocking (2)

Noisy

Scratchy



Collecting an Annotated Music Corpus

Survey

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

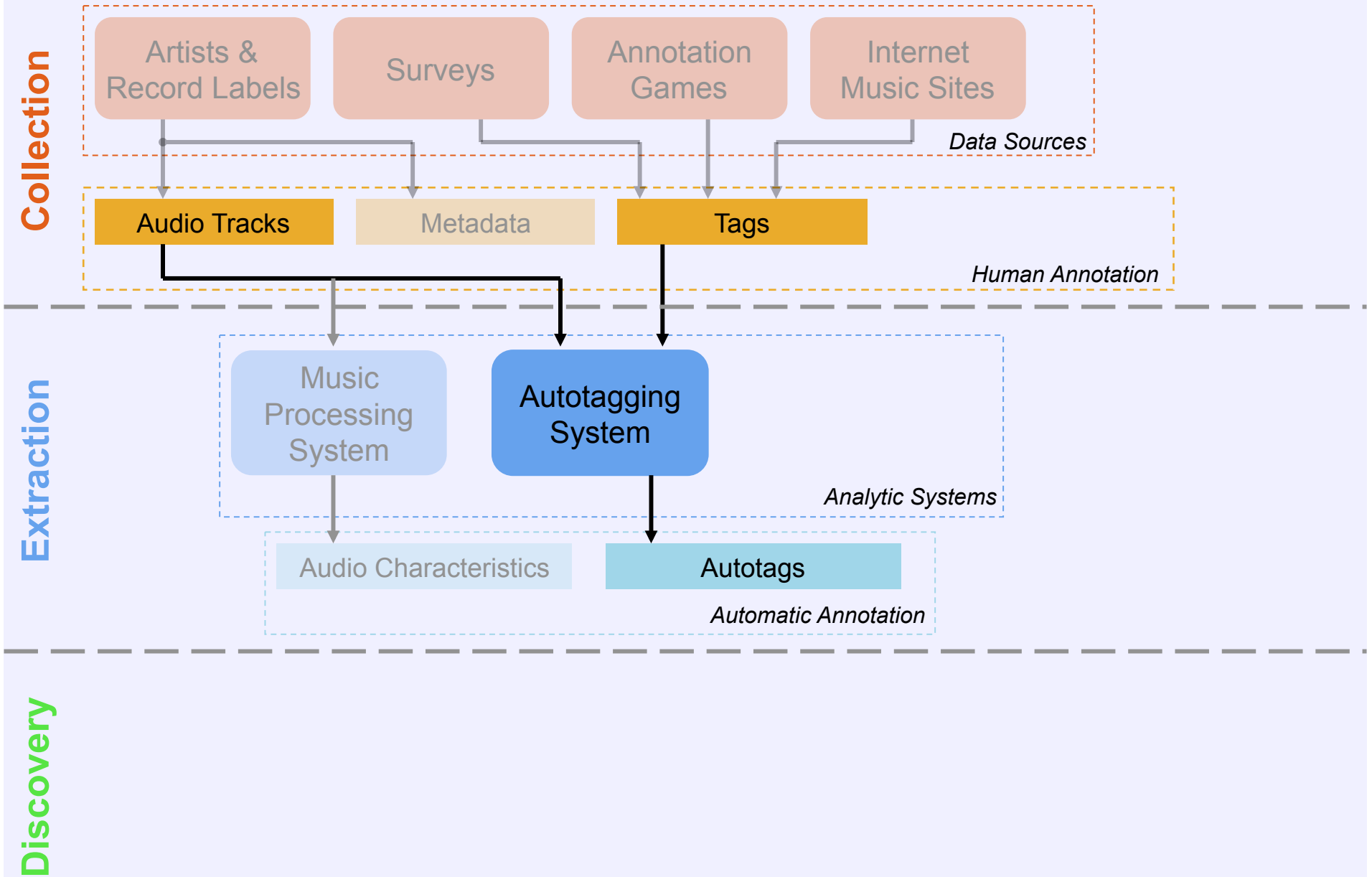
Annotation Game

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

Music Web Sites

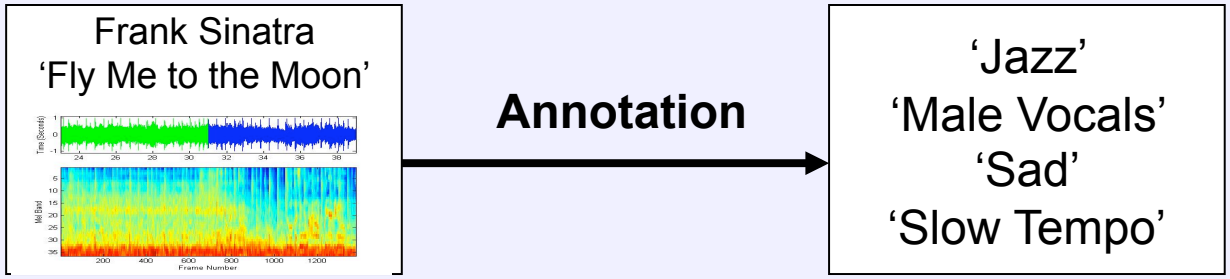
- ✓ Cheap, Annotations for short-head
- X Noisy, long-tail is poorly represented

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

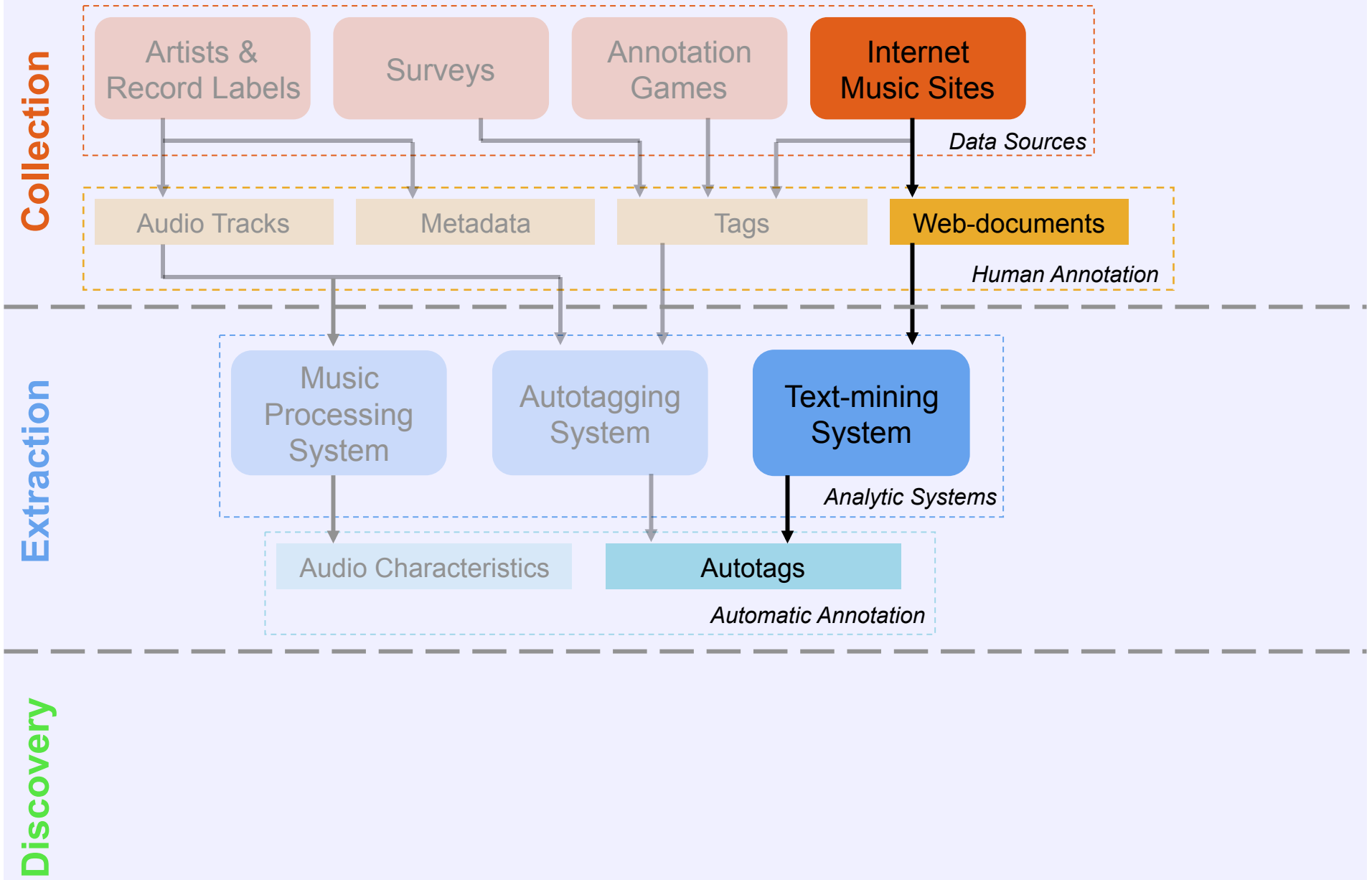
Dr. Dre (feat. Snoop Dogg) - Nuthin' but a 'G' thang

