

COW14 Tokenization, Tagging, and Lemmatization

Roland Schäfer

Linguistic Web Characterization (DFG)
Freie Universität Berlin
email@domain

Felix Bildhauer

Grammar Department
Institut für Deutsche Sprache Mannheim
email@domain

Abstract

Ab stract.

1 Original Design Goals

(Schäfer and Bildhauer, 2012)

2 Implementation

3 Problems and Outlook

Acknowledgments

Roland Schäfer's contribution to this paper was funded by the German Research Council (Deutsche Forschungsgemeinschaft, DFG) through the personal grant SCHA1916/1-1.

References

Roland Schäfer and Felix Bildhauer. 2012. Building large corpora from the web using a new efficient tool chain. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, pages 486–493, Istanbul. ELRA.