From Web 2.0 to Web 3.0 using Data Mining

Steffen Staab & Rabeeh Abbasi
Web 2.0 Folksonomies

- Folksonomies facilitate sharing resources
  - Pictures
  - Bookmarks
  - Publications

- Users annotate resources with tags (keywords)
- Tags can be used for searching and browsing resources
Why do you want to have a Semantic Web at all?

- 4 billion photos!
- Many of them tagged
Why do you want to have a Semantic Web at all?

Solid growth

Growth of Flickr

Number of Images

Source: http://www.kullin.net/2009/10/flickr-reaches-4-billion-photos.html
Why I want to have a Semantic Web!

- Intelligent queries
  - Thesauri

Many photos are unsearchable!

Requirement: Know that “40ies” = “forties”
Why I want to have a Semantic Web!

Faceted Search & Browsing: Semaplorer (btc.uni-koblenz.de)

Requirement: Know what is a city, country, landmark
Is there a quantifiable benefit in Ontology Learning?
Problems and Features of Folksonomies

Low-Level Features

Geotags
51°30'03"N 0°7'24"W

Folksonomy

Add Tags

Search Resources Browse Resources

Groups

Add Tags

Low-Level Features
Discovering and Exploiting Semantics

1. Geotags
   - 51°30'03"N 0°7'24"W

2. Improve Semantic Browsing

3. Recommend Tags

flickr Groups

Low-Level Features

Folksonomy
Improving Search in Folksonomies
(Can we improve search, particularly for sparse results?)
Search in Folksonomies

- Billions of resources in Folksonomies

  Flickr: 4,000,000,000 Photos And Counting

  October 12th, 2009 | by Christina Warren
  @ 36 Comments and 413 Reactions

  On Saturday, photo number four billion was uploaded to photo sharing site Flickr. This comes just five and a half months after the 3 billionth and nearly 18 months after photo number two billion.

  As impressive as four billion photos is, that figure is trounced by Facebook, which reported more than 15 billion photos in its k in April and currently adds two billion more-month.

- But still many photos are unsearchable!

  ![flickr search screenshot]
Still many photos are unsearchable, because

- they have been tagged with other (semantically related) tags!
Most of the pictures in Flickr have only a few tags
- Leads into extremely sparse data!
Consider the tagging system as a Vector Space Model (VSM)

- Each resource is represented as a vector of tags

<table>
<thead>
<tr>
<th></th>
<th>1942 cent</th>
<th>flower</th>
<th>forties</th>
<th>hibiscus</th>
<th>nature</th>
<th>penny</th>
</tr>
</thead>
<tbody>
<tr>
<td>1942</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cent</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>flower</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>fortes</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>hibiscus</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>nature</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>penny</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Discovering Semantically Related Tags

- Enrich the existing Vector Space Model using semantically related tags
  - Dimensions of Semantically Related Tags
    - Context of tags (Resource / User)
    - Distribution of tags (Similar / Generalized)

```
Tag Context

Resource          User

Similar          Generalized

SR_G                    -
SR_C       SU_C

Resource Context & Generalized Tags
Resource Context & Similar Tags
User Context & Similar Tags
```
Resource vs. User Context

- **Resource Context**
  - Narrow context, based on the resources associated to a tag

- **User Context**
  - Broader context, based on the vocabulary of a user

<table>
<thead>
<tr>
<th>Resource</th>
<th>User 1</th>
<th>User 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1942</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>cent</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>flower</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>forties</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>hibiscus</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>nature</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>penny</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Semantically Related Tags – Tag Distribution

• Tag Distribution
  • Similar Tags (Symmetric Similarity)
    – Using Cosine Similarity: \( \text{cosine}(r_{1*}, r_{2*}) = \frac{r_{1*} \cdot r_{2*}}{||r_{1*}|| \cdot ||r_{2*}||} \)

<table>
<thead>
<tr>
<th></th>
<th>1942</th>
<th>cent</th>
<th>flower</th>
<th>forties</th>
<th>hibiscus</th>
<th>nature</th>
<th>penny</th>
</tr>
</thead>
<tbody>
<tr>
<td>1942</td>
<td>1</td>
<td>0.71</td>
<td>0</td>
<td>0.71</td>
<td>0</td>
<td>0</td>
<td>0.71</td>
</tr>
<tr>
<td>cent</td>
<td>0.71</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>flower</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.71</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>forties</td>
<td>0.71</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>hibiscus</td>
<td>0</td>
<td>0</td>
<td>0.71</td>
<td>0</td>
<td>1</td>
<td>0.71</td>
<td>0</td>
</tr>
<tr>
<td>nature</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.71</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>penny</td>
<td>0.71</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Semantically Related Tags – Tag Distribution