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

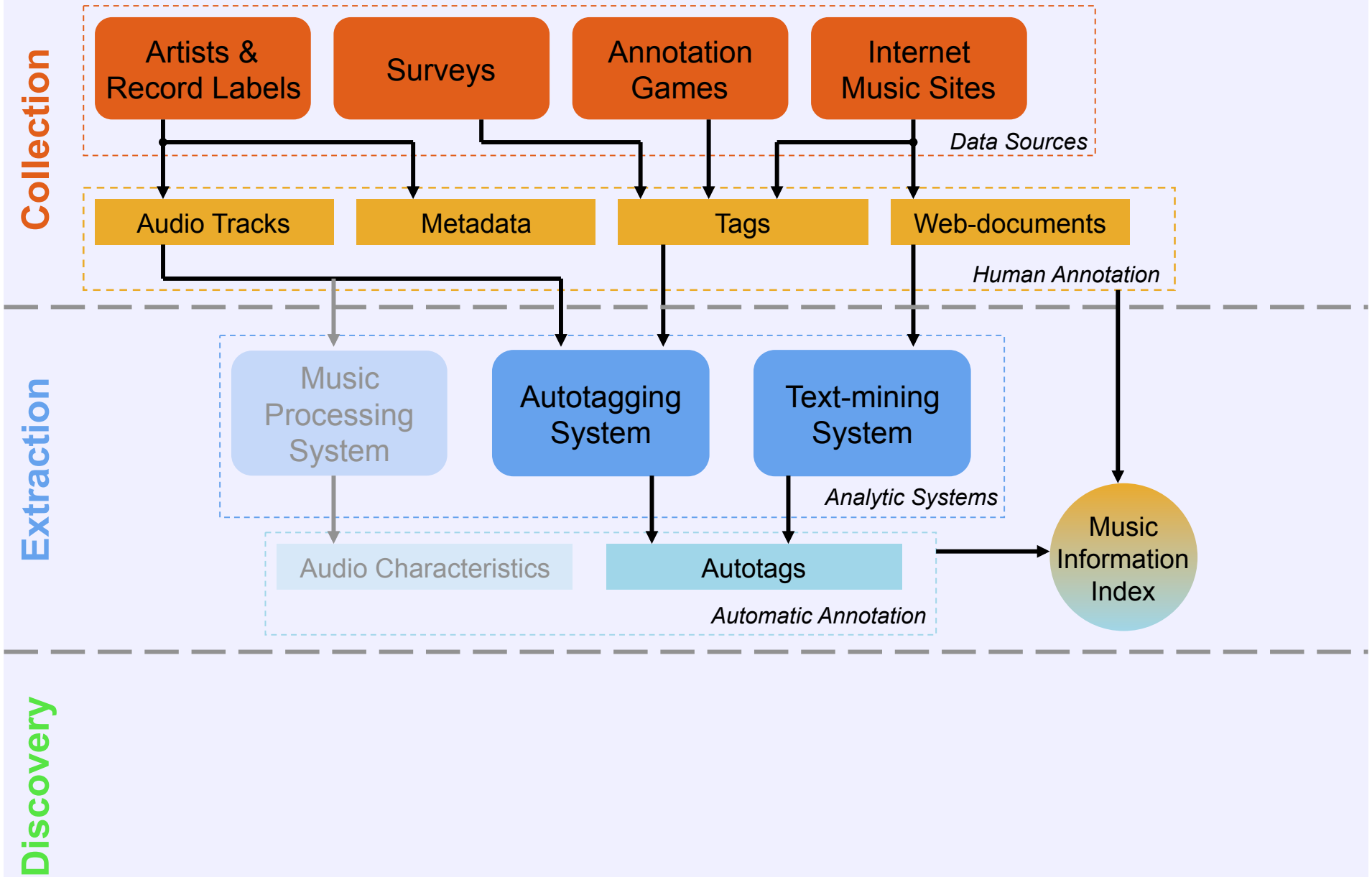
For each song, use a **search engine** to **retrieve web pages**:

- 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



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)

