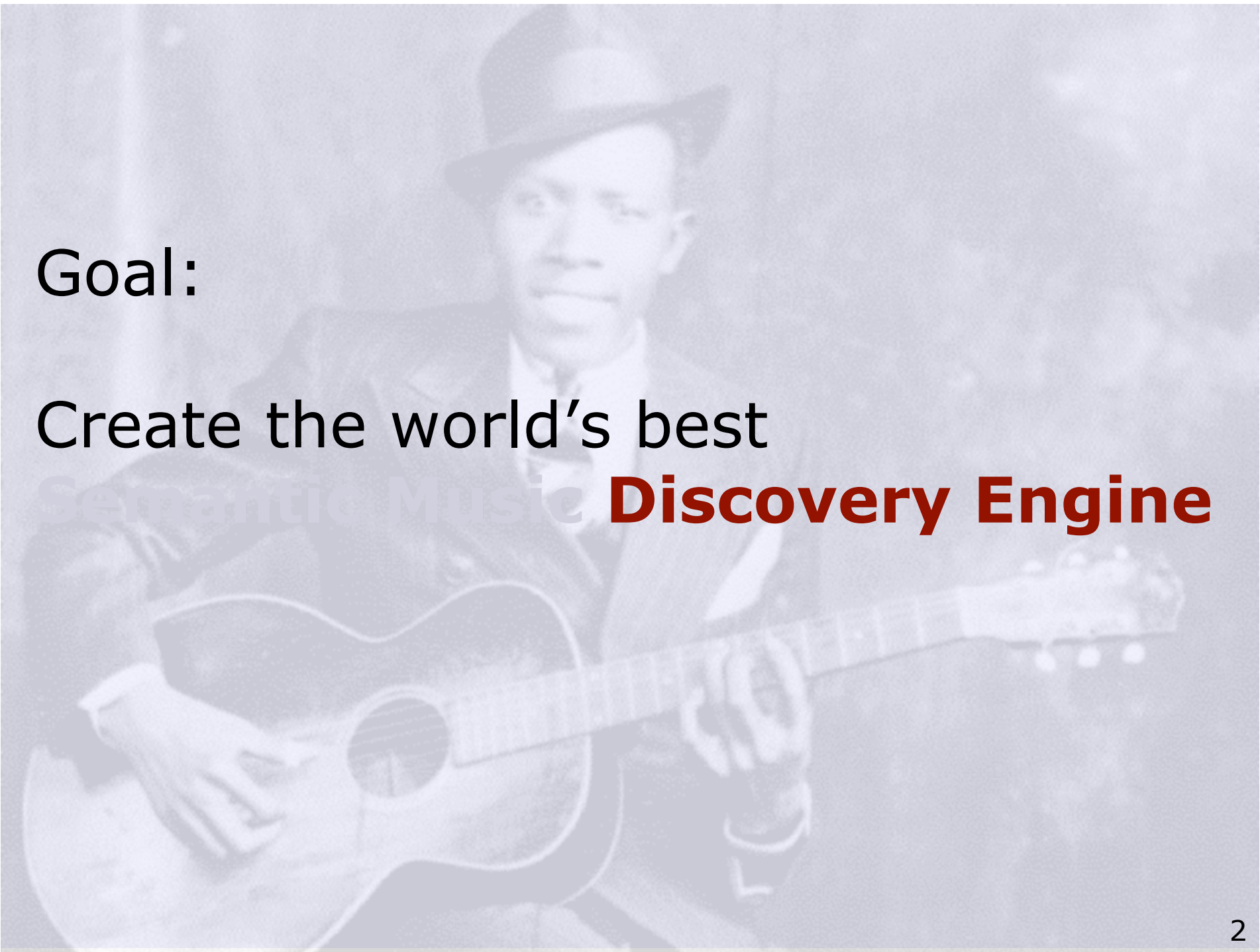Goal:

Create the world's best
Semantic Music Discovery Engine

Goal:

Create the world's best
Discovery Engine

Goal:

Create the world's best
Music Discovery Engine

Goal:

Create the world's best
Semantic Music Discovery Engine

Meerkat

I Made A Lovers Prayer by Gillian Welch

Blue Diamond Mines by The Johnson Mountain Boys

Whiskey Lullaby by Alison Krauss

Tramp On The Street by Dian & The Greenbriar Boys

Count Me In by Deana Carter

'Whiskey Lullaby' Tags:

folk influences

country women

chicks

country

easy listening

female-vocalists

Station Tags:

[   ] (Add Tag)
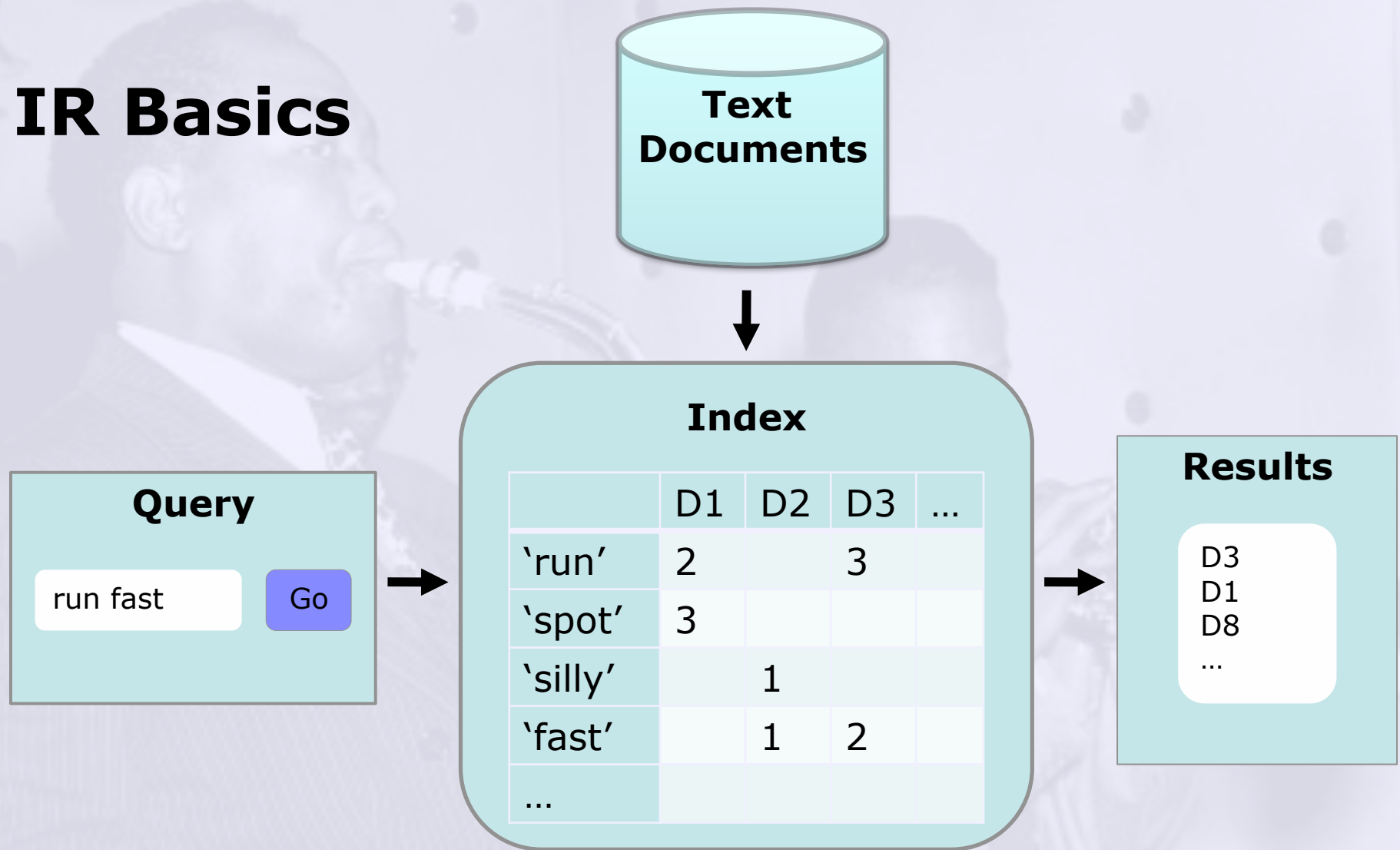