GATE and Social Media: References

Leon Derczynski
Kalina Bontcheva

References (A-D)

Attardi. 2010. Phratris - A Phrase Annotation Game. In INSEMTIVES Game Idea Challenge

Balahur, Steinberger 2009. Rethinking Sentiment Analysis in the News: from Theory to Practice and back. WOMSA

Baldwin, Lui 2010. Language Identification: The Long and the Short of the Matter. NAACL

Barbosa and Feng. 2010. Robust sentiment detection on Twitter from biased and noisy data. COLING '10.

Basave, Varga, Rowe, Stankovic, Dadzie 2013. Making sense of microposts (#MSM2013) concept extraction challenge. WWW MSM2013 workshop

Bontcheva, Derczynski, Funk, Greenwood, Maynard, Aswani 2013. TwitIE: An open-source information extraction pipeline for microblog text. RANLP

Bontcheva, Gorrell, Wessels 2013. Social Media and Information Overload: Survey Results.

Brody, Diakopoulos 2011. Cooooooooooooooollllllllllllll!!!!!!!!!!!!!: using word lengthening to detect sentiment in microblogs. EMNLP

Callison-Burch and M. Dredze, editors. 2010b. Proc. of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk .

Callison-Burch, Dredze. 2010a. Creating Speech and Language Data with Amazon's Mechanical Turk. In (Callison-Burch and Dredze, 2010b), pages 1–12.

Carter, Weerkamp, Sagkias 2013. Microblog language identification: Overcoming the limitations of short, unedited and idiomatic text. JLRE

Celli, Pianesi, Stillwell, Kosinski 2013. Workshop on Computational Personality Recognition: Shared Task. ICWSM

Crooks, Croitoru, Stefanidis, Radzikowski 2012. #Earthquake: Twitter as a Distributed Sensor System. T GIS 17(1):124

Derczynski, Maynard, Aswani, Bontcheva 2013. Microblog-genre noise and impact on semantic annotation. Hypertext

Derczynski, Ritter, Clark, Bontcheva 2013b. Twitter part-of-speech tagging for all: overcoming sparse and noisy data. RANLP

Derczynski, Yang, Jensen 2013. Towards Context-Aware Search and Analysis on Social Media Data. EDBT

Downey, Etzioni, Soderland 2010. Analysis of a probabilistic model of redundancy in unsupervised information extraction. AI 174(11):726

Dresner, Herring 2010. Functions of the non-verbal in CMC: Emoticons and illocutionary force. Communication Theory, 20(3):249

Drouin, Davis 2009. R u txting? Is the use of text speak hurting your literacy? Journal of Literacy Research, 41(1):46

References (E-K)

Eisenstein 2013. What to do about bad language on the internet. NAACL

Eisenstein, O'Connor, Smith and Xing 2010. Discovering sociolinguistic associations with structured sparsity. ACL

Finin, Murnane, Karandikar, Keller, Martineau, Dredze 2010. Annotating Named Entities in Twitter Data with Crowdsourcing. NAACL Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk

Finin, W. Murnane, A. Karandikar, N. Keller, J. Martineau, and M. Dredze. 2010. Annotating Named Entities in Twitter Data with Crowdsourcing. In Callison-Burch and Dredze (Callison-Burch and Dredze, 2010b), pages 80–88.

Fort, G. Adda, and K.B. Cohen. 2011. Amazon Mechanical Turk: Gold Mine or Coal Mine? Computational Linguistics , 37(2):413 –420.

Ganesan, Zhai, and Viegas 2012. Micropinion generation: an unsupervised approach to generating ultra-concise summaries of opinions. In Proceedings of the 21st International Conference on World Wide Web, WWW '12, pages 869–878.

Go, Bhayani, Huang 2009. Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford.

Halliday, Matthiessen 2004. An introduction to functional grammar.

Han, Baldwin 2011. Lexical Normalisation of Short Text Messages: Makn Sens a# twitter. ACL

Hangya & Farkas 2013. Filtering and Polarity Detection for Reputation Management on Tweets. CLEF 2013.