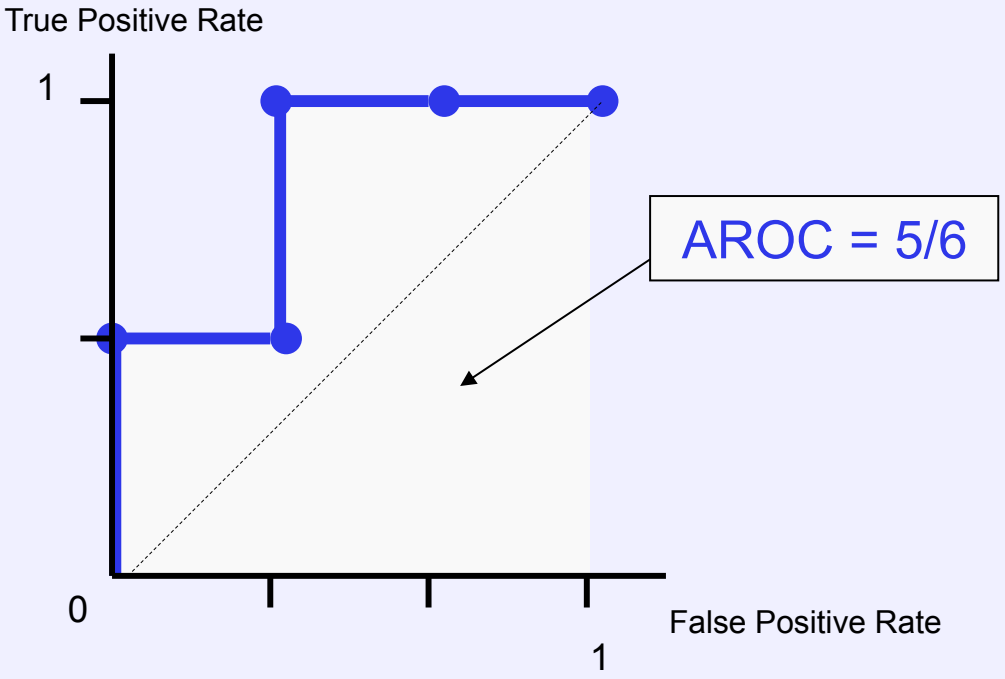
Calculate mean AROC

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

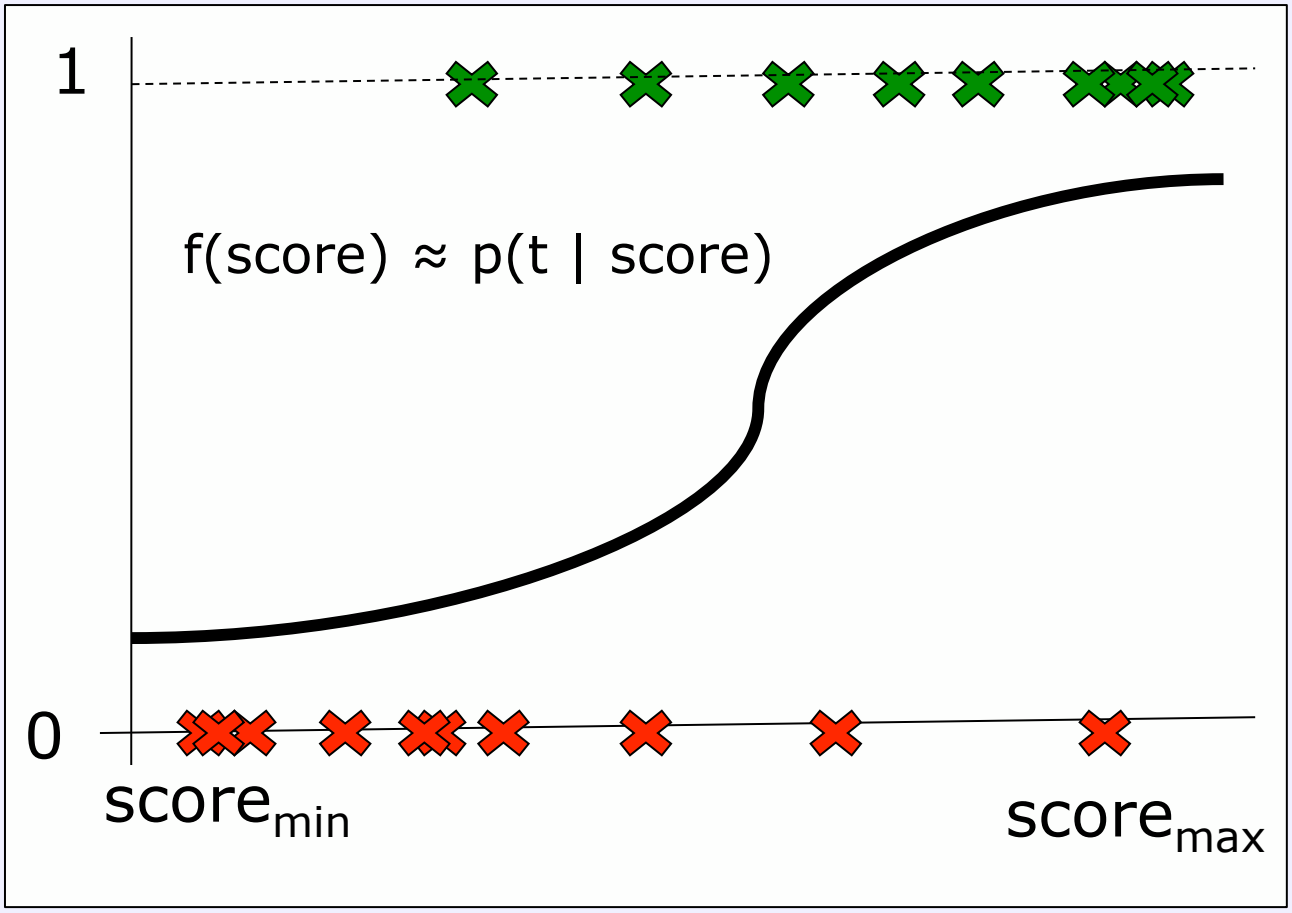
Approach	Songs	AROC
Social Tags	**CAL500**	**0.62**
	Long Tail	0.54
Game	**CAL500**	**0.65**
	Long Tail	*
Web Autotags	**CAL500**	**0.66**
	Long Tail	0.56
Audio Autotags	**CAL500**	**0.69**
	Long Tail	0.70

Combining Data Sources

Approaches

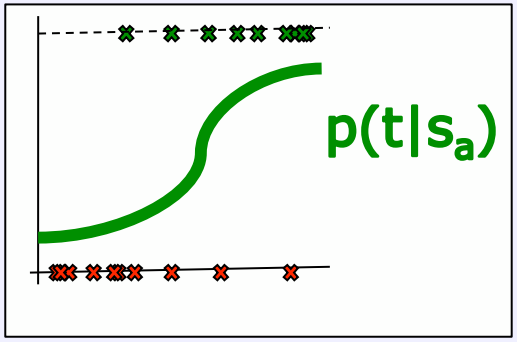
1. **Calibrated Score Averaging** - [Zadrozny 02]
2. **RankBoost** - [Freund 03]
3. **Kernel Combination SVM**- [Lanckriet 04]

1. Calibrated Score Averaging

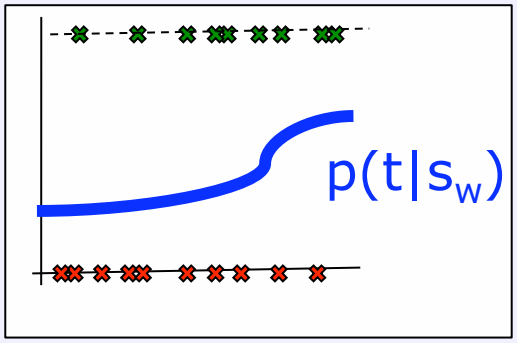


1. Calibrated Score Averaging

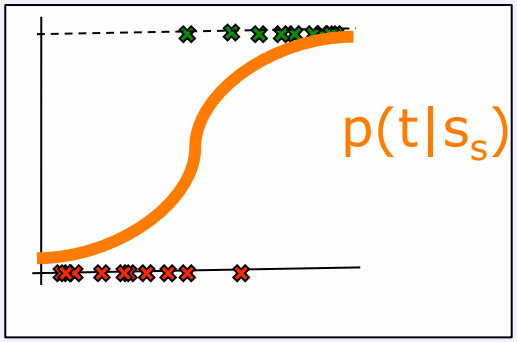
Autotag Score



Web Document Score



Social Tag Score



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

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

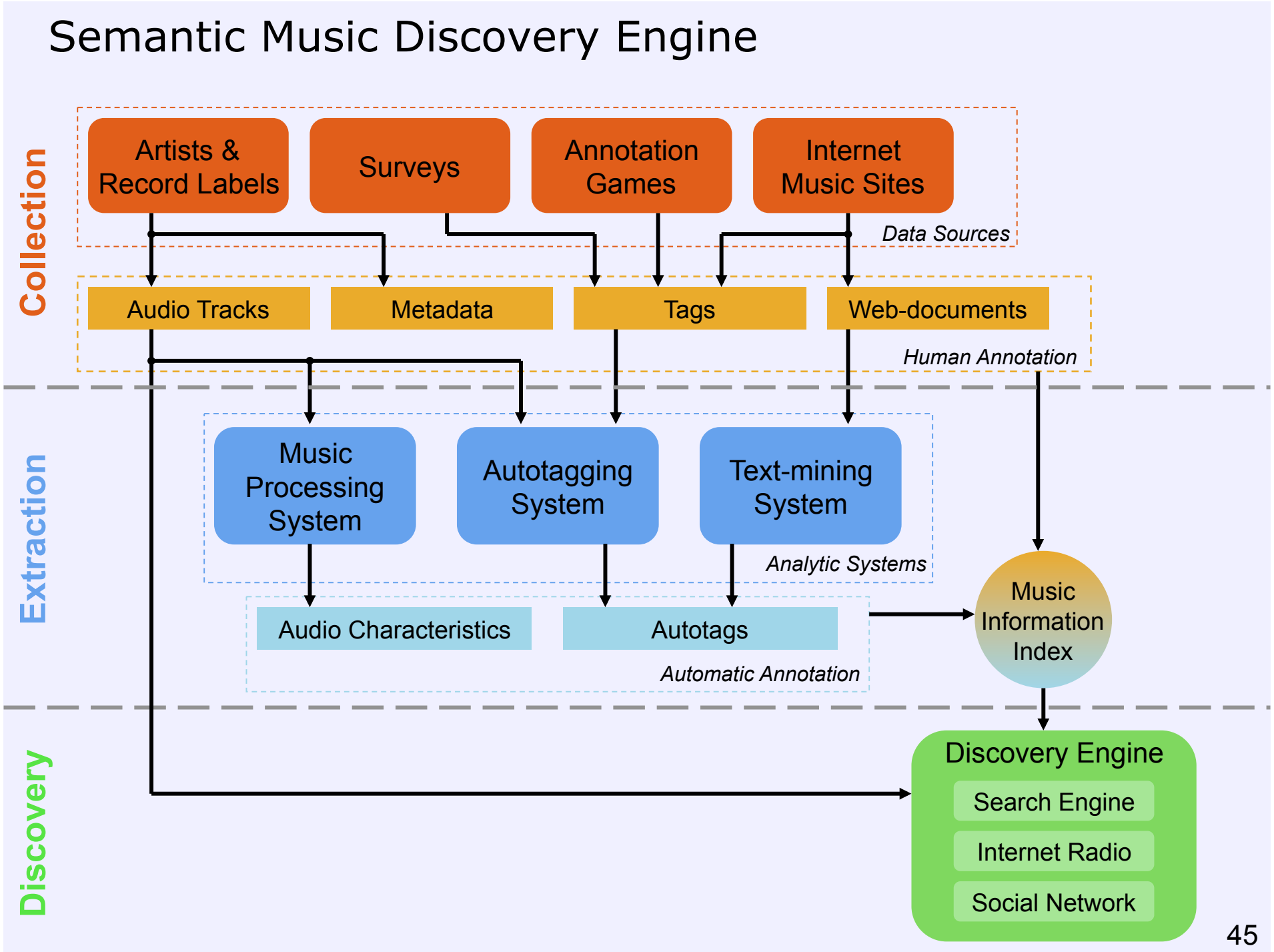
2. Learn an optimal **linear combination of the kernel matrices** using convex optimization
   - Produces single kernel matrix for SVM