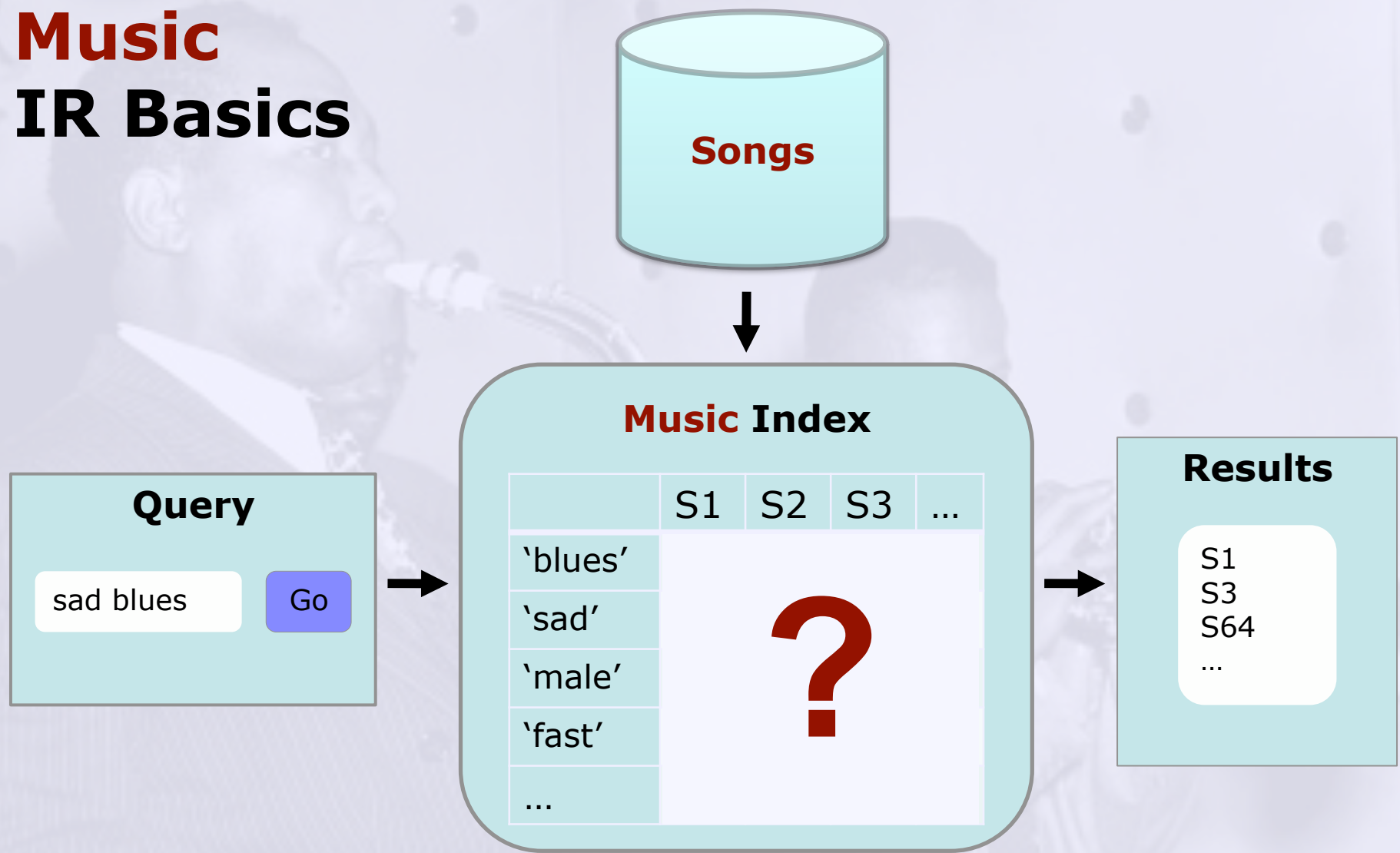
old time country ✖