- Tag Distribution
  - Generalized Tags (Asymmetric Similarity)
    - Using Modified Overlap Coefficient

\[
gen(r_{1*}, r_{2*}) = \begin{cases} 
\frac{|r_{1*} \cap r_{2*}|}{|r_{1*}|} & \text{if } |r_{1*}| \leq |r_{2*}|; \\
0 & \text{otherwise.}
\end{cases}
\]

<table>
<thead>
<tr>
<th>Tag</th>
<th>Generalized Tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>hibiscus</td>
<td>hibiscus(1.00), flower(0.61), flowers(0.25), red(0.17), macro(0.15)</td>
</tr>
<tr>
<td>hibiskus</td>
<td>hibiskus(1.00), hibiscus(0.67), flower(0.47), blume(0.33), garten(0.29)</td>
</tr>
<tr>
<td>ibisco</td>
<td>ibisco(1.00), flower(0.65), hibiscus(0.52), flor(0.43), red(0.30)</td>
</tr>
<tr>
<td>ibiscus</td>
<td>ibiscus(1.00), flower(0.55), hibiscus(0.45), nature(0.41), flowers(0.41)</td>
</tr>
<tr>
<td>hibisco</td>
<td>hibisco(1.00), flor(0.66), flower(0.65), hibiscus(0.39), macro(0.21)</td>
</tr>
<tr>
<td>rosemallow</td>
<td>rosemallow(1.00), hibiscus(0.52), flower(0.48), malvaceae(0.38), garden(0.24)</td>
</tr>
<tr>
<td>gumamela</td>
<td>gumamela(1.00), flower(0.58), philippines(0.38), hibiscus(0.35), red(0.23)</td>
</tr>
<tr>
<td>roseofsharon</td>
<td>roseofsharon(1.00), flower(0.66), macro(0.28), flowers(0.28), hibiscus(0.21)</td>
</tr>
<tr>
<td>malvaceae</td>
<td>malvaceae(1.00), flower(0.73), hibiscus(0.62), flowers(0.28), macro(0.22)</td>
</tr>
</tbody>
</table>
Use semantically related tags to enrich the existing vector space model

<table>
<thead>
<tr>
<th></th>
<th>1942</th>
<th>cent</th>
<th>flower</th>
<th>forties</th>
<th>hibiscus</th>
<th>nature</th>
<th>penny</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Vector Space</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1942</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cent</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>flower</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>forties</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>hibiscus</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>nature</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>penny</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

| Enriched Vector Space |
| 1942   | 1.71 | 2.41 | 0      | 0       | 0        | 0      | 0     |
| cent   | 0.71 | 2.71 | 0      | 0       | 0        | 0      | 0     |
| flower | 0    | 0    | 1.71   | 0.71    | 0        | 0      | 0     |
| forties| 0.71 | 2.71 | 0      | 0       | 0        | 0      | 0     |
| hibiscus| 0  | 0    | 1.71   | 1.71    | 0        | 0      | 0     |
| nature | 0    | 0    | 0      | 0       | 0.71     | 1.71   | 0     |
| penny  | 1.71 | 0.71 | 0      | 0       | 0        | 0      | 0     |
## Examples – Context and Distribution of Tags

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>brick</strong></td>
<td>brick(1.00) wall(0.11)</td>
<td>brick(1.00) wall(0.19) building(0.13) red(0.11)</td>
<td>brick(1.00) wall(0.37) fence(0.36) window(0.36) rust(0.35)</td>
</tr>
<tr>
<td><strong>bromelia</strong></td>
<td>bromelia(1.00) airplant(0.32) bromeliad(0.17) tillandsia(0.15)</td>
<td>bromelia(1.00) bromeliad(0.35) tillandsia(0.30) flower(0.27) nature(0.12)</td>
<td>bromelia(1.00) lirio(0.18) tibouchina(0.15) strelitzia(0.15)</td>
</tr>
<tr>
<td><strong>pub</strong></td>
<td>pub(1.00) crawl(0.17)</td>
<td>pub(1.00) beer(0.11) bar(0.11)</td>
<td>pub(1.00) beer(0.25) bar(0.24) sign(0.22) london(0.22)</td>
</tr>
<tr>
<td><strong>style</strong></td>
<td>style(1.00) crave(0.14) arian(0.13) fashion(0.11) persians(0.10)</td>
<td>style(1.00) fashion(0.26) beauty(0.11)</td>
<td>style(1.00) fashion(0.16) hair(0.13) woman(0.12) man(0.12)</td>
</tr>
</tbody>
</table>
Evaluation

