

A background image of a woman with long dark hair, wearing a dark jacket, playing an acoustic guitar. The image is semi-transparent and serves as a backdrop for the text.

Tag Propagation based on Artist Similarity

Joon Hee Kim
Brian Tomasik
Douglas Turnbull
Swarthmore College

ISMIR 2009

Artist Annotation with Tags

Ani DiFranco

Acoustic Instrumentation
Folk Rock
Feminist Lyrics

Jack Johnson

Acoustic Guitar
Folk
Mellow

New Artist

???

Ben Harper

Slide Guitar
Rock
Mellow

Three Annotation Problems

Cold Start

- new artists are not labeled

Data Sparsity

- most artists have been annotated with very few tags

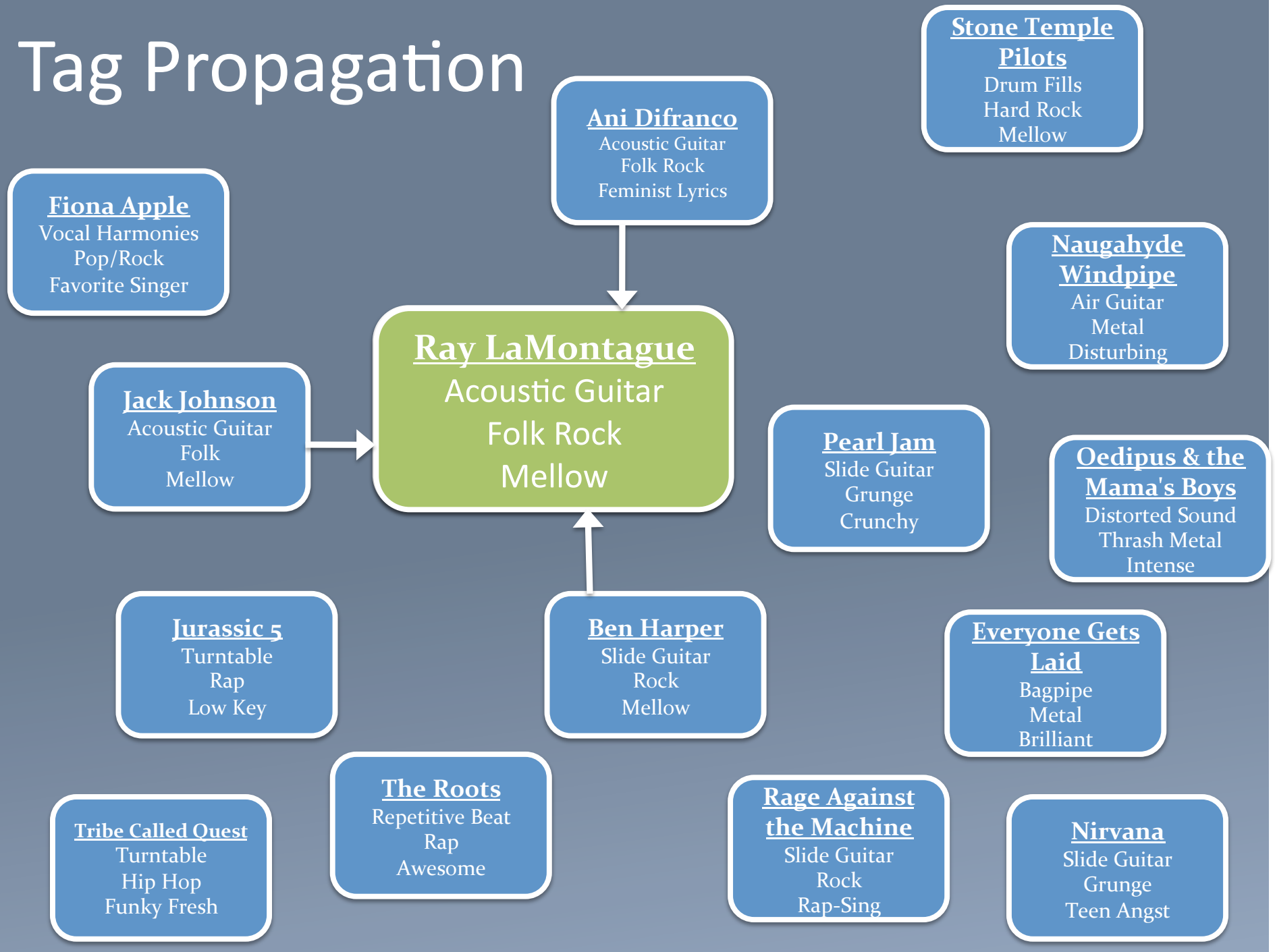
Data Reliability

- some artists have been annotated with incorrect tags

Tag Propagation address these problem

- originally proposed by Sordo et al. at ISMIR 2007

Tag Propagation



Tag Propagation

Vocabulary	A1	A2	A3	A4	A5	A6
Acoustic Guitar	1		1	1		
Electric Guitar		1		1		1
Slide Guitar					1	
Rock		1				1
Folk	1			1		1
Mellow	1		1			
...



New Artist
3/3
1/3
2/3
2/3
...

Similarity Score	A1	A2	A3	A4	A5	A6
New Artist	.1	.4	.3	.2	.5	.5



3-Nearest Neighbors

Calculating Similarity

Four Data Sources:

1. Preference Information
2. Social Tags
3. Web Documents
4. Audio-Content