3. Rank songs based on score from SVM
   - **positive distance** from **separating hyperplane**.

Combining Tags

Approach	AROC
Audio Autotags	**0.69**
Calibrated Score Averaging	**0.75**
Rankboost	**0.75**
Kernel Combo SVM	**0.74**

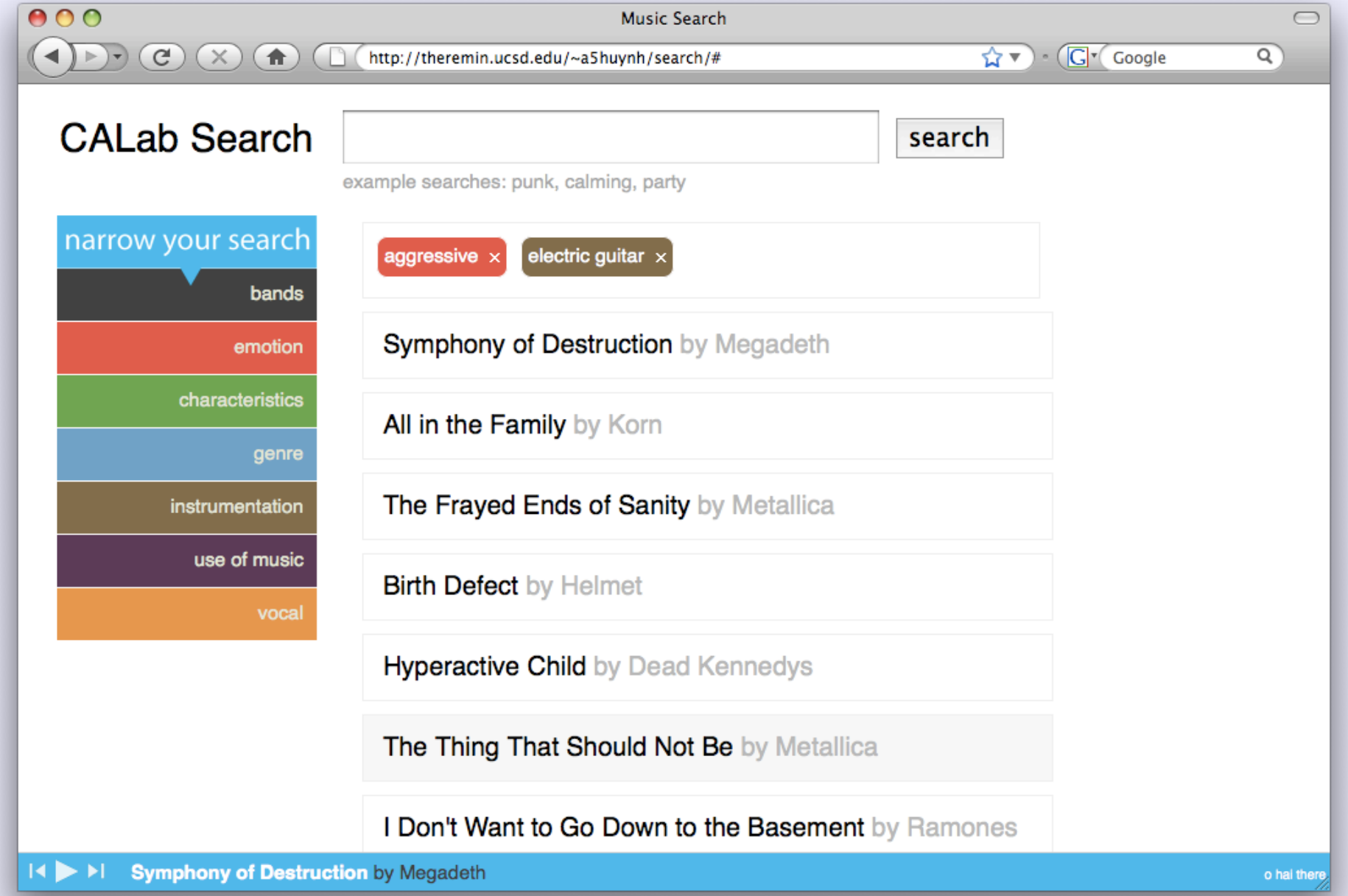
Semantic Music Discovery Engine



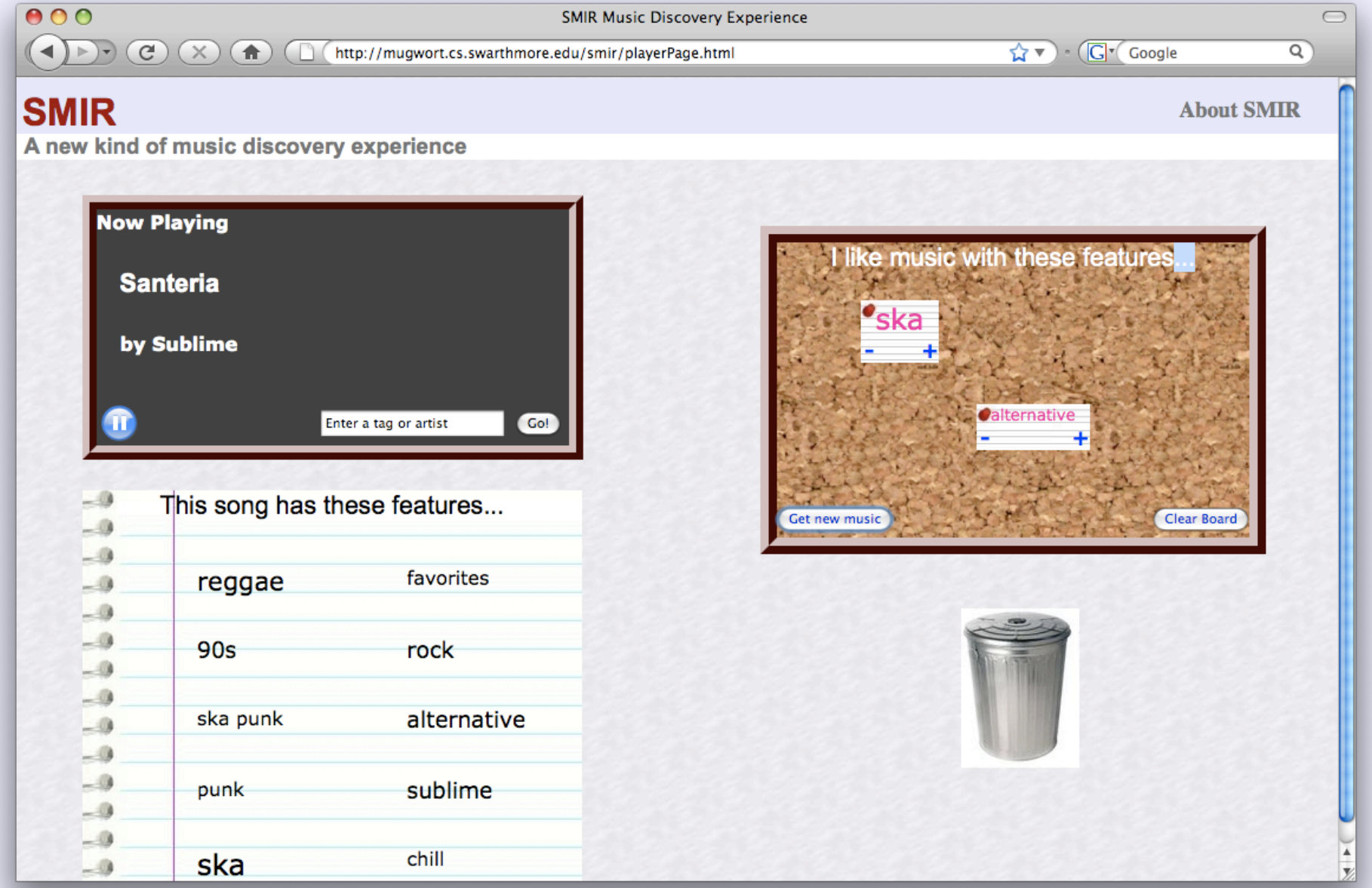
Chapter 3

The User Experience

CALab Search



SMIR – Swarthmore Music IR



Herd-it

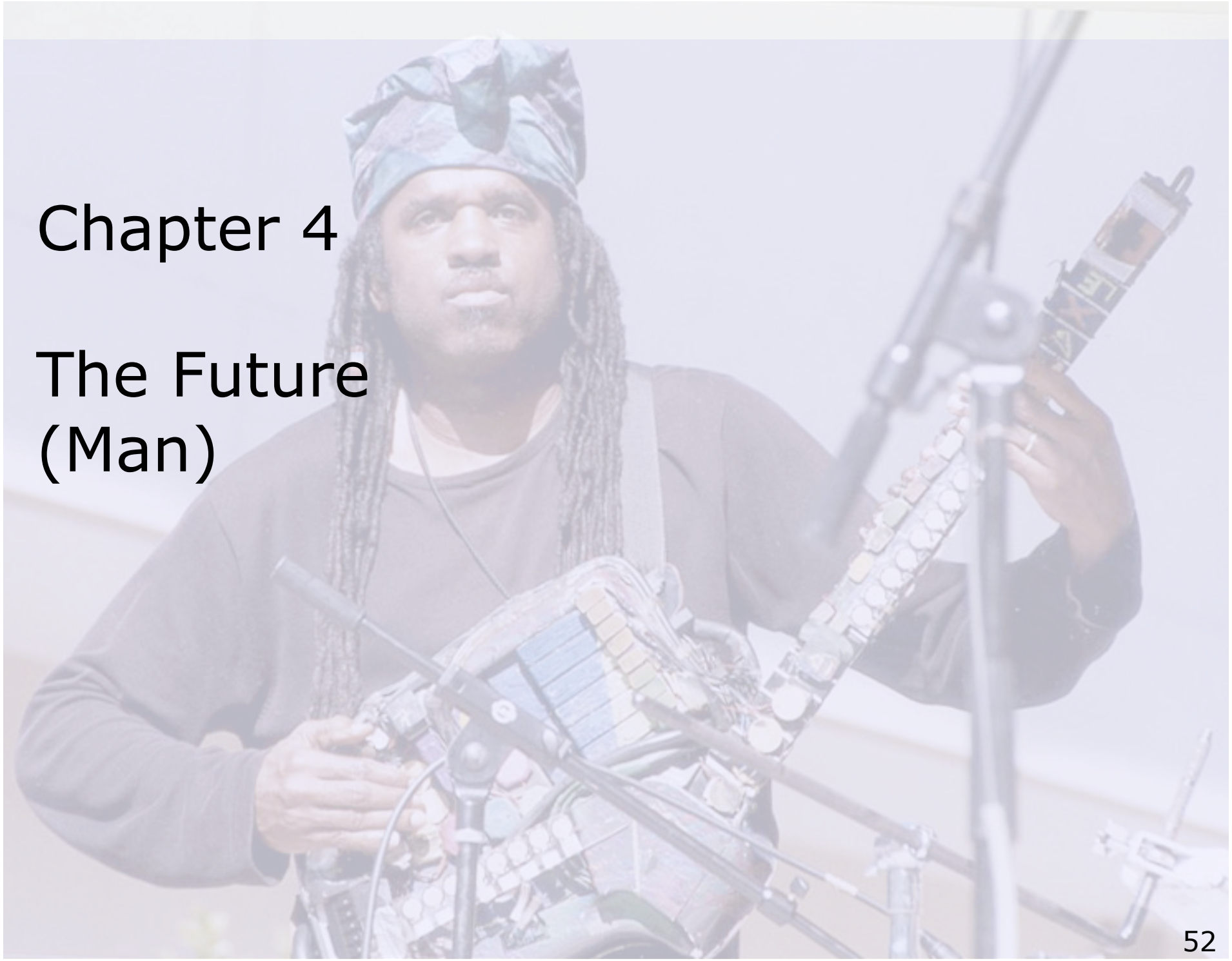

Herd It on Facebook
http://apps.facebook.com/herd-it/?ref=ts
Google
facebook
Home
Profile
Friends
Inbox
Douglas Turnbull
Settings
Logout
10-Question Survey
Report a bug
Herd It
Hi Douglas
Start by choosing a musical style
Everything
8
players online
Electronic
8
players online
Pop
3
players online
Rock
8
players online
Blues
5
players online
Hip-Hop
7
players online
Are you ready to rock out with the world?
