

# Semantic Role Labelling for Dutch Law Texts

## Motivation

- ▶ Legal texts are difficult to interpret
- ▶ Formalising them will make an interpretation explicit
- ▶ Creating formalisations is very labour-intensive
- ▶ **We propose a method for automated extraction of knowledge representations using two different NLP approaches**

## Related work

- ▶ Rule-based approaches with syntactic parsing [1]
- ▶ Flint frames as a framework for representing norms in acts and facts [2]
- ▶ State-of-the-art performance with BERT for semantic role labelling [3]

## Architecture



## Method

- ▶ Law text cleaning and preprocessing
- ▶ Law text annotation
- ▶ Finetuning BERTje (Dutch BERT)
- ▶ *Syntactic tagging + rules VS Semantic role labeling*
- ▶ Evaluation: performance and comparison

## Data

- ▶ 1854 unique sentences from law texts
- ▶ Annotated with four semantic roles: action, actor, object, recipient
- ▶ 409 sentences with actions, 324 actors, 337 objects, 85 recipients
- ▶ Inter-annotator agreement: Fleiss  $\kappa = .785$
- ▶ *The processor [actor] collects [action] personal data [object] of the data subject [recipient] (GDPR art. 5, simplified)*

## Results: Confusion matrices

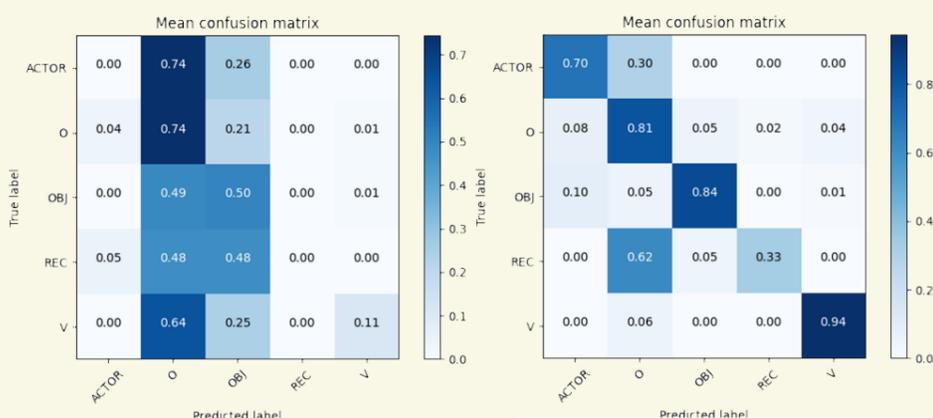


Figure: Rule-based matrix and Transformer-based matrix on test set

## Results: Mean accuracy per method and test set

| Dataset                | Rule | Transformer |
|------------------------|------|-------------|
| Annotated dataset      | 0.42 | -           |
| Test set               | 0.52 | 0.81        |
| Aliens Act (positives) | 0.42 | 0.80        |
| Aliens Act (all)       | 0.74 | 0.69        |

## Conclusion and Future Work

- ▶ The transformer-based method outperforms the rule-based method and is promising with an accuracy of 81%
- ▶ Future work: extraction of other elements such as recipients, improve the rules and combine both methods, test methodology on English law texts

## References

- [1] Ron van Gog and Tom M van Engers. Modeling legislation using natural language processing. In *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236)*, volume 1, pages 561–566. IEEE, 2001.
- [2] Robert van Doesburg and Tom M van Engers. Explicit interpretation of the dutch aliens act. In *Proceedings of the Workshop on Artificial Intelligence and the Administrative State co-located with 17th International Conference on AI and Law*, pages 27–37, 2019.
- [3] Peng Shi and Jimmy Lin. Simple bert models for relation extraction and semantic role labeling. *ArXiv*, abs/1904.05255, 2019.