Preference Information

1. Collect 400,000 user preference data from last.fm
 - Implicit - listening habit
 - Explicit - list of loved tracks
2. build artist co-occurrence matrix
3. calculate similarity by popularity normalization
 - equivalent to User-by-Item rating matrix [Celma 08]

	artist 1	artist 2	artist 3	artist 4	popularity
artist 1		3	1	6	$\sqrt{10}$
artist 2	3		2	2	$\sqrt{7}$
artist 3	1	2		0	$\sqrt{3}$
artist 4	6	2	0		$\sqrt{8}$
popularity	$\sqrt{10}$	$\sqrt{7}$	$\sqrt{3}$	$\sqrt{8}$	

Social Tags

1. Collect top tags for an artists
 - 0.5% pruning gives 949 unique tags
2. Represent each artist as a tag vector
 - 1 if tag present, 0 otherwise
3. Compute Cosine Similarity

Web Documents

1. Collect 50 web pages: "Artist Name" "music"
2. Append documents
3. Represent as a bag-of-words
4. Re-represent as TF-IDF vector
5. Calculate Cosine Similarity

Audio Content

1. Pick one song from each artist
2. Extract bag-of-MFCC-vectors from song
3. Two Content-based Approaches:
 - CB Acoustic [Mandel 05]
 - Represent each song by the mean and (full) covariance matrix
 - Calculate Symmetric (Gaussian) KL Divergence
 - CB Semantic [Barrington 07]
 - Calculate normalized vector of **autotag** probabilities
 - Calculate (Multinomial) KL Divergence between vectors