Challenge Your Friends!
How do I play Herd It?
Play an interactive demo:
play demo
or
View quick instructions:
instructions
Top Scoring Friends
Leaderboard
Musical Profile
Antoni
466
Rock
Luke
459
Rock
Oscar
447
Rock
Damien
439
Rock
Challenge Your Friends to Play Herd It!
Applications
Bookmark Herd It
Online Friends (12)

Herd-it


Herd It on Facebook
http://apps.facebook.com/herd-it/?ref=ts
Google
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
electronic
irreverent
brash
thuggish
sardonic
theatrical
100%
50%
0%
agree-O-meter
CHAT
Applications
Bookmark Herd It
Online Friends (12)

Herd-it


Herd It on Facebook
http://apps.facebook.com/herd-it/?ref=ts
Google
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
Chris
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]

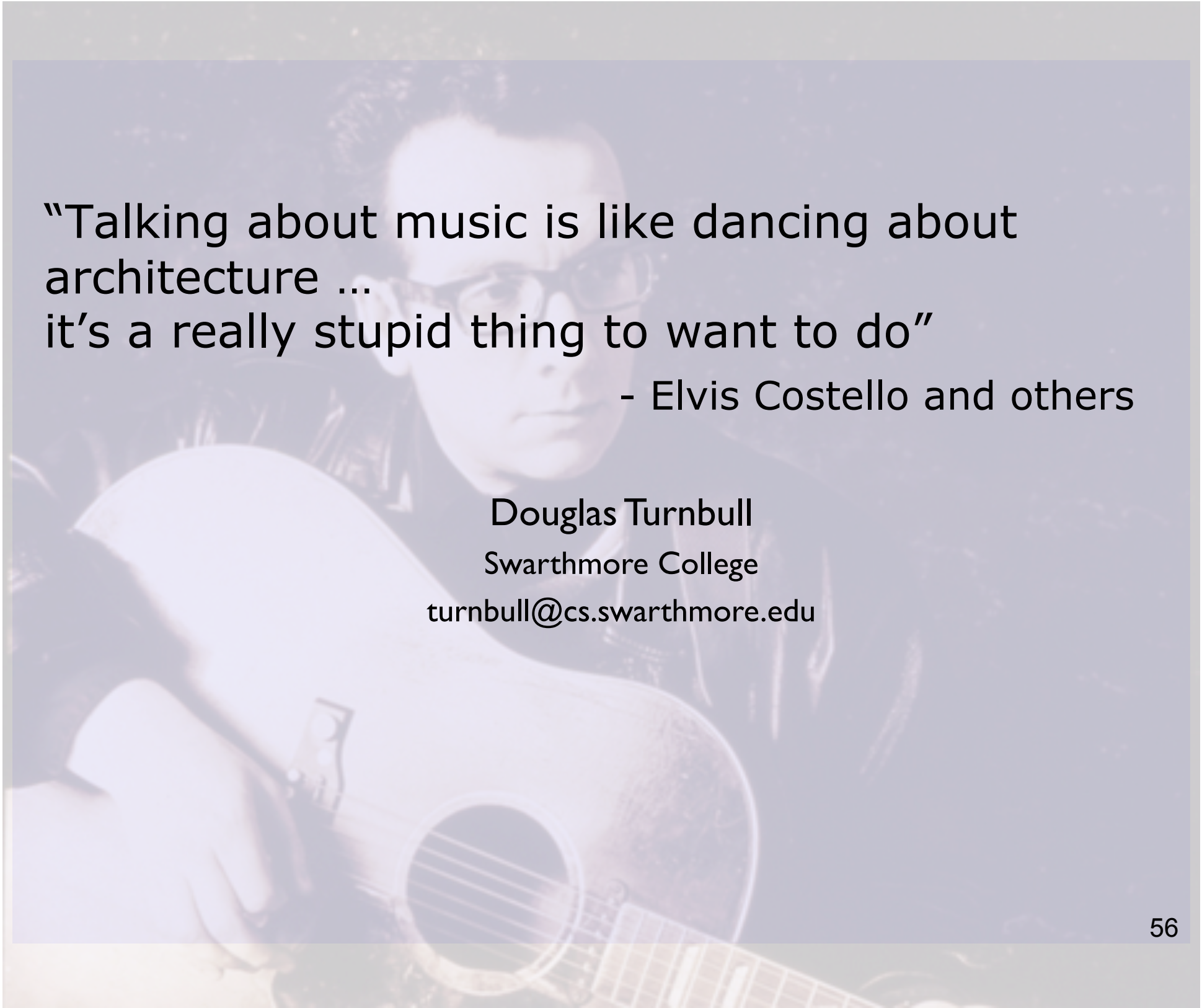
Music Annotation Games [ISMIR 07a]