- Dataset
  - Image uploaded to Flickr between 2004/05
  - Tags used by at least 10 different users
  - Dataset statistics
    - Photos: ~27 Million
    - Tags: ~92 Thousand
    - Users: ~0.3 Million
    - Tag Assignments: ~94 Million
- Queries
  - Subset of AOL query log
Evaluation

- User based evaluation
  - done by 18 users
- 150 queries split into 3 sets of 50
  - having exact matches:
    - 1 to 10 (sparse results)
    - 11 to 50 (normal)
    - more than 50 (too many results)
- Search results were presented to users for different vector space models
- Calculated Precisions @ 5, 10, 15, 20
Results – Sparse Results

1 to 10 Relevant Resources

- Simple
- Res./Similar
- Res./Gen.
- User/Similar
- Semi Random
- Random

Number of Relevant Resources

Precision at 20

- Results
- Significant Improvement for sparse results
Results

11 to 50 Relevant Resources

- Simple
- Res./Similar
- Res./Gen.
- User/Similar
- Semi Random
- Random

<table>
<thead>
<tr>
<th>Number of Relevant Resources</th>
<th>Precision at 20</th>
</tr>
</thead>
<tbody>
<tr>
<td>11-50</td>
<td>0.68</td>
</tr>
</tbody>
</table>
More than 50 Relevant Resources

- **Simple**
- **Res./Similar**
- **Res./Gen.**
- **User/Similar**
- **Semi Random**
- **Random**

**Number of Relevant Resources**
- Precision at 20

- More than 50 Relevant Resources

<table>
<thead>
<tr>
<th>Number of Relevant Resources</th>
<th>Simple</th>
<th>Res./Similar</th>
<th>Res./Gen.</th>
<th>User/Similar</th>
<th>Semi Random</th>
<th>Random</th>
</tr>
</thead>
<tbody>
<tr>
<td>50+</td>
<td>0.71</td>
<td>0.72</td>
<td>0.69</td>
<td>0.69</td>
<td>0.68</td>
<td>0.66</td>
</tr>
</tbody>
</table>
Browsing (landmarks) in Folksonomies
One prominent photo returned by Flickr when searching for „Koblenz“

Is this what you expect when searching for a place?
Usage context: Faceted Browsing

- Semaploer (btc.uni-koblenz.de)
  - Handles billions of triples,
  - Lacks semantics for individual photos
    - Cannot identify landmark photos
Finding Landmark Photos

Which photos are (non-)representative for a given place?
Building a Classifier

- For a classifier we need
  - Features of the photos
    - Tags
  - Positive and Negative training examples
    - Exploit Flickr groups
Positive Flickr Groups

- Landmarks
- City Landmark

Positive Training Examples

Normalization

SVM

Classification Model

Classifier

+VE

-VE

Negative Flickr Groups

- CAR [directory] 車
- Kids and nature

Examples

- flickr.com/photos/conner395/1018557535/, flickr.com/photos/66164549@N00/2508246015/, flickr.com/photos/kupkup/1356487670/, flickr.com/photos/asam/432194779/, flickr.com/photos/michaelfoleyphotography/392504158/
User Evaluation

- Datasets: Training 430282 photos, Test Dataset 232265 photos
- Comparison with state of the art WorldExplorer
- 20 Users evaluated both methods for 50 different cities
- Improvement of 12% in precision for our method

![Precision per User](image)

![Graph showing precision per user for World Explorer and Our Method](image)
Semantic Browsing in Folksonomies
I want to “brow**se**” vehicle images!!!
- how can I do it?
  - can I do it using a Tag Cloud?

Perhaps I need to structure the tags and resources!
- how can I do it?
  - Put them into categories (like Vehicles, People, etc)!
    - Do it Manually or with Training?
      » Might not be possible on a large scale!
    - Automatically and without any training!
      » Using T-ORG!
Organize **resources** by putting their **tags** into **categories** depending upon their **context**.