bluegrass ✖

female vocals ✖

Problem:

Music Annotation is Hard

IR Basics

Text Documents

Index

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

Query

run fast   Go

Results

D3
D1
D8
...

Music
IR Basics

Songs

Music Index

	S1	S2	S3	...
'blues'				
'sad'		?		
'male'				
'fast'				
...				

Query

sad blues   Go

Results

S1
S3
S64
...

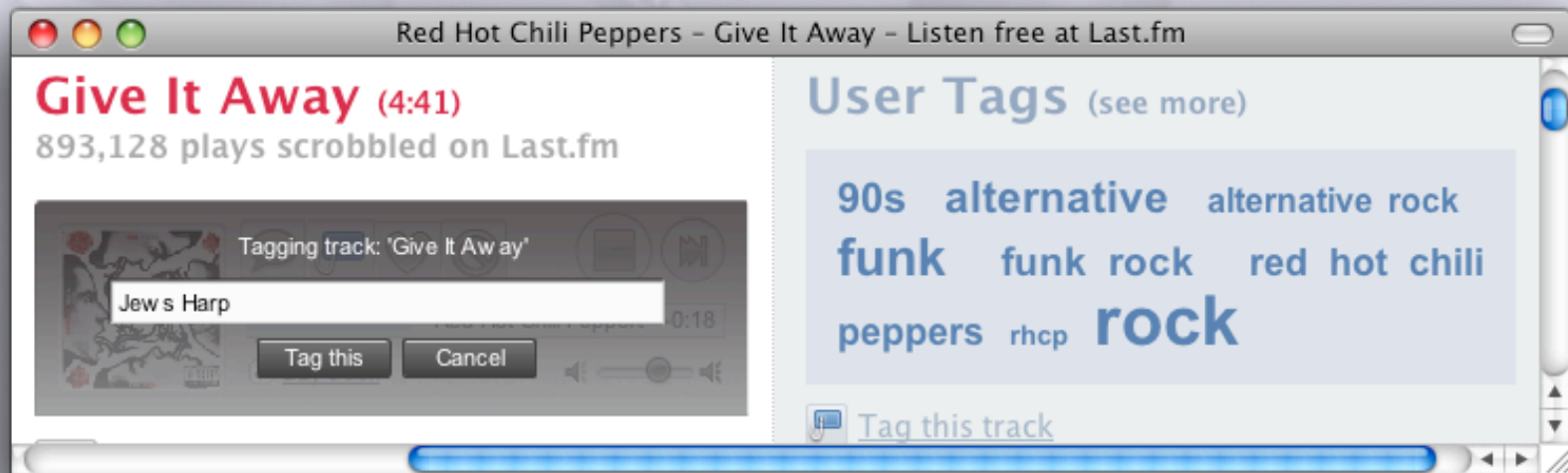# Four Sources of Music Information

1. Human Survey

CAL500 [SIGIR 07]

- 174 tags
- 55 undergrads, 120 hours
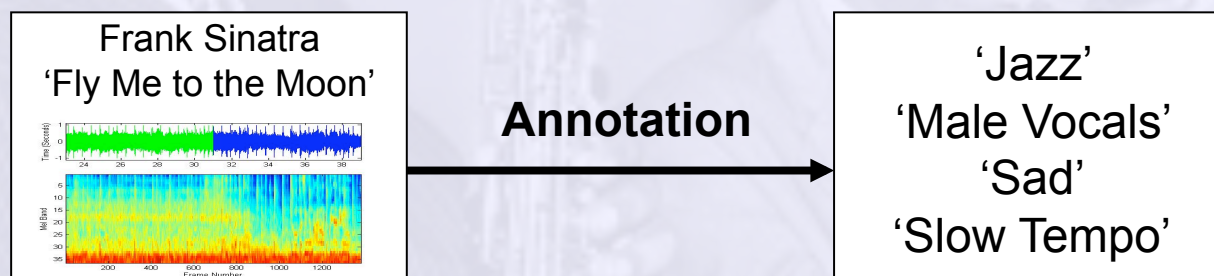- 500 songs annotated by 3+ people

2. Social Tagging

3. Mining Web Documents [Knees 07]

4. Content-based AutoTagging [SIGIR 07]

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

Frank Sinatra 'Fly Me to the Moon'	Annotation →	'Jazz' 'Male Vocals' 'Sad' 'Slow Tempo'

4. Content-Based AutoTagging

Supervised Multi-class Labeling model [SIGIR 07]

- One Gaussian Mixture Model (GMM) per tag
- Mixture Hierarchies EM Algorithm

Notes:

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

Comparing Data Sources

Groundtruth

- **CAL500** - {72 tags} x {500 songs} binary labels

Approaches

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

Comparing Data Sources

For each 3 approach:

    For each 72 tag:

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

    Calculate mean AUC

Constructing a Music Index

Approach	AUC	Best for # Tags
Social Tags	0.623	9
Text Mining Web Documents	0.625	12
Content-Based Autotagging	**0.731**	51
Single Source Oracle	**0.756**	

