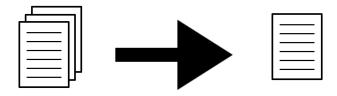
Automatic Detection of Linguistic Quality Violations

Jonathan Oberländer

Bachelor Thesis Defense Universität des Saarlandes 21.08.2014



- Single-Document: One document
- ▶ Multi-Document: Multiple documents on the same topic

- ▶ Single-Document: One document
- ▶ Multi-Document: Multiple documents on the same topic
- ▶ **Abstractive:** Internal semantic representation + generation
- **Extractive:** New summary from source sentences

- ▶ Single-Document: One document
- ▶ Multi-Document: Multiple documents on the same topic
- ▶ **Abstractive:** Internal semantic representation + generation
- **Extractive:** New summary from source sentences

	Single-document	Multi-document
Abstractive		
Extractive		

Linguistic Quality Violations

Friedrich et al. (2014)

detecting ungrammaticality

detecting ungrammaticality:subtypes

detecting ungrammaticality:unknowntokens

detecting ungrammaticality:unknowntokens

detecting ungrammaticality:UT_evaluation

detecting ungrammaticality:decision trees

detecting ungrammaticality:evaluation

detecting datelines

detecting datelines:method

detecting datelines:eval

detecting redundancies

detecting redundancies:method

detecting redundancies:eval

Friedrich, A., Valeeva, M., and Palmer, A. (2014). Lqvsumm: A corpus of linguistic quality violations in multi-document summarization.