Users can browse categories to retrieve required resources.
Tag Organization using T-ORG

- Select ontologies related to the categories (e.g. Vehicle, People, etc.)
- Prune and refine these ontologies according to the desired categories (add missing concepts, filter existing concepts)
- Apply the classification algorithm T-KNOW to classify the tags and resources
- Browse the categories to explore the tags and resources
Classifying the tags using T-KNOW

Use well-known linguistic patterns to generate queries

Search these patterns on Google and download search results

Select the concept which has the highest similarity with the context of the tag

Compare each Google search result with the context of the tag and extract the concept
Compute similarity using cosine measure between Bag of Words (BOW) representation of “Tag Context” and “Search Result”.

**Tag Context**
- singen
- cars
- motors
- ford
- 1955

**Search Result**
- TCC Group | About TCC Group | Our Clients
- international organizations such as Ford Foundation’s Cons 
- Agricultural Research, and Monsanto Foundation ...

**BOW**
- 1955 = 1
- as = 0
- cars = 1
- ford = 1
- foundation = 0
- international = 0
- motors = 1
- organizations = 0
- singen = 1
- such = 0

**BOW**
- 1955 = 0
- as = 1
- cars = 0
- ford = 1
- foundation = 2
- international = 1
- motors = 0
- organizations = 1
- singen = 0
- such = 1

\[
\cos(\hat{c}, \hat{a}) = \frac{\hat{c} \times \hat{a}}{|\hat{c}| |\hat{a}|} = 0.15
\]

**Only consider the results having similarity above a certain Threshold**
- Result having the highest similarity is considered as final
Experimental Setup

**4+1 Categories**
- Person
- Location
- Vehicle
- Organization
- Other

**932 Concepts**
- Author, Singer, Human, …
- Country, District, City, Village, …
- Vehicle, Car, Truck, Motorbike, Train, …
- Company, Organization, Firm, Foundation, …

**189 random Images from 9 Flickr groups**

**1754 Tags**
Experimental Setup – Classifiers

- Two human classifiers: \( K \) (gold standard) and \( S \)
- T-KNOW
Experimental Setup – Evaluation

- F-Measure
  
  \[ A = \text{set of correct classification by test (user S or T-KNOW)} \]
  \[ B = \text{set of all classification by Gold Standard (user K)} \]
  \[ C = \text{set of all classifications by test} \]

- Precision = \( \frac{A}{C} \)
- Recall = \( \frac{A}{B} \)
- F-Measure = \( \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \)
- Results not far away from Human Classification

51% 79%

51% 79%
Tag Recommendations in Folksonomies
Simplifying the annotation process
### Image

![Image](http://www.flickr.com/photos/64255998@N00/283890351/)

47° 4’ 32” N
15° 26’ 12” E

### Tags

- Clocktower
- Uhrturn
- Graz
- Austria
- Europe
- Herbst
- Autumn
- Night
- Nacht
Tag Recommendation

Input

47° 4’ 32" N
15° 26’ 12" E

Tag Recommender

Output

Clocktower
Uhrturm
Graz
Austria
Europe
Herbst
Autumn
Night
Nacht

http://www.flickr.com/photos/64255998@N00/283890351/
Tag Recommendation

**Input**
- Geographical Coordinates: 47° 4’ 32” N, 15° 26’ 12” E
- Tags
- Low-level Features

**Output**
- Clocktower
- Uhrturn
- Graz
- Austria
- Europe
- Herbst
- Autumn
- Night
- Nacht

Tag Recommender
Tag Recommendation

Input

Output

Clocktower
Uhrturm
Graz
Austria
Europe
Herbst
Autumn
Night
Nacht

Input

Geographical Coordinates

47° 4’ 32" N
15° 26’ 12" E

Tags

Low-level Features

Which image features to use?
Features Available in Large Datasets

- CoPhIR Dataset contains
  - **106** million images from Flickr with
    - Title, location, tag, comment, etc
    - Five standard MPEG-7 Features