Experimental Setup

3,5000 artists

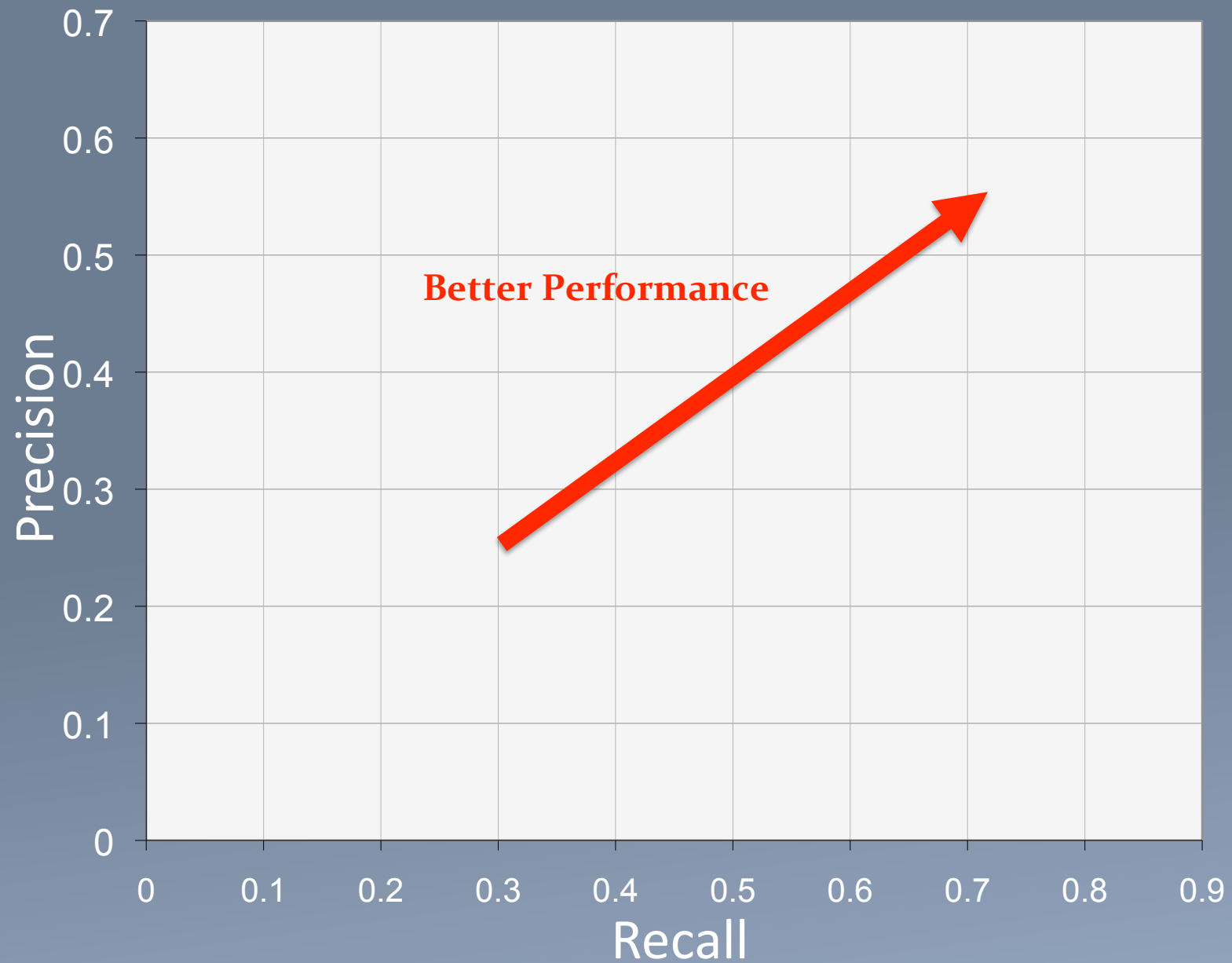
Three Vocabularies (1982)

