



## Machine-Learning Aided Annotation Transfer from Sentence- To Token-Level

### Motivation

Many tasks require annotations on sentence AND on token level. This can e.g. be overall sentiment and phrase sentiment, overall argument stance and pro/con argument spans or intent and slots.

A large chunk of time during token- or character-level annotation is spent on deciding for the correct span, and inputting that span into an annotation system. Furthermore, reading, understanding, and remembering annotation guidelines which explain the correct annotation of spans imposes considerable mental load on annotators. The goal of this thesis is training machine learning models that help human annotators with making token-level annotations given already made sentence level annotations.

Topic: Death Penalty

It does not deter crime and  
it is extremely expensive to administer .

Topic: Gun Control

Yes , guns can be used for protection  
but laws are meant to protect us , too .

### Task Description

- Analyze which tasks would benefit from sentence to token-level annotation transfer
- Research existing methods, collect data, reimplement existing methods
- Implement an interactive simulation with simulated annotators
- If time, conduct a user study and test the implemented methods in the wild

### References

- Socher, Richard & Perelygin, A. & Wu, J.Y. & Chuang, J. & Manning, C.D. & Ng, A.Y. & Potts, C.. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. EMNLP
- Toprak, Cigdem & Jakob, Niklas & Gurevych, Iryna. (2010). Sentence and Expression Level Annotation of Opinions in User-Generated Discourse.
- Trautmann, Dietrich & Daxenberger, Johannes & Stab, Christian & Schütze, Hinrich & Gurevych, Iryna. (2020). Fine-Grained Argument Unit Recognition and Classification.

### Contact

Analysis



Programming



Literature



Prof. Dr. Iryna Gurevych

Jan-Christoph Klie

[thesis@ukp.informatik.tu-darmstadt.de](mailto:thesis@ukp.informatik.tu-darmstadt.de)