Combining Data Sources

Combining Data Sources

Approaches

1. Calibrated Score Averaging - [Zadrozny 02]

2. RankBoost - [Freund 03]

3. Kernel Combination SVM- [Lanckriet 04]

1. Calibrated Score Averaging

$f(\text{score}) \approx p(t \mid \text{score})$

Ground Truth Label

Predicted Score
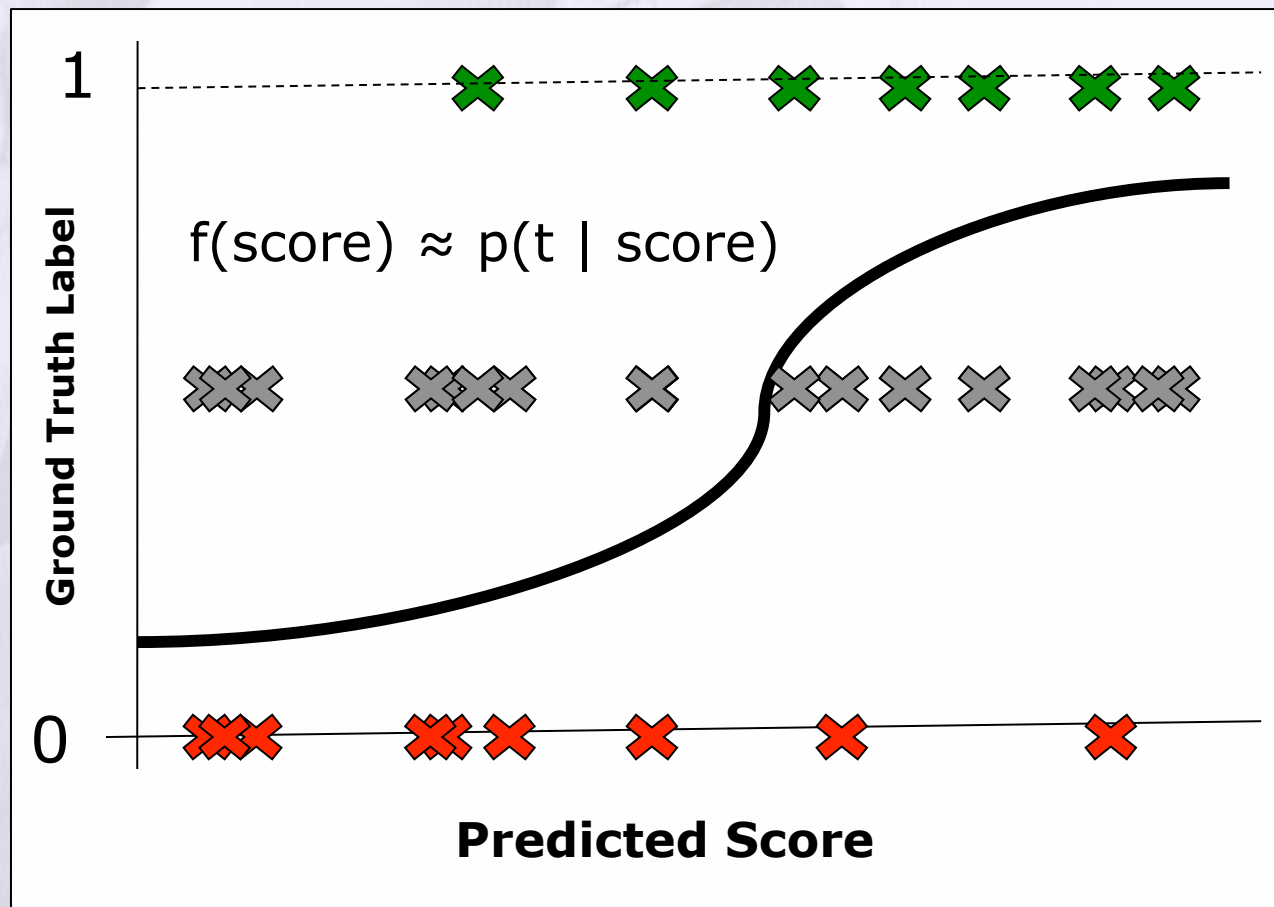
1. Calibrated Score Averaging

Autotag Score (green)

$p(t|s_a)$
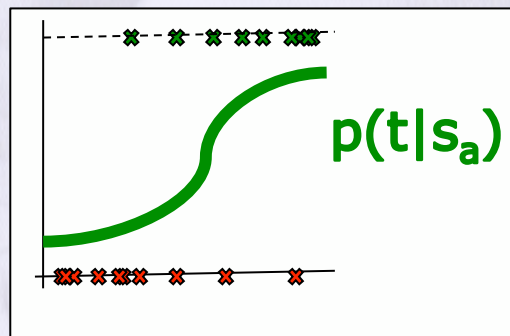
Web Document Score (blue)

$p(t|s_w)$

Social Tag Score (orange)

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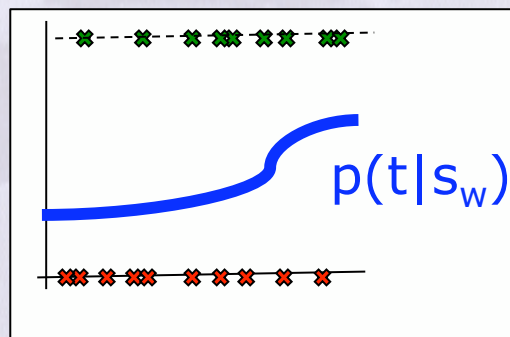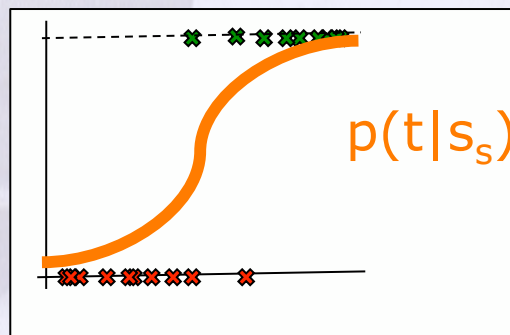
$$p(t|\mathbf{s}) = \Pi_i \, p(t|s_i)$$

2. RankBoost [Freund 03]

- Greedy iterative algorithm:
  - pick **weak ranker** & assigns **weight** based on loss function
- **Strong ranker** is a weighted voted by weak rankers

2. RankBoost

Weak Ranker = (Source, Threshold, Weight)

AutoTags		
id	Score	Label
8	**0.82**	1
2	**0.77**	0
13	**0.76**	1
7	**0.69**	1
3	**0.68**	0
...		

w_3 0

w_1 0

Web Docs		
id	Score	Label
7	**19**	1
14	**18**	0
9	**16**	1
11	**16**	0
21	**15**	1
...		

w_4 0

Social Tags		
id	Score	Label
13	**100**	1
4	**95**	1
17	**95**	0
11	**89**	0
8	**82**	1
...		

w_2 0

2. RankBoost

Strong Ranker: f(s)

Test Song s = [s_{audio}, s_{web}, s_{social}]

$$f(s) = w_1 \, I(s_{audio} > t_1) + w_2 \, I(s_{social} > t_2) + \ldots$$

3. Kernel Combo SVM

1. Compute **kernel matrix** for each data source.
   - 3 Song X Song similarity matrices

$$w_1 \; \boxed{K_{audio}} \; + \; w_2 \; \boxed{K_{web}} \; + \; w_3 \; \boxed{K_{social}} \; = \; \boxed{K_{combo}}$$

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**.

Summary

Approach	AUC
Single Source Oracle	0.756
Kernel Combination	0.756
RankBoost	0.760
Calibrated Score Averaging	**0.763**

Summary

Data Sources provide complementary information

- Sparse **Social Context**

- Dense **Audio Content**

Calibrated Score Averaging & RankBoost work well

- Good performance, easy to implement, easy to tune, quick to train & evaluate

Chapter 4

The Future (Man)

Research Challenges

1. ## Combine Music Information Sources

   - Games, Social Networks, Popularity Info

2. ## Improving **Autotagging**

   - Discriminative Approaches [Mandel 08, Eck 07]

3. ## **Personalization**

   - Demographic Groups

   - Psychographic Groups

   - Individual Preference

   - Emotional states of Individual
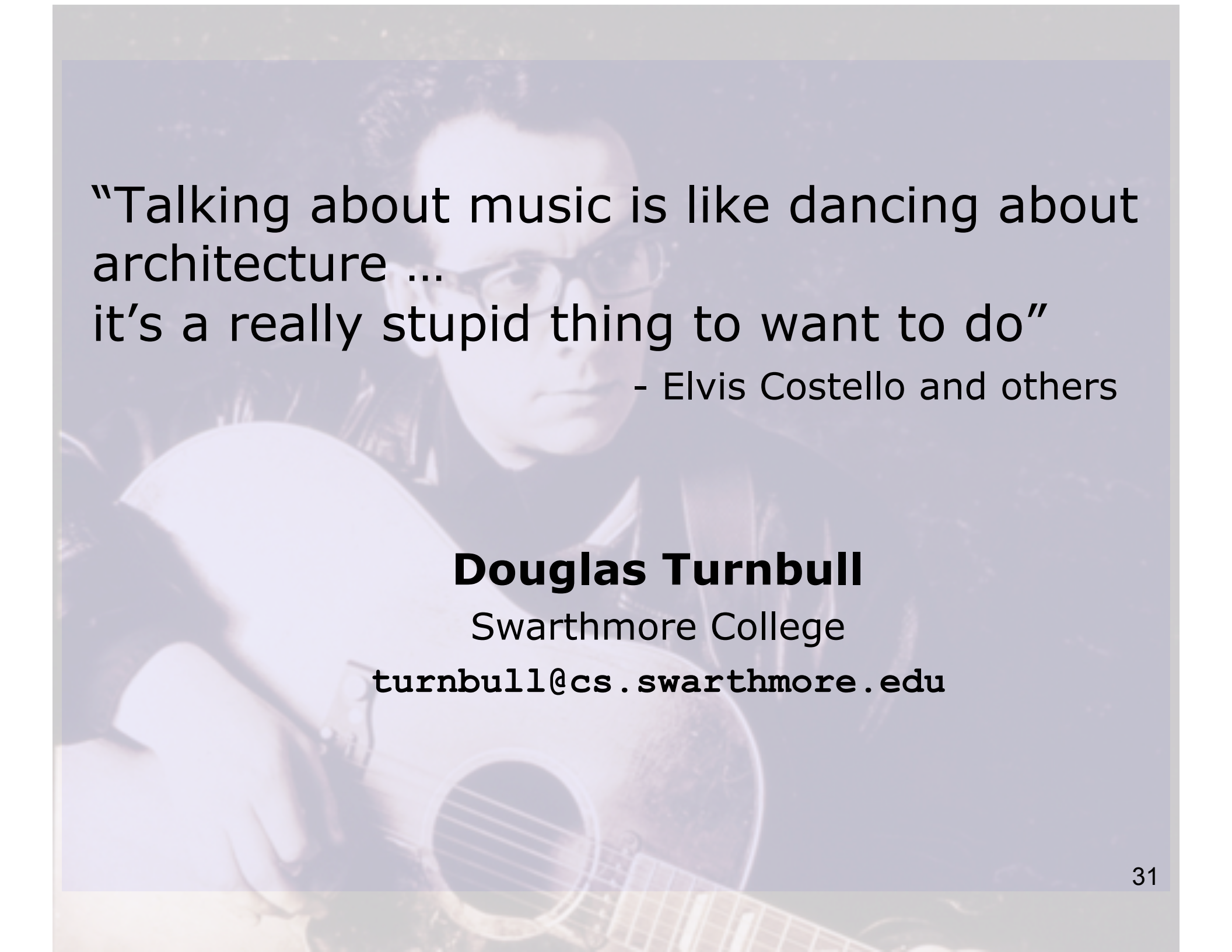
What's on tap

1. ## Big new data set

   - 10,000 songs

   - Acoustic, Genre, Social Tags

2. ## Herd-it Game [HCOMP 09]

   - **http://apps.facebook.com/herd-it**

   - Analyzing Game Data

3. ## New Everything

   - Autotagging Approaches

   - Content-Context Approach

   - User Interfaces

"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`

References

Social Context-Audio Content [Tech, SIGIR 09, ISMIR 08]
Autotagging [IEEE TASLP 08, SIGIR 07]

Related:

Music Annotation Games [HCOMP 09, ISMIR 07a]

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

Tag Vocabulary Selection with Sparse CCA [ISMIR 07b]

Supervised Music Boundary Detection [ISMIR 07c]

Collaborative Filtering & Tag Propagation [Tech]