Overcoming the Imbalance Between Tag Recommendation Approaches and Real-World Folksonomy Structures with Cognitive-Inspired Algorithms

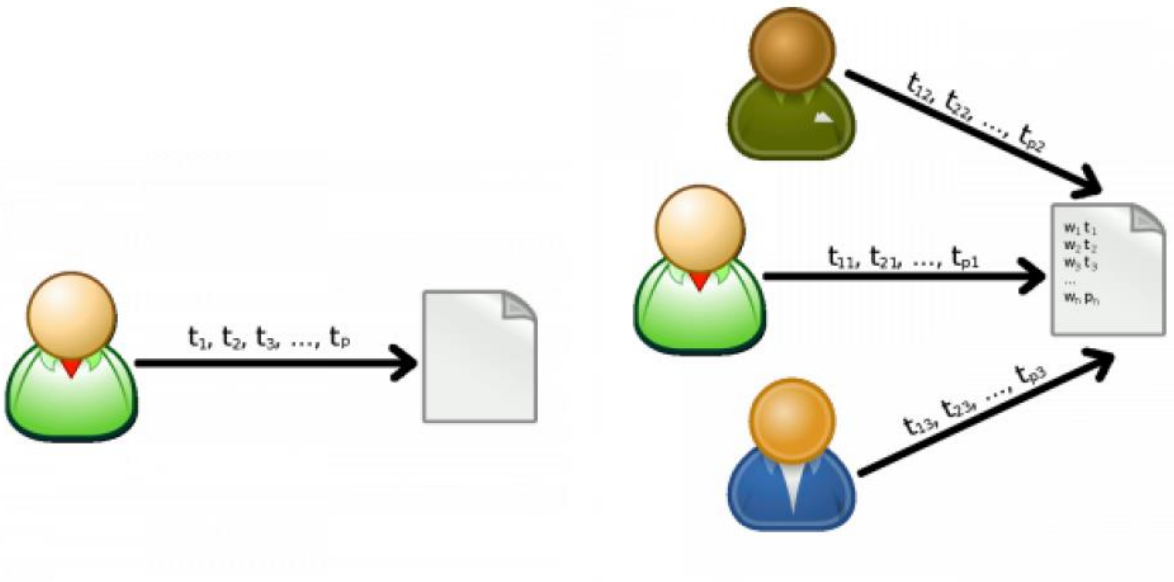
Dominik Kowald & Elisabeth Lex

Know-Center & Graz University of Technology

ESCSS, London, 15.11. – 17.11.2017

Social Tagging

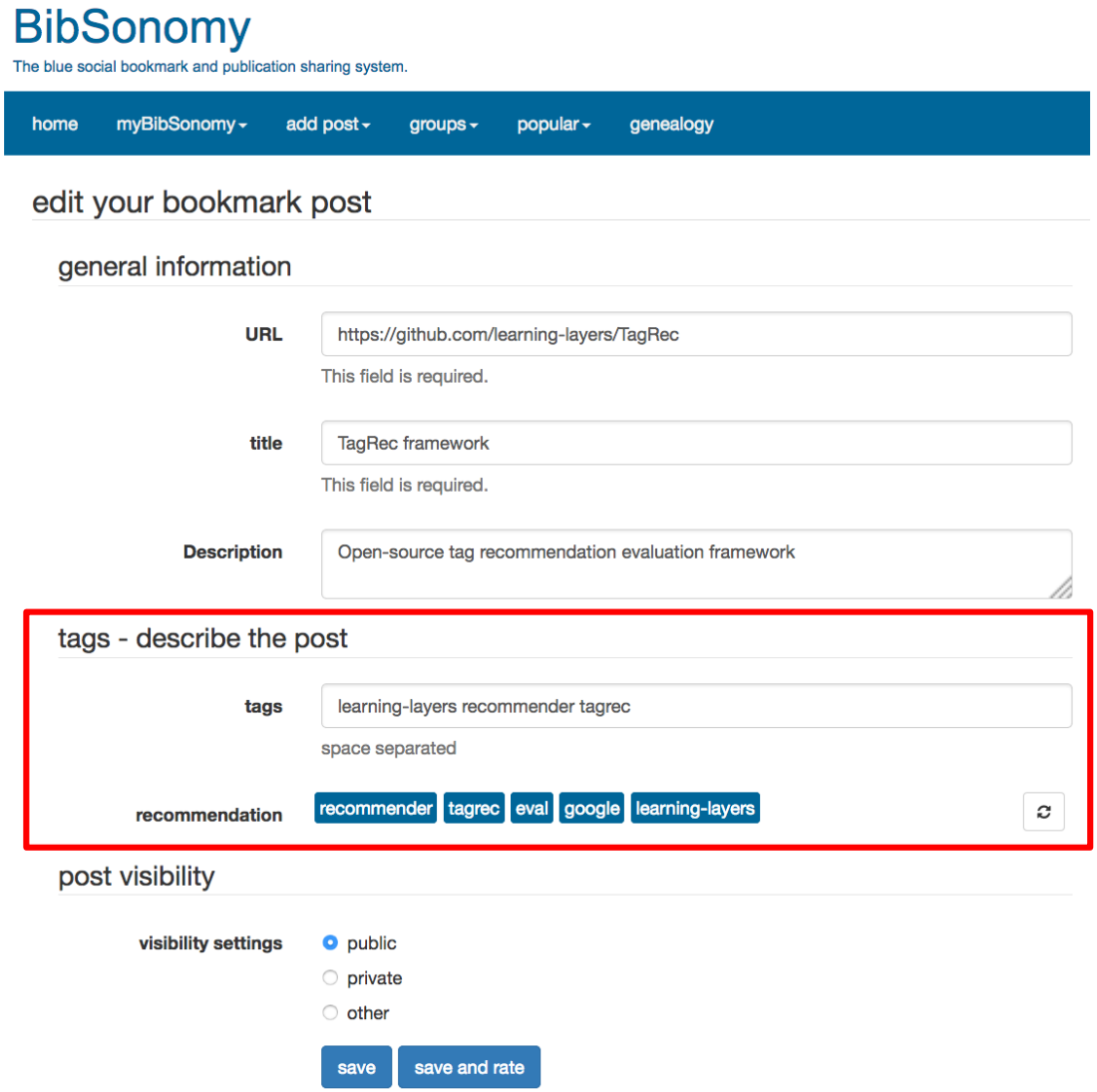
- Social tagging is the process of **collaboratively annotating content** with keywords (i.e., *tags*)
- Essential instrument of Web 2.0 to **structure and search Web content**



(a) Narrow folksonomy (b) Broad folksonomy

[Zubiaga, 2009]

Tag Recommendations



[BibSonomy, 2017]

Dominik Kowald & Elisabeth Lex
Know-Center & Graz University of Technology

Imbalance

- Current tag recommendation algorithms are designed in a purely **data-driven** way
  - Tag popularity, user similarities, topic modeling, factorization of resource features, etc.
  - Rely on **dense / broad** folksonomy structures
- Most real-world folksonomies are **sparse / narrow**

Dataset	$\|U\|$	$\|R\|$	$\|T\|$	$\|Y\|$	$\|B\|$	$\|B\|/\|U\|$	$\|B\|/\|R\|$
Flickr	9,590	856,755	125,119	3,328,590	856,755	89.338	1.000
CiteULike	18,474	811,175	273,883	3,446,650	900,794	48.760	1.110
BibSonomy	10,179	683,478	201,254	2,986,396	772,108	75.853	1.129
Delicious	15,980	963,741	184,012	4,266,206	1,447,267	90.567	1.501
LastFM	1,892	12,522	9,748	186,474	71,062	37.559	5.674
MovieLens	4,009	7,601	15,238	95,580	55,484	13.839	7.299

Approach

- The way users choose tags for their resources strongly corresponds to **processes in human memory** and its cognitive structures [Fu, 2008; Seitlinger & Ley, 2012]
  - **Activation processes** in human memory → **ACT-R** [Anderson et al., 2004]
  - **Activation equation** → usefulness of memory unit depends on **general usefulness** (i.e., frequency and recency) and usefulness in **current semantic context**

$$A_i = B_i + \sum_j (W_j \cdot S_{j,i})$$

$$B_i = ln(\sum_{j=1}^{n} t_j^{-d})$$

Dominik Kowald & Elisabeth Lex
Know-Center & Graz University of Technology

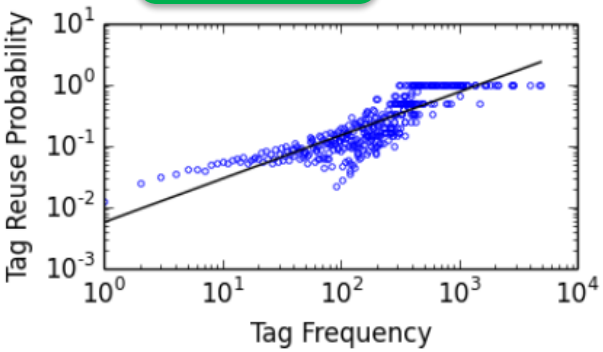
RQ1

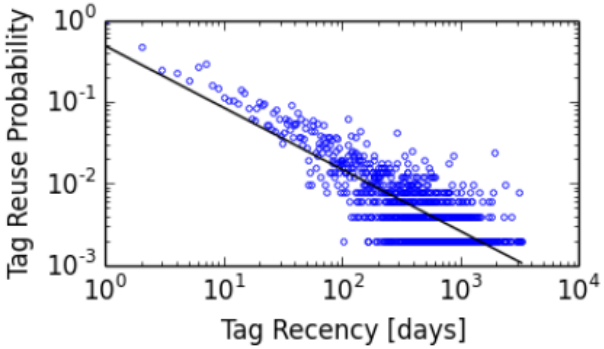
*How are **activation processes in human memory** influencing the **tag reuse behavior** of users in social tagging systems?*

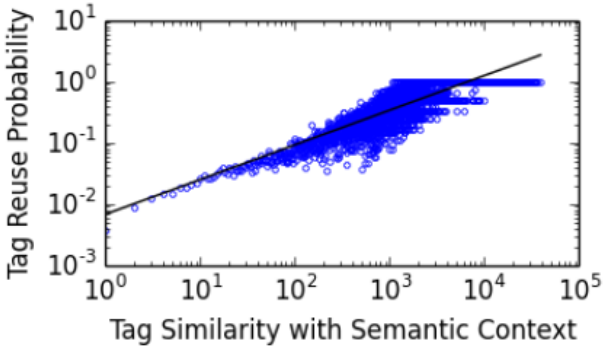
Kowald, D. (2015). Modeling cognitive processes in social tagging to improve tag recommendations. In *Proceedings of the 24th International Conference on World Wide Web, WWW '15 Companion*, ACM

Kowald, D. and Lex, E. (2016). The influence of frequency, recency and semantic context on the reuse of tags in social tagging systems. In *Proceedings of the 27th ACM Conference on Hypertext and Social Media, HT '16*, ACM.

RQ1 Results







[CiteULike, 2016]

- The **more frequently** a tag was used in the past **(k > 0)**, the higher its reuse probability is.
- The **more recently** a tag was used in the past **(k < 0)**, the higher its reuse probability is.
- The **more similar** a tag is to tags of the **current sem. context (k > 0)**, the higher its reuse probability is.

→ The **activation equation of ACT-R** models these factors

RQ2

*Can the **activation equation** of the cognitive architecture ACT-R be exploited to develop a **tag recommendation algorithm**, which is capable of overcoming the imbalance current approaches and real-world folksonomy structures?*

Kowald, D., Seitlinger, P., Trattner, C., and Ley, T. (2014). Long time no see: The probability of reusing tags as a function of frequency and recency. In *Proceedings of the 23rd International Conference on World Wide Web, WWW '14 Companion*, ACM

Kowald, D. and Lex, E. (2015). Evaluating tag recommender algorithms in real-world folksonomies: A comparative study. In *Proceedings of the 9th ACM Conference on Recommender Systems, RecSys '15*, ACM

RQ2 Results (nDCG@10)

Dataset	MP_r	$MP_{u,r}$	CF	LDA	PITF	FR	GIRPTM	ACT-R
Flickr	-	.569	.666	.280	.535	.561	.686	.711
CiteULike	.063	.392	.359	.138	.294	.392	.422	.438
BibSonomy	.091	.407	.369	.219	.327	.408	.409	.434
Delicious	.187	.358	.356	.271	.302	.292	.393	.431
LastFM	.283	.386	.317	.388	.414	.399	.397	.425
MovieLens	.271	.328	.254	.296	.324	.319	.326	.338

- **ACT-R** outperforms related tag recommendations methods in **narrow and broad** folksonomy settings

→ Cognitive-inspired approaches can **overcome the imbalance** between tag recommendations and folksonomies

RQ3

Given that activation processes in human memory can be modeled to improve tag recommendations, can they also be ***utilized for hashtag recommendations in Twitter****?*

Kowald, D., Pujari, S., and Lex, E. (2017). Temporal effects on hashtag reuse in Twitter: A cognitive-inspired hashtag recommendation approach. In *Proceedings of the 26th International Conference on World Wide Web, WWW'17*, ACM.

Kowald, D., Kopeinik, S., & Lex, E. (2017). The TagRec Framework as a Toolkit for the Development of Tag-Based Recommender Systems. In *Proc. of the 25th Conference on User Modeling, Adapation and Personalization, UMAP'2017*. ACM.

RQ3 Results (nDCG@10)

- **Scenario 1**: Hashtag recommendations w/o current tweet

Dataset	MP_I	MR_I	BLL_I	MP_S	MR_S	BLL_S	MP	FR	CF	**ACT-R**
CompSci	.175	.218	.225	.046	.154	.235	.012	.169	.196	**.324**
Random	.323	.352	.370	.144	.205	.280	.035	.324	.333	**.434**

- **Scenario 2**: Hashtag recommendations w/ current tweet

Dataset	SR	TCI	**ACT-R**
CompSci	.299	.385	**.446**
Random	.388	.507	**.562**

→ **Activation processes in human memory** can be utilized for **hashtag recommendations in Twitter**

Conclusion

RQ1 Activation processes in human memory (i.e., **frequency, recency and semantic context**) have an influence on tag usage practices

RQ2 The **activation equation of ACT-R** can be used to design a tag recommendation algorithm that **overcomes the imbalance** between current algorithms and the structure of real-world folksonomies

RQ3 This approach can also be generalized for **hashtag recommendations in Twitter**

- **Future Work**
  - **Adapt** approach for other types of cognitive-inspired recommender systems (e.g., resource recommendation)
  - Validate offline results with **online studies**

Thank you for listening!

Questions / suggestions? → **Poster**

Dominik Kowald

Elisabeth Lex

Know-Center & Graz University of Technology

→ All evaluations have been conducted using the open-source **TagRec tag recommendation benchmarking framework**

- https://github.com/learning-layers/TagRec

References (i)

- **[Anderson et al., 2004]** J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin. An integrated theory of the mind. *Psychological review*, 111(4):1036, 2004.
- **[Anderson & Schooler, 1991]** Anderson, J. R. and Schooler, L. J. Reflections of the environment in memory. *Psychological science*, 2(6), 1991
- **[Dellschaft & Staab, 2012]** Dellschaft, K. and Staab, S. (2012). Measuring the influence of tag recommenders on the indexing quality in tagging systems. In *Proceedings of Hypertext'12*, pages 73-82. ACM.
- **[Floeck et al., 2010]** Floeck, F., Putzke, J., Steinfels, S., Fischbach, K., and Schoder, D. (2010). Imitation and quality of tags in social bookmarking systems - collective intelligence leading to folksonomies. In *On collective intelligence*, pages 75-91. Springer.
- **[Font et al., 2015]** Font, F., Serrà, J., & Serra, X. (2015). Analysis of the impact of a tag recommendation system in a real-world folksonomy. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *7*(1), 6.
- **[Fu, 2008]** Fu, Wai-Tat. The microstructures of social tagging: a rational model. In *Proceedings of CSCW' 08*, pages 229-238. ACM, 2008
- **[Hadgu & Jäschke, 2014]** A. T. Hadgu and R. Jäschke. Identifying and analyzing researchers on twitter. In *Proceedings of WebSci '14*, pages 23-30, New York, NY, USA, 2014.

References (ii)

- **[Harvey & Crestani, 2015]** M. Harvey and F. Crestani. Long time, no tweets! Time-aware personalised hashtag suggestion. In *Proceedings of ECIR'15*. Springer, 2015.
- **[Hotho et al., 2006]** Hotho, A., Jäschke, R., Schmitz, C., and Stumme, G. (2006). Information retrieval in folksonomies: search and ranking. In Proceedings of ESCW'06, pages 411-426. Springer.
- **[Rendle & Schmidt-Thieme, 2010]** Rendle, S. and Schmidt-Thieme, L. (2010). Pairwise interaction tensor factorization for personalized tag recommendation. In *Proceedings of WSDM'10*, pages 81-90. ACM.
- **[Seitlinger & Ley, 2012]** Seitlinger, P. and Ley, T. (2012). Implicit imitation in social tagging: familiarity and semantic reconstruction. In *Proceedings of CHI'12*, pages 1631-1640. ACM.
- **[Wagner et al., 2014]** Wagner, C., Singer, P., Strohmaier, M., and Huberman, B. A. (2014). Semantic stability in social tagging streams. In *Proceedings of WWW'14*, pages 735-746. ACM.
- **[Wang et al., 2012]** Wang, M., Ni, B., Hua, X. S., & Chua, T. S. (2012). Assistive tagging: A survey of multimedia tagging with human-computer joint exploration. *ACM Computing Surveys (CSUR)*, *44*(4), 25.
- **[Zhang et al., 2012]** Zhang, L., Tang, J., and Zhang, M. (2012). Integrating temporal usage pattern into personalized tag prediction. In *Web Technologies and Applications*, pages 354-365. Springer.