Related:

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

Tag Vocabulary Selection with Sparse CCA [ISMIR 07]

Supervised Music Boundary Detection [ISMIR 07]



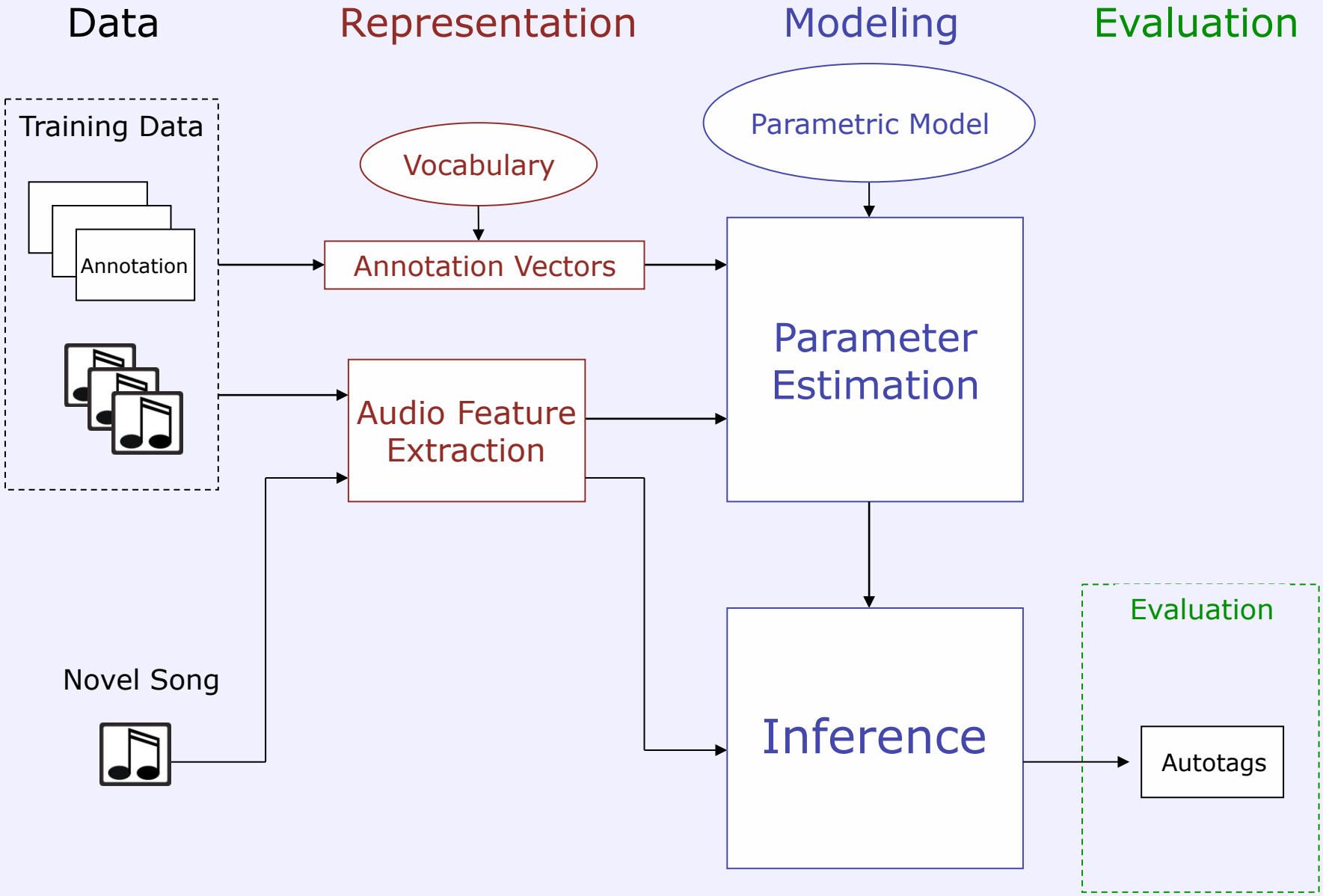
"Talking about music is like dancing about architecture …
it's a really stupid thing to want to do"

- Elvis Costello and others

Douglas Turnbull
Swarthmore College
turnbull@cs.swarthmore.edu

Questions?