Hargittai, Neuman, Curry 2012. Taming the Information Tide: Perceptions of Information Overload in the American Home. The Information Society 28(3)

Hecht, Hong, Suh, Chi 2011. Tweets from Justin Bieber's Heart: The Dynamics of the "Location" Field in User Profiles. CHI

Hoffmann, L. 2009. Crowd Control. Communications of the ACM , 52(3):16 –17.

Hovy, Berg-Kirkpatrick,Vaswani, Hovy. 2013. Learning Whom to Trust with MACE. Proc. NAACL

Hsueh, Melville, Sindhwani. 2009. Data Quality from Crowdsourcing: A Study of Annotation Selection Criteria. In Proc. of the Workshop on Active Learning for Natural Language Processing , pages 27–35.

Hu, Talamadupula and Kambhampati 2013. Dude, srsly?: The Surprisingly Formal Nature of Twitter's Language. ICWSM

Ito, Tagliamonte 2003. Well weird, right dodgy, very strange, really cool. Language in Society 32(2):257

Ji and Grishman 2011. Knowledge Base Population: Successful Approaches and Challenges. ACL/HLT 2011

Jones 2010. The changing face of spelling on the internet.

Kwak, Lee, Park, Moon 2010. What is Twitter, a social network or a news media? WWW

References (L-O)

Labov 1972. Sociolinguistic patterns.

Lampos, Preotiuc-Pietro, Cohn 2013. A user-centric model of voting intention from Social Media. Proc. of ACL 51, 993–1003.

Laorden, Galan-Garcia, Santos, Sanz, Gomez Hidalgo, Garcia Bringas 2012. Negobot: A Conversational Agent Based on Game Theory for the Detection of Paedophile Behaviour. Advances in Intelligent Systems and Computing 189

Lawson, K. Eustice, M. Perkowitz, and M. Yetisgen-Yildiz. 2010. Annotating Large Email Datasets for Named Entity Recognition with Mechanical Turk. In Callison-Burch and Dredze (Callison-Burch and Dredze, 2010b), pages 71–79.

Li, Bontcheva, Cunningham 2009. Adapting SVM for data sparseness and imbalance: A case study in information extraction. JNLE 1(1):1

Liu 2012. Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers

Liu, Wang, Li, and Liu. 2010. Improving blog polarity classification via topic analysis and adaptive methods .HLT 2010.

Lui, Baldwin 2012. langid. py: An off-the-shelf language identification tool. ACL

Magnani, Montesi and Rossi 2012. Conversation Retrieval from Social Media. Information Retrieval Journal

Maynard 2014. Who cares about sarcastic tweets? Investigating the impact of sarcasm on sentiment analysis. LREC

Meij, Weerkamp & de Rijke 2012. Adding Semantics to Microblog Posts. In Proceedings of the 5th International Conference on Web Search and Data Mining (WSDM'12)

O'Connor, Balasubramanyan, Routledge, Smith: From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series. ICWSM 2010

O'Connor, Krieger, Ahn 2010. TweetMotif: Exploratory Search and Topic Summarization for Twitter. ICWSM

Ott, Choi, Cardie, and Hancock 2011. Finding deceptive opinion spam by any stretch of the imagination. ACL.

Owoputi, O'Connor, Dyer, Gimpel, Schneider, Smith 2013. Improved Part-of-Speech Tagging for Online Conversational Text with Word Clusters. NAACL

References (P-S)

Pak and Paroubek 2010. Twitter Based System: Using Twitter for Disambiguating Sentiment Ambiguous Adjectives. 5th International Workshop on Semantic Evaluation.

Parent and M. Eskenazi. 2011. Speaking to the Crowd: Looking at Past Achievements in Using Crowdsourcing for Speech and **Predicting Future Challenges.** In Proc. of INTERSPEECH , pages 3037– 3040.

