

Wann	Thema	Literatur
15/04/10	Einführung, Themenbesprechung	
<b>I: Überblick über verschiedene Arten linguistischer Annotation</b>		
22/04/10	Semantisch annotierte Korpora: FrameNet/Salsa	
29/04/10	Syntaktisch annotierte Korpora (Baumbanken): Penn Treebank, TiGer Treebank, TüBa-D/Z	
<b>II. Der Annotationsprozess</b>		
06/05/10	Der Annotationsprozess – Richtlinien, Training, ... Besprechung der Gruppenarbeit	
13/05/10	<i>Himmelfahrt</i>	
20/05/10	Gruppenarbeit	
27/05/10	WordNet, Propbank, Besprechung der Ergebnisse der Gruppenarbeit, Erfahrungen bei der Annotationen	
<b>III: Evaluation</b>		
03/06/10	<i>Feiertag</i>	
10/06/10	Überblick über verschiedene Methoden der Evaluation	1
17/06/10	Inter-Annotator Agreement, Verlässlichkeit der Annotation, Intra-Annotator Agreement, Sind Goldstandards wirklich golden? Lernen mit fehlerbehafteten Daten	2, 3, 4, 5
24/06/10		
<b>IV. Wie kommen wir an mehr Daten?</b>		
01/07/10	Sentiment Analysis (Unterschiede in der Annotation, spezifische Probleme), Einfluss von Annotationsumgebung (Tools) und Erfahrung der Annotator/innen	6, 7
08/07/10	Semi-automatische Annotation, Bootstrapping	8, 9, 10, 16, 17, 18
15/07/10	Active Learning,	11, 12, 13, 14, 15

## Literatur:

1. Massimo Poesio and Ron Artstein. Inter-Coder Agreement for Computational Linguistics. *Computational Linguistics*, Volume 34 , Issue 4 (December 2008).
2. Maxwell, A. 1977. [Comparing the classification of subjects by two independent judges](#). *British Journal of Psychiatry*, 116,651-655.
3. Bayerl, Petra Saskia und Harald Lungen und Ulrike Gut und Karsten I. Paul (2003). [Methodology for reliable schema development and evaluation of manual annotations](#). In: *Proceedings of the Workshop on Knowledge Markup and Semantic Annotation at the Second International Conference on Knowledge Capture (K-CAP 2003)*. Sanibel Island, Florida.
4. Eyal Beigman and Beata Beigman Klebanov. Learning with annotation noise. *ACL 2009*. Singapore.
5. Cecily Jill Duffield, Jena D. Hwang, Susan Windisch Brown, Dmitriy Dligach, Sarah E. Vieweg, Jenny Davis, Martha Palmer. 2007. Criteria for the Manual Grouping of Verb Senses. *Proceedings of the Linguistic Annotation Workshop*, pages 49–52, Prague, June 2007.
6. Wiebe, Janys, Theresa Wilson and Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*. Volume. 39, no. 2, 165-210.
7. Dandapat, S., Biswas, P., Choudhury, M., and Bali, K. 2009. Complex linguistic annotation - no easy way out!: a case from Bangla and Hindi POS labeling tasks. In *Proceedings of the Third Linguistic Annotation Workshop* (Suntec, Singapore, August 06 - 07, 2009). *Annual Meeting of the ACL. Association for Computational Linguistics*, Morristown, NJ, 10-18.
8. Chou, W., Tsai, R. T., Su, Y., Ku, W., Sung, T., and Hsu, W. 2006. A semi-automatic method for annotating a biomedical proposition bank. In *Proceedings of the Workshop on Frontiers in Linguistically Annotated Corpora 2006* (Sydney, Australia, July 22 - 22, 2006). *ACL Workshops. Association for Computational Linguistics*, Morristown, NJ, 5-12.
9. Xue, N., Chiou, F., and Palmer, M. 2002. Building a large-scale annotated Chinese corpus. In *Proceedings of the 19th international Conference on Computational Linguistics - Volume 1* (Taipei, Taiwan, August 24 - September 01, 2002). *International Conference On Computational Linguistics. Association for Computational Linguistics*, Morristown, NJ, 1-8.
10. Kuzman Ganchev, Fernando Pereira, Mark Mandel, Steven Carroll, Peter White. Semi-Automated Named Entity Annotation. *LAW 2007*.
11. Katrin Tomanek and Fredrik Olsson. A Web Survey on the Use of Active Learning to support Annotation of Text Data. *Workshop on Active Learning for NLP at NAACL 2009*.
12. Roi Reichart, Katrin Tomanek , Udo Hahn, and Ari Rappoport. Multi-task active learning for linguistic annotations. In *ACL'08 - Proceedings of the 46th Annual Meeting of the Association of Computational Linguistics . Association for Computational Linguistics*, 2008.
13. Zhu and Hovy. Active Learning for Word Sense Disambiguation with Methods for Addressing the Class Imbalance Problem. *EMNLP 2007*.
14. Eric Ringger, Peter McClanahan, Robbie Haertel, George Busby, Marc Carmen, James Carroll, Kevin Seppi and Deryle Lonsdale. Active Learning for Part-of-Speech Tagging: Accelerating Corpus Annotation. *LAW 2007*.
15. Jason Baldridge and Alexis Palmer How well does active learning actually work? Time-based evaluation of cost-reduction strategies for language documentation, *EMNLP 2009*, Singapore.

16. Hwa, R., Resnik, P., Weinberg, A., Cabezas, C., and Kolak, O. 2005. Bootstrapping parsers via syntactic projection across parallel texts. *Nat. Lang. Eng.* 11, 3 (Sep. 2005), 311-325.
17. Riloff, E., Wiebe, J., and Wilson, T. 2003. Learning subjective nouns using extraction pattern bootstrapping. In *Proceedings of the Seventh Conference on Natural Language Learning At HLT-NAACL 2003 - Volume 4* (Edmonton, Canada). Human Language Technology Conference. Association for Computational Linguistics, Morristown, NJ, 25-32.
18. [Feiyu Xu](#), Daniela Kurz, [Jakub Piskorski](#), [Sven Schmeier](#).  
A Domain Adaptive Approach to Automatic Acquisition of Domain Relevant Terms and their Relations with Bootstrapping  
*Proceedings of the 3rd International Conference on Language Resources an Evaluation (LREC'02), May 29-31, Las Palmas, Canary Islands, Spain, 2002.*
19. a) Phrase Detectives: A Web-based Collaborative Annotation Game.  
Jon Chamberlain, Massimo Poesio, Udo Kruschwitz.  
b) ANAWIKI: Creating Anaphorically Annotated Resources Through Web Cooperation.  
Chamberlain, Poesio, Kruschwitz.  
c) (Linguistic) Science Through Web Collaboration in the ANAWIKI Project.  
Kruschwitz, Chamberlain & Poesio, 2009. *Proc. WebSci09.*, Athens.
20. Online Word Games for Semantic Data Collection.  
D. Vickrey, A. Bronzan, W. Choi, A. Kumar, J. Turner-Maier, A. Wang, and D. Koller.  
*EMNLP 2008.*
21. Open data collection for training intelligent software in the Open Mind Initiative.  
David G. Stork. *IEEE Expert Systems and Their Applications*. 2000, volume 14, 16-20.
22. Timothy Chklovski (2005), [1001 Paraphrases: Incenting Responsible Contributions in Collecting Paraphrases from Volunteers](#), Proceedings of AAAI Spring Symposium on Knowledge Collection from Volunteer Contributors (KCV05), Stanford, CA. 2005.
23. a) Push Singh (2001). [The Open Mind Common Sense project](#). *KurzweilAI.net*  
b) Push Singh (2002). [The public acquisition of commonsense knowledge](#). *Proceedings of AAAI Spring Symposium on Acquiring (and Using) Linguistic (and World) Knowledge for Information Access. Palo Alto, CA.*  
c) Push Singh, Thomas Lin, Erik T. Mueller, Grace Lim, Travell Perkins and Wan Li Zhu (2002). [Open Mind Common Sense: Knowledge acquisition from the general public](#). *Proceedings of the First International Conference on Ontologies, Databases, and Applications of Semantics for Large Scale Information Systems*. Irvine, CA.
24. Snow, R., O'Connor, B., Jurafsky, D., and Ng, A. Y. 2008. Cheap and fast---but is it good?: evaluating non-expert annotations for natural language tasks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (Honolulu, Hawaii, October 25 - 27, 2008). Annual Meeting of the ACL. Association for Computational Linguistics, Morristown, NJ, 254-263.
25. a) Luis von Ahn. Games with a Purpose.  
b) Siorpaes, K. and Hepp, M. 2008. Games with a Purpose for the Semantic Web. *IEEE Intelligent Systems* 23, 3 (May. 2008), 50-60.  
c) Elena Simperl and Katharina Siorpaes: [Human Intelligence in the Process of Semantic Content Creation](#) World Wide Web Journal (WWW), Springer, December 2009.