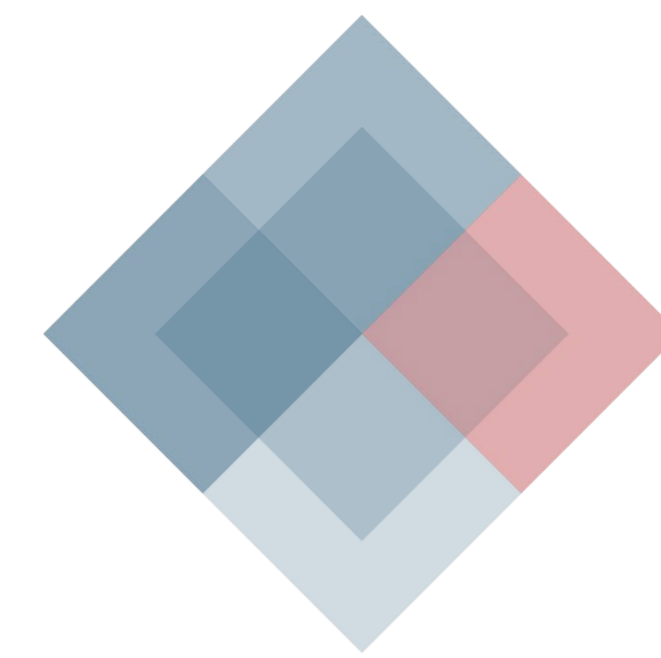


# Robuste Text-Einbettung und deren Visualisierung



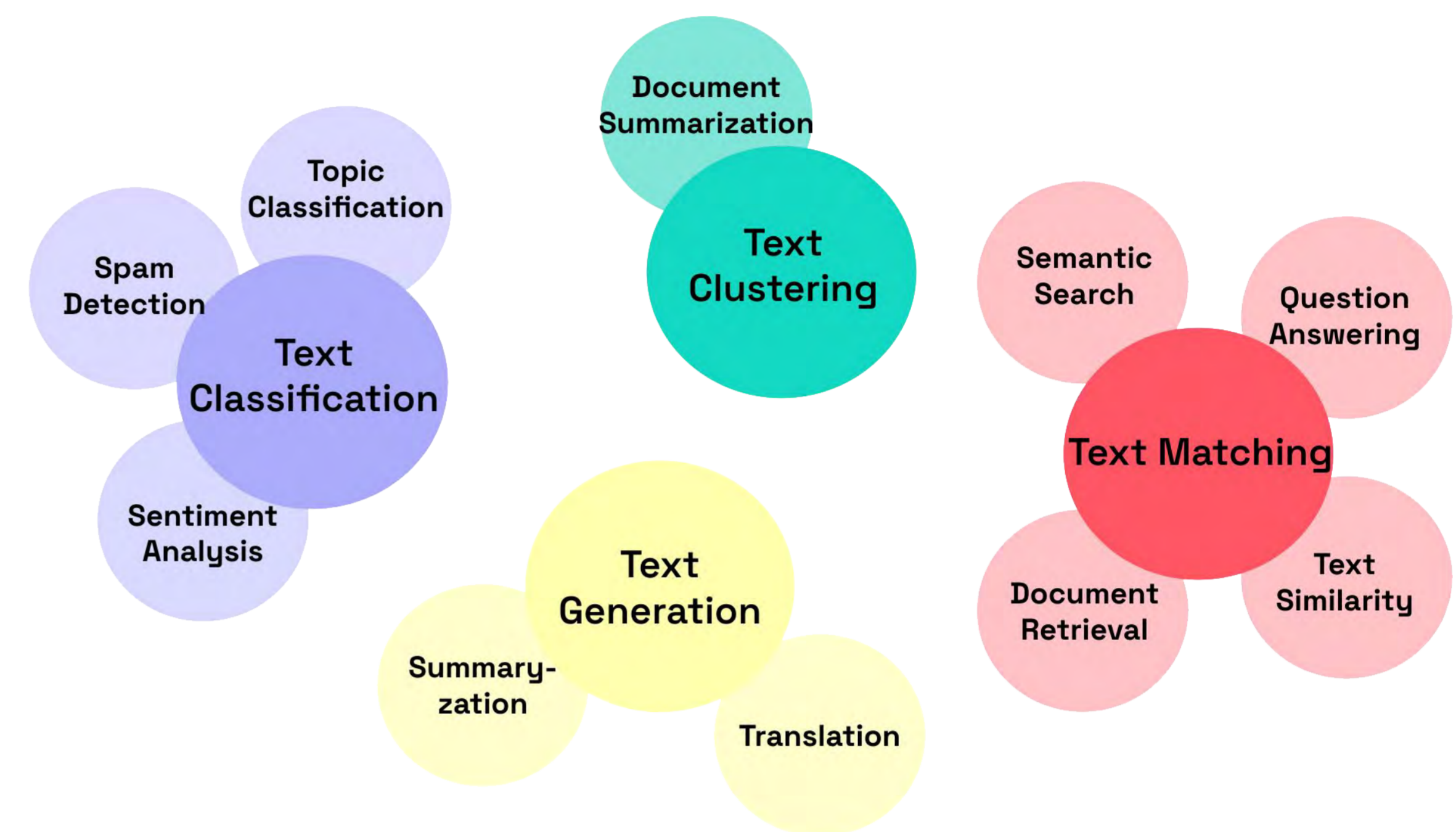
**ResKriVer**  
Resiliente Versorgungsnetze

Erlernen und Visualisieren von modernen Einbettungsverfahren zur Verarbeitung von verrauschten Social-Media-Textdaten



## Einführung - Wozu Text-Einbettungen?

- Deep-Learning-Methoden stellen ein mächtiges Tool zum Erlernen von allgemeinen Text-Repräsentationen dar.
- Einbettungen sind eine spezielle Repräsentationsform, bei der ein Neuronales Netz, zumeist ein Transformer, eine Eingabe-Textsequenz in einen Vektor verwandelt.
- Auf Basis dieser Vektoren können zahlreiche Textverarbeitungsaufgaben adressiert werden.
- Besonders relevant für das ResKriVer-Projekt sind robuste Einbettungen von verrauschten Texten (z.B. Tweets).



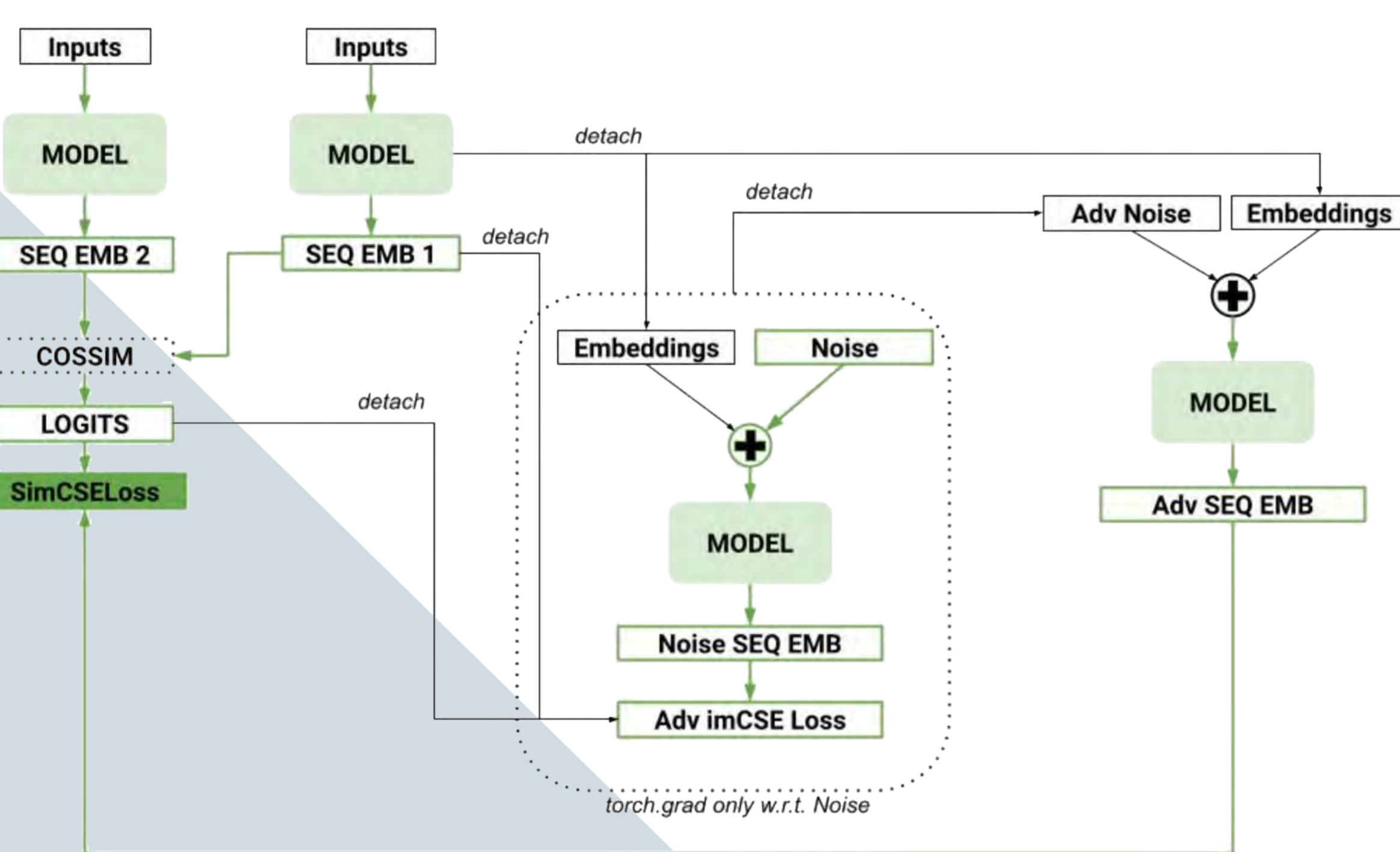
Beispiele von Verarbeitungsaufgaben, die mit Text-Einbettungen gelöst werden können.

## Die AdvCSE-Methode

- Neuartiger Ansatz inspiriert von SimCSE<sup>1</sup> & MixCSE<sup>2</sup>, basierend auf kontrastivem Lernen:



- Kernidee: **negative Beispiele** werden durch **adversarial Noise** im Training so modifiziert, dass der Encoder gezwungen wird, **robustere Einbettungen** zu erlernen.
- Wirksamkeit und schnellere Konvergenz wurde auf dem populären STS-B Benchmark-Datensatz nachgewiesen.