1. Acoustic (891) Pandora Music Genome Genes
2. Genre (142) Pandora Radio Stations
3. Social (949) Last.fm social tags

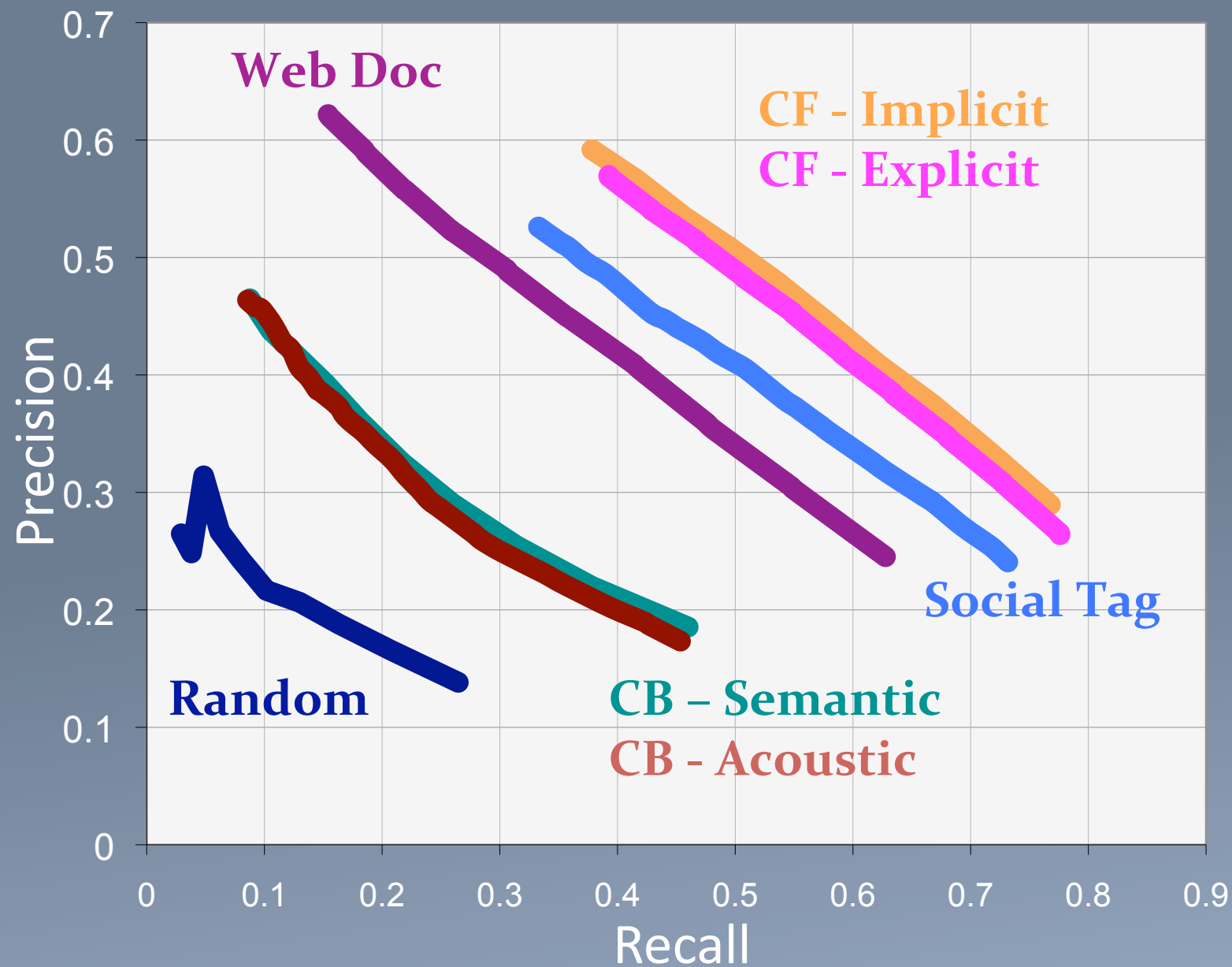
Artist Annotation Task Setup

- Leave-one-out cross-validation
- Estimate tag vector for each artist
 1. find k most similar artists (k=32)
 2. average 32 tag vectors
 3. transform the vector into binary score
 4. calculate precision, recall with varying threshold

Annotation



Annotation



Retrieval Task setup

- Leave-one-out cross-validation
- For each tag,
 1. generate a ranked-list of 3,500 artists
(based on estimated association strength)
 2. calculate r-precision, 10-precision, MAP, AUC

Retrieval

approach	Acoustic (891 Tags)			
	r-precision	10-precision	MAP	AUC
random	0.025	0.023	0.029	0.495
CB Semantic	0.115	0.123	0.091	0.714
CB Acoustic	0.118	0.132	0.088	0.692
Web Documents	0.244	0.300	0.200	0.814
Social Tags	0.267	0.274	0.237	0.874
CF Explicit	0.282	0.304	0.246	0.878
CF Implicit	0.281	0.306	0.254	0.882

Retrieval

approach	Genre (142 Tags)			
	r-precision	10-precision	MAP	AUC
random	0.012	0.015	0.017	0.499
CB Semantic	0.087	0.103	0.069	0.687
CB Acoustic	0.101	0.127	0.076	0.701
Web Documents	0.321	0.393	0.282	0.861
Social Tags	0.344	0.349	0.311	0.889
CF Explicit	0.362	0.388	0.329	0.909
CF Implicit	0.362	0.381	0.342	0.914

Retrieval

approach	Social (949 Tags)			
	r-precision	10-precision	MAP	AUC
random	0.030	0.029	0.033	0.498
CB Semantic	0.107	0.126	0.084	0.662
CB Acoustic	0.117	0.159	0.092	0.661
Web Documents	0.318	0.478	0.286	0.797
Social Tags	0.428	0.584	0.413	0.874
CF Explicit	0.410	0.562	0.396	0.869
CF Implicit	0.409	0.543	0.394	0.876

Conclusion

Collaborative Filtering (CF) works well

- Surprising since it does not explicitly model semantics

Content-Based (CB) doesn't work well

- Unfortunate because if it had, we would have solved the autotagging problem

Tag propagation is an alternative way to evaluate music similarity