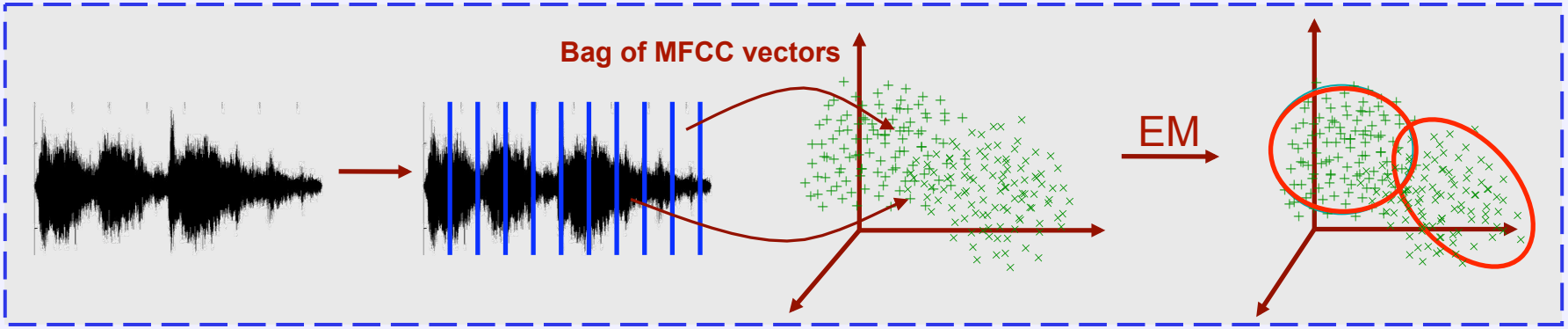
System Overview



Modeling a Song

Algorithm

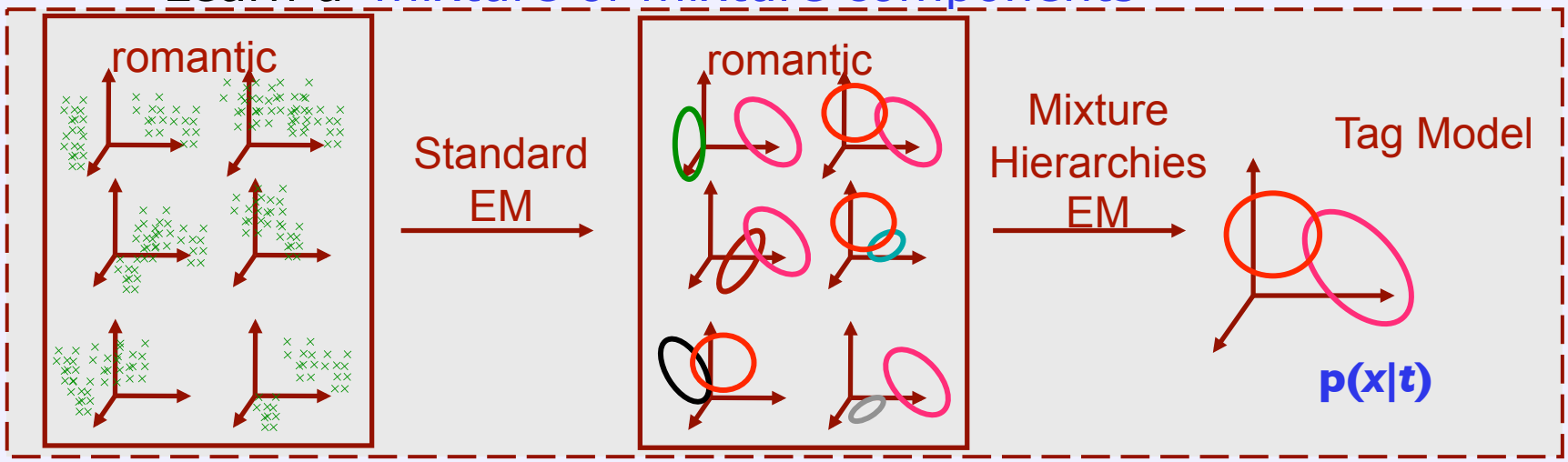
1. Segment audio signals
2. Extract short-time feature vectors
3. Estimate GMM with EM algorithm



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 → better density estimate

Annotation

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

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

Assuming

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

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

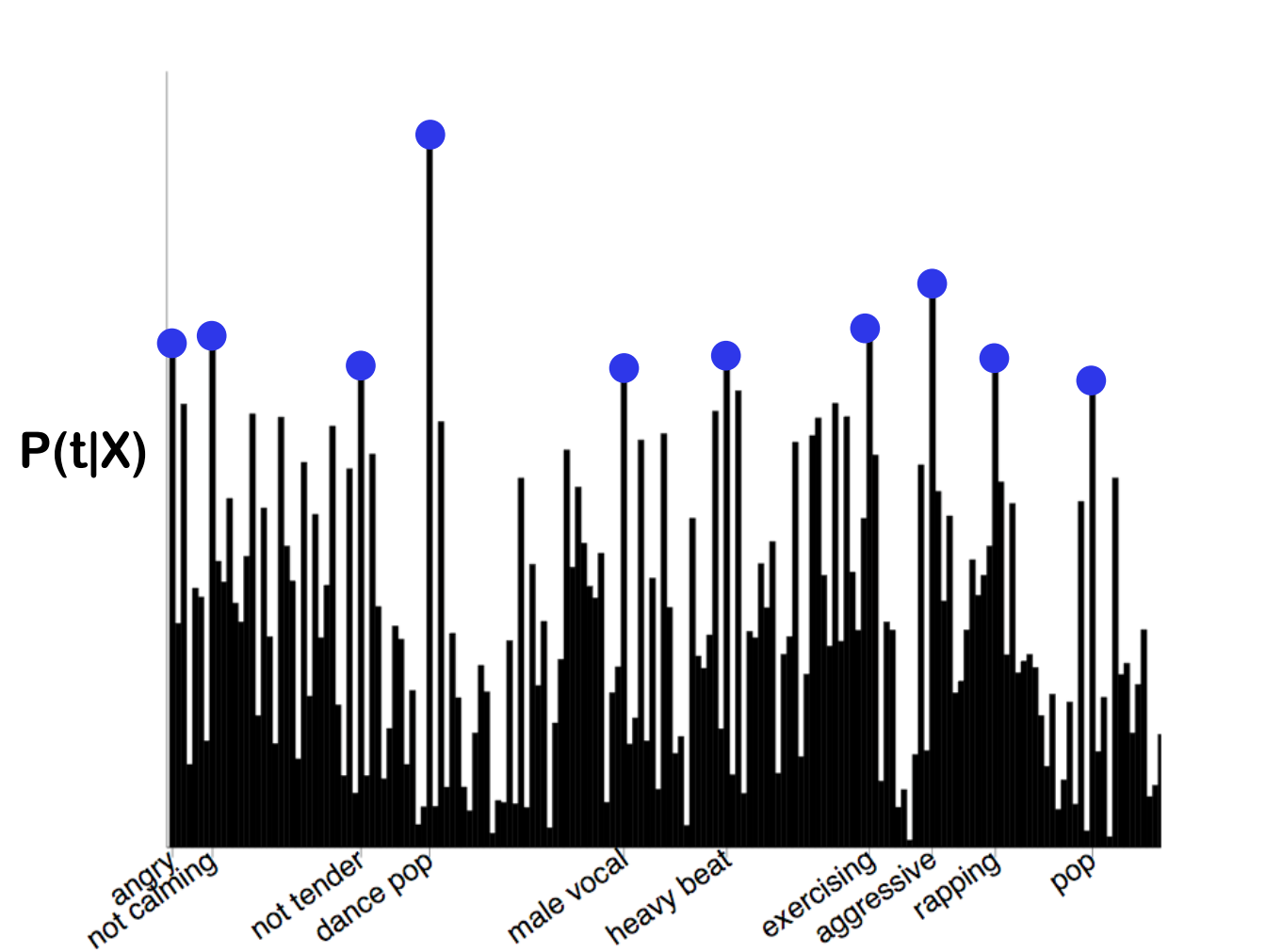
Semantic Multinomial:

- P(*t*|**X**)'s →multinomial distribution over the tag vocabulary

Annotation: peaks of multinomial

Annotation

Semantic Multinomial for “Give it Away” by the Red Hot Chili



Retrieval

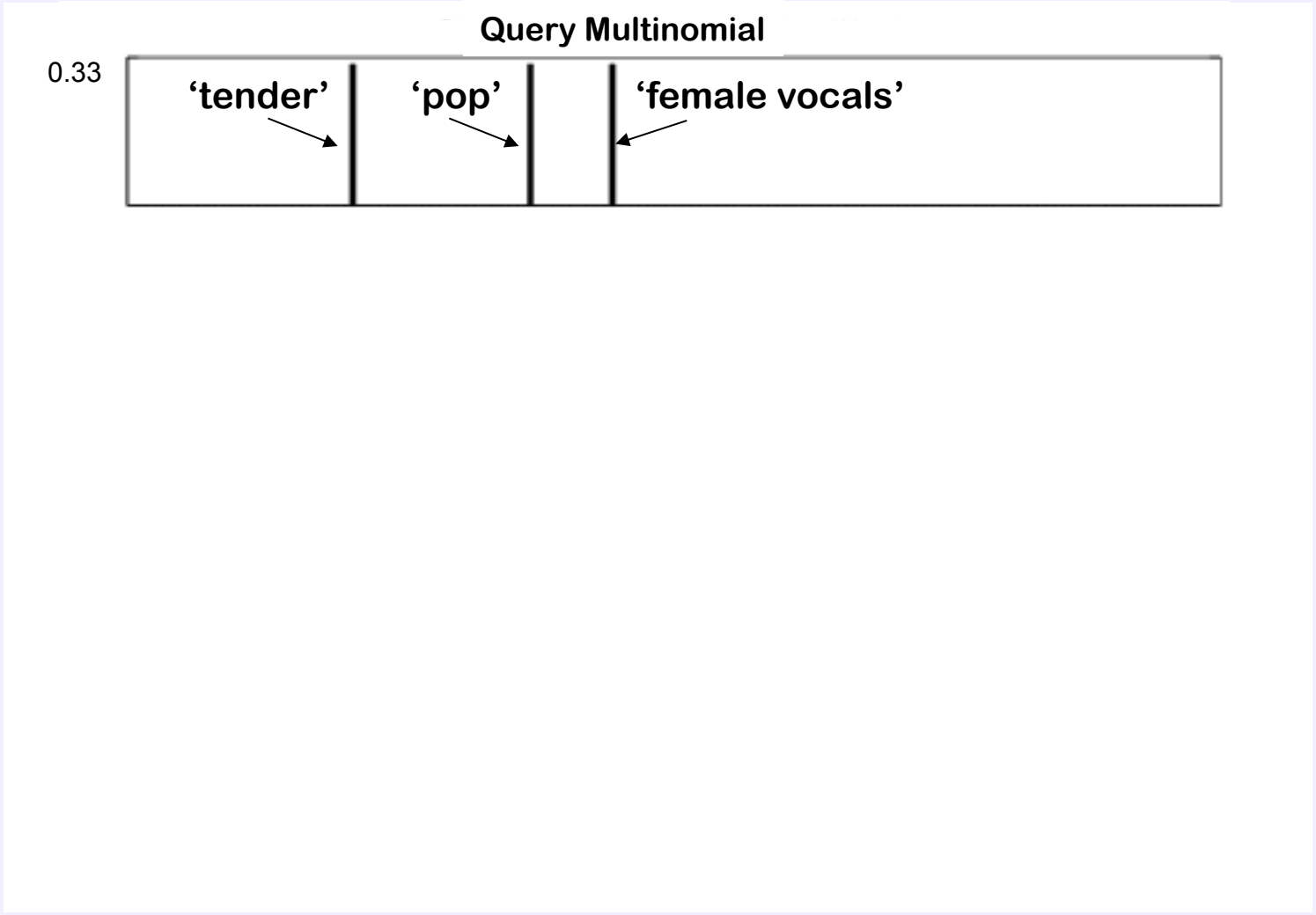
1. Annotate each song in corpus with a semantic multinomial **p**
   - **p** = {$P(t_1|X)$, …, $P(t_{|V|}|X)$}
2. Given a text-based query, construct a query multinomial **q**
   - $q_i = 1/|t|$, if tag t appears in the query string
   - $q_i = 0$, otherwise
3. Rank all sor

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

(KL) divergence

Retrieval

Query: ‘a tender pop song with female vocals’



Retrieval

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

The Age of Music Proliferation

Production:

- 5M artist page
- 150M distinct song

Distribution

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

Consumption

- 11M Internet radio us
- 110M iPods sold

Comparing Tags

Approach	Songs	Density	AROC
Ground Truth CAL500	All	0.15	1.00
	Long-Tail	0.15	1.00
Social Tags Last.fm	All	0.23	0.62
	Long-Tail	0.03	0.54
Game Listen Game	All	0.37	0.65
	Long-Tail	*	*
Web Autotags	All	0.67	0.66
	Long-Tail	0.25	0.56
Audio Autotags	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 $\mathbf{M_{s,d}}$ = mapping of **songs** to **documents**

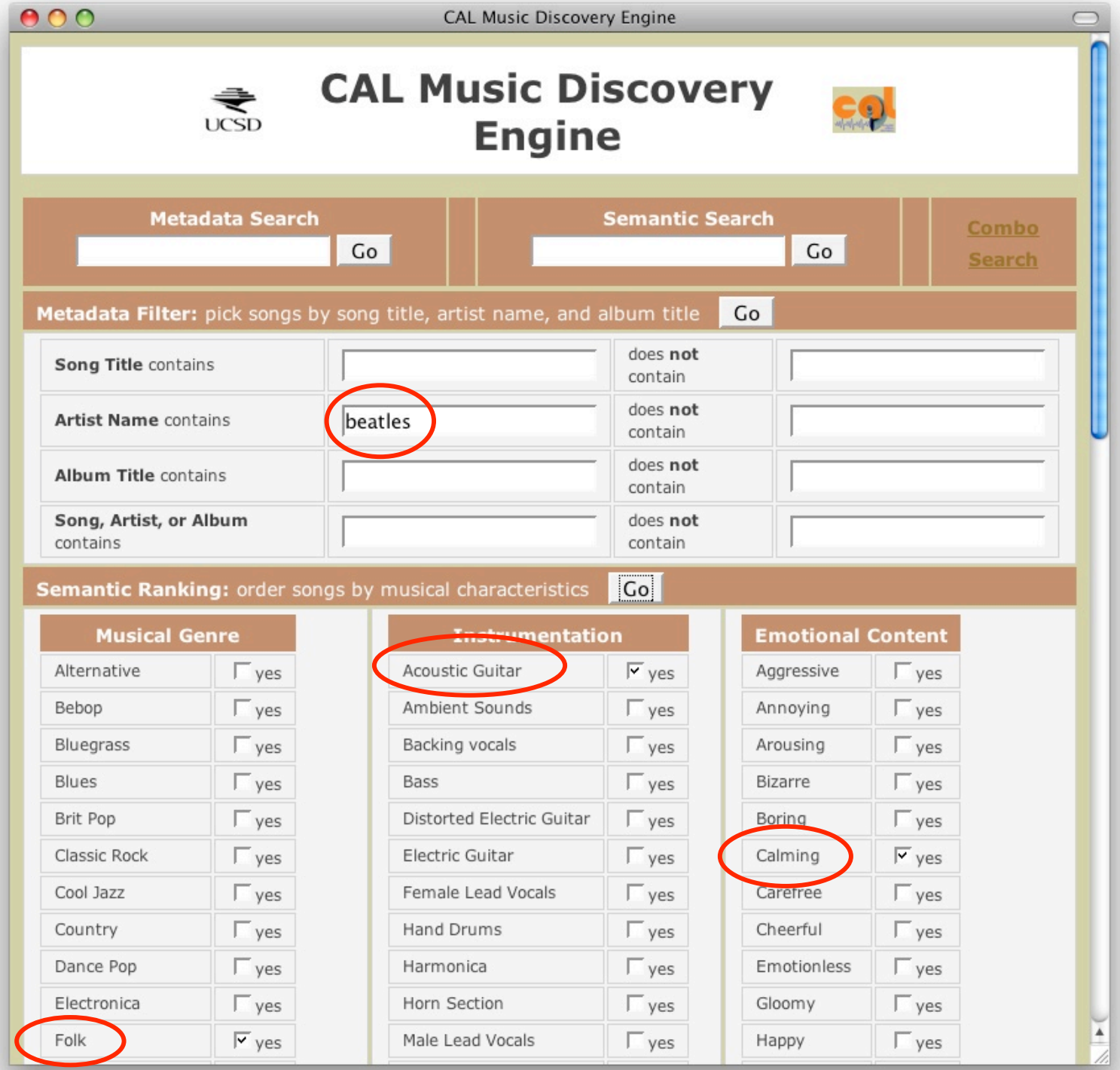
Text-mining System

Step 2: Autotag songs

For each tag *t*:

1. Query corpus with tag ***t*** to find relevant documents
   - $w_{t,d}$ → 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
Metadata Search
Go
Semantic Search
Go
Combo Search
Metadata Filter: pick songs by song title, artist name, and album title
Go
Song Title contains
does not contain
Artist Name contains
beatles
does not contain
Album Title contains
does not contain
Song, Artist, or Album contains
does not contain
Semantic Ranking: order songs by musical characteristics
Go
Musical Genre
Alternative yes
Bebop yes
Bluegrass yes
Blues yes
Brit Pop yes
Classic Rock yes
Cool Jazz yes
Country yes
Dance Pop yes
Electronica yes
Folk yes
Instrumentation
Acoustic Guitar yes
Ambient Sounds yes
Backing vocals yes
Bass yes
Distorted Electric Guitar yes
Electric Guitar yes
Female Lead Vocals yes
Hand Drums yes
Harmonica yes
Horn Section yes
Male Lead Vocals yes
Emotional Content
Aggressive yes
Annoying yes
Arousing yes
Bizarre yes
Boring yes
Calming yes
Carefree yes
Cheerful yes
Emotionless yes
Gloomy yes
Happy yes


CAL Music Discovery Engine
UCSD
CAL Music Discovery Engine
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 Ecstacy
'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%s' 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

http://www.allmusic.com/cg/amg.dll?p=amg&sql=11:difrxqr5ldje