<table>
<thead>
<tr>
<th>Low-level Feature</th>
<th>Properties</th>
<th>Dims</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalable Color</td>
<td>Color histogram</td>
<td>64</td>
</tr>
<tr>
<td>Color Structure</td>
<td>Localized color distributions</td>
<td>64</td>
</tr>
<tr>
<td>Color Layout</td>
<td>Color and spatial information</td>
<td>12</td>
</tr>
<tr>
<td>Edge Histogram</td>
<td>Local-edge distribution</td>
<td>80</td>
</tr>
<tr>
<td>Homogeneous Texture</td>
<td>Texture</td>
<td>62</td>
</tr>
</tbody>
</table>
Tag Recommender – System Overview

Training Data

Location 1

Features

Location 1 Clusters

Location N

Features

Location N Clusters

Model
System Overview

**Input**

- Image Features
- Tags
- Location 1
- Location N

**Output**

- Clocktower
- Graz
- Austria
- Europe
- Autumn
- Night
- Uhrturm
- Nacht
- Herbst
- HDR
- Photomatix

**Model**

- Training Data
- Location 1
- Location N

**Features**

- Image
- Tags

**Location 1 Clusters**

- Features

**Location N Clusters**

- Features
Clustering

- Used Standard K-Means clustering

**Input:** Feature space $F \in \{G, L, T, D\}$, Number of clusters $k$

**Output:** A set of $k$ clusters

**Method:**

1. Randomly choose $k$ images from feature space $F$ as the initial cluster centers
2. Assign each image to the closest cluster
3. Update cluster means
4. If cluster means are changed, then repeat step 2-3
Given a cluster containing homogeneous images

• Find most representative tags
  – Tag Rank ~ Frequency of Users

Cluster Tags

Clocktower (4)
Graz (4)
Austria (3)
Europe (3)
Autumn (2)
Night (2)
Uhrturm (1)
Nacht (1)
Herbst (1)
HDR (1)
Photomatix (1)
Classification

- Map a new image to one of the existing clusters
- Find the nearest cluster centroid
  - Assign representative tags of the mapped cluster to the image

Representative Tags
Nearest Centroid
Evaluation

- Evaluated on tags of ~400,000 Flickr images
  - 75% training data
  - 25% test data / gold standard
    - Evaluation based on the correctly identified tags of test data
  - Data is randomly split based on users (not resources)
- Ignored
  - 10 most frequent tags for each location (city) and
  - Tags used by less than three users
- Calculated
  - Precision
  - Recall
  - F-Measure
Micro Precision

![Graph showing precision vs. number of tags recommended. The graph compares Geo Coordinates, Low-level Features, Tags, and Baseline (Random).]

- **Geo Coordinates**
- **Low-level Features**
- **Tags**
- **Baseline (Random)**

**Geo EHD Tags Random**

**Precision**

**Number of tags recommended**

- Precision values range from 0.0000 to 0.1600.
- The graph shows a decrease in precision as the number of tags recommended increases.

**Geo Coordinates** and **Tags** show a significant drop in precision compared to **Low-level Features** and **Baseline (Random)**.
Micro Recall

Recall vs. Number of tags recommended

- Geo
- EHD
- Tags
- Random
WeST Micro F-Measure

The diagram shows the F-Measure values for different numbers of tags recommended. The x-axis represents the number of tags recommended, ranging from 1 to 20. The y-axis represents the F-Measure values, ranging from 0 to 0.1. The graph lines are labeled as Geo, EHD, Tags, and Random, each indicating different methods or models for tag recommendation.
Coverage

- Number of different tags correctly suggested

![Coverage Diagram](image)
Summary

We can discover semantics in Web 2.0 and exploit them for
- Improved search
- Semantic Browsing
- Tag recommendation
Lessons Learned

- There is a quantifiable value in ontology learning techniques

- Improvement of user experience is about
  - Semantics
    - A little bit of semantics goes a long way!
      - Who?
      - What?
      - When?
      - Where?
  - Representativity / Prototypicality
    - Do not ignore!
    - Not represented in SemWeb!
Most interesting result for „Koblenz“:

A Castle half an hour away from Koblenz

Much work left to be done…. 
More details in:


- Abbasi, Grzegorzek, Staab: Large Scale Tag Recommendation Using Different Image Representations. (SAMT 2009)

- Abbasi, Chernov, Nejdl, Paiu, Staab: Exploiting Flickr Tags and Groups for Finding Landmark Photos. (ECIR 2009)

- Abbasi, Staab, Cimiano: Organizing Resources on Tagging Systems using T-ORG. (ESWC 2007)
Thank you for your attention!