

# Understanding the Motivation behind Tagging

Author: Christian Körner - Advisor: Markus Strohmaier

July 1, 2009



# Research Questions

- Can we measure tagging motivation?

# Research Questions

- Can we measure tagging motivation?
- Does tagging motivation vary on different tagging systems?

# Research Questions

- Can we measure tagging motivation?
- Does tagging motivation vary on different tagging systems?
- How do different types of tagging motivation influence tags and resulting folksonomies?

# Research Questions

- Can we measure tagging motivation?
- Does tagging motivation vary on different tagging systems?
- How do different types of tagging motivation influence tags and resulting folksonomies?
- **Important:** This work focuses on tag pragmatics rather than tag semantics

# Two Different Types of Tagging Motivation

- Two vastly different types of motivations

# Two Different Types of Tagging Motivation

- Two vastly different types of motivations

	<i>Categorizer</i>	<i>Describer</i>
Goal	later browsing	later retrieval
Change of Vocabulary	costly	cheap
Size of Vocabulary	limited	open
Tags	subjective	objective

Table 1 - Characteristics of Categorizers and Describers

# Quantitative Measures (Answer to RQ1)

- Tag-Resource Ratio

$$\text{tagRatio} = \frac{|T|}{|R|} \quad (1)$$



# Quantitative Measures (Answer to RQ1)

- Tag-Resource Ratio

$$\text{tagRatio} = \frac{|T|}{|R|} \quad (1)$$

- “Orphaned” Tags - Proportion of tags which only annotate a small amount of resources

$$D_n = \frac{|T_n|}{|T|}, T_n = \{t \mid |R(t)| \leq n\}, n = \lceil \frac{|R(t_{max})|}{100} \rceil \quad (2)$$

# Quantitative Measures (Answer to RQ1)

- Tag-Resource Ratio

$$\text{tagRatio} = \frac{|T|}{|R|} \quad (1)$$

- “Orphaned” Tags - Proportion of tags which only annotate a small amount of resources

$$D_n = \frac{|T_n|}{|T|}, T_n = \{t \mid |R(t)| \leq n\}, n = \lceil \frac{|R(t_{max})|}{100} \rceil \quad (2)$$

- Tag Entropy - effectiveness of the encoding process of tagging

$$H(R|T) = - \sum_{r \in R} \sum_{t \in T} p(r, t) \log_2(p(r, t)) \quad (3)$$

$$C_n = \frac{H(R|T) - H_{opt}(R|T)}{H_{opt}(R|T)} \quad (4)$$

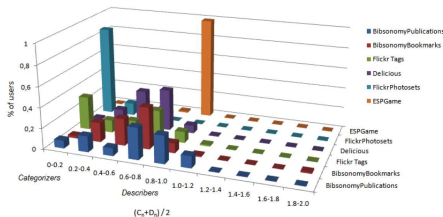
# The used Data Sets

<i>Data Set</i>	$ U $	$ T $	$ R $	$ R_u _{min}$	$\frac{ T }{ R }$
ESP Game	82	27.872	86.669	1.000	0,3216
Flickr Sets	173	6.286	245.282	500	0,0256
Flickr Tags	74	37.889	160.717	1.000	0,2357
Del.icio.us	445	112.173	585.580	1.000	0,1916
Bibsonomy Publications	26	11.006	23.696	500	0,4645
Bibsonomy Bookmarks	84	29.176	93.309	500	0,3127

Table 2 - Characteristics of the Data Sets

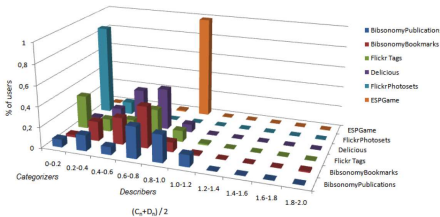
# Preliminary Results

- Answer to RQ2: Diversity of the different tagging systems

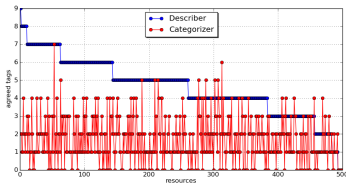


# Preliminary Results

- Answer to RQ2: Diversity of the different tagging systems



- Answer to RQ3: Tag Agreement of categorizer vs describer



# Implications...

- *Tag Recommendation*: Based on the two types of motivation different recommendation methods are needed
  - Categorizers need recommender based on their personal vocabulary.
  - Describers need recommender that capture content of the resources.

# Implications...

- *Tag Recommendation*: Based on the two types of motivation different recommendation methods are needed
  - Categorizers need recommender based on their personal vocabulary.
  - Describers need recommender that capture content of the resources.
- *User Interfaces*: Different user interfaces aid different motivation types (e.g. tag clouds for categorizer)

# Implications...

- *Tag Recommendation*: Based on the two types of motivation different recommendation methods are needed
  - Categorizers need recommender based on their personal vocabulary.
  - Describers need recommender that capture content of the resources.
- *User Interfaces*: Different user interfaces aid different motivation types (e.g. tag clouds for categorizer)
- Influence of tag pragmatics on tag semantics



# Outlook

- Development of a formal model describing how likely a given personomy represents the history of a categorizer or a describer.

# Outlook

- Development of a formal model describing how likely a given personomy represents the history of a categorizer or a describer.
- Expanding the types of tagging motivation (and other measurements)

# Outlook

- Development of a formal model describing how likely a given personomy represents the history of a categorizer or a describer.
- Expanding the types of tagging motivation (and other measurements)
- Further evaluation and validation

# Outlook

- Development of a formal model describing how likely a given personomy represents the history of a categorizer or a describer.
- Expanding the types of tagging motivation (and other measurements)
- Further evaluation and validation
- **Question:** Does tagging motivation influence the performance of recommendation systems?

# Take Home Message

- Measuring tagging motivation appears feasible
- Tagging motivation varies within and across different tagging systems
- Tagging motivation has influence on fabric of folksonomies