



HATE SPEECH-DETEKTOREN IM JOURNALISMUS?

SYLVIA JAKI

UNIVERSITÄT HILDESHEIM

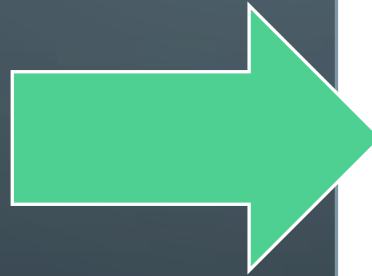
@SYLVIAJAKI

JAKISY@UNI-HILDESHEIM.DE

1. UNSERE HATE SPEECH-FORSCHUNG

**Automatisierte
Erkennung von Hate
Speech
(maschinelles
Lernen)**

**Quantitative und
qualitative Analyse
von Hate Speech in
sozialen Medien**



**Beispiel:
Hate Speech-Detektor für
deutschsprachige
rechtsradikale Tweets
(84%)**

Phase 1:

Datenauswahl Hate vs. Nicht-Hate



Phase 2:

**Algorithmus wird mit diesem Datenset trainiert =
Algorithmus analysiert Datenset eigenständig mit
Hilfe statistischer Verfahren**



Phase 3:

**Algorithmus kann bei neuem Datenset zwischen
Hate vs. Nicht-Hate unterscheiden**

2. WEM KÖNNTEN SOLCHE ALGORITHMEN NUTZEN?

- **Social Media-Plattformen**
- **Behörden**
- **Wissenschaft**
- **Journalismus**
- **...**

3. EINSATZMÖGLICHKEITEN IM JOURNALISMUS UND DAMIT VERBUNDENE HERAUSFORDERUNGEN

Einsatzmöglichkeit	Kommentarspalten	Datenjournalismus
Vorteile	<ul style="list-style-type: none">• Schnelle Beseitigung von Hate Speech, ohne dass jemand manuell große Datenmengen sichten muss	<ul style="list-style-type: none">• Schneller Überblick über große Datenmengen• Möglicherweise verbessertes Verständnis politisch motivierter Kommunikation in den sozialen Medien
Größte Herausforderungen	<ul style="list-style-type: none">• Technisch: Reliabilität• Ethisch: Entfernung nach welchen Grundsätzen?	<ul style="list-style-type: none">• Handhabbarkeit ohne Spezialwissen