Petrovic, Osborne, Lavrenko 2013. I Wish I Didn't Say That! Analyzing and Predicting Deleted Messages in Twitter. arXiv cs.SI 1305.3107

Plank, Hovy, Søgaard 2014. Learning POS taggers with inter-annotator agreement loss. EACL

Poesio, M., U. Kruschwitz, J. Chamberlain, L. Robaldo, and L. Ducceschi. 2012. Phrase Detectives: Utilizing Collective Intelligence for Internet-Scale Language Resource Creation. Transactions on Interactive Intelligent Systems.

Power, Robinson, Ratcliffe 2013. Finding Fires with Twitter. Australasian Language Technology Association Workshop

Preotiuc-Pietro, Samangooei, Cohn, Gibbins, Naranjan 2012. Trendminer: An architecture for real time analysis of social media data. RAMSS

Ramage, Hall, Nallapati, Manning 2009. Labeled LDA: a supervised topic model for credit attribution in multi-labeled corpora. EMNLP

Ritter, Clark, Mausam, Etzioni 2011. Named Entity Recognition in Tweets: An Experimental Study. EMNLP

Roth, Ratinov 2009. Design challenges and misconceptions in named entity recognition. CoNLL

Rout, Preotiuc-Pietro, Bontcheva, Cohn 2011. Where's @wally? A Classification Approach to Geolocating Twitter Users. Hypertext

Sakaki, Okazaki, Matsuo 2010. Earthquake shakes Twitter users: real-time event detection by social sensors. WWW

Scharl, Sabou, Gindl, Rafelsberger, Weichselbraun. 2012. Leveraging the wisdom of the crowds for the acquisition of multilingual language resources. Eight Int. Conf. on Language Resources and Evaluation Conference (LREC12)

Snow, R. B. O'Connor, D. Jurafsky, and A. Y. Ng. 2008. Cheap and Fast—but is it Good?: Evaluating Non-Expert Annotations for Natural Language Tasks. In Proc. of the Conference on Empirical Methods in Natural Language Processing (EMNLP'08)

Stede and C.R. Huang. 2012. Inter-operability and reusability: the science of annotation. Language Resources and Evaluation , 46:91–94. 10.1007/s10579-011-9164-x.

Stewart 2014. Now We Stronger Than Ever: African-American Syntax in Twitter. EACL

Sugumaran, Voss 2012. Real-time spatio-temporal analysis of West Nile Virus using Twitter Data. Proc. Computing for Geospatial Research and Applications

References (S-Z)

Taboada et al 2011. Lexicon-Based Methods for Sentiment Analysis. Computational Lingusitics. Vol. 37. Num 2.

Tagliamonte, **Denis** 2008. Linguistic ruin? lol! instant messaging and teen language. American Speech 83(1):3

Turian, Ratinov, Bengio 2010. Word representations: a simple and general method for semi-supervised learning. ACL

von Ahn and L. Dabbish. 2008. Designing games with a purpose. Commun. ACM , 51(8):58–67

Wang, Hoang, Kan. 2012. Perspectives on Crowdsourcing Annotations for Natural Language Processing. Language Resources and Evaluation

Wang, Wei, Liu, Zhou, and Zhang. 2011. Topic sentiment analysis in Twitter: a graph-based hashtag sentiment classification approach. CIKM '11.

Weikum, Ntarmos, Spaniol, Triantafillou, Benczur, Kirkpatrick, Rigaux, Williamson 2011. Longitudinal Analytics on Web Archive Data: It's About Time! CIDR

Xu, Bhargava, Nowak, Zhu 2012. Socioscope: Spatio-temporal signal recovery from social media. ECML PKDD

Yates 1996. Oral and written linguistic aspects of computer conferencing. Pragmatics and beyond

Zaidan and C. Callison-Burch. 2011. Crowdsourcing Translation: Professional Quality from Non-Professionals. In Proc. of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL:HLT'11), pages 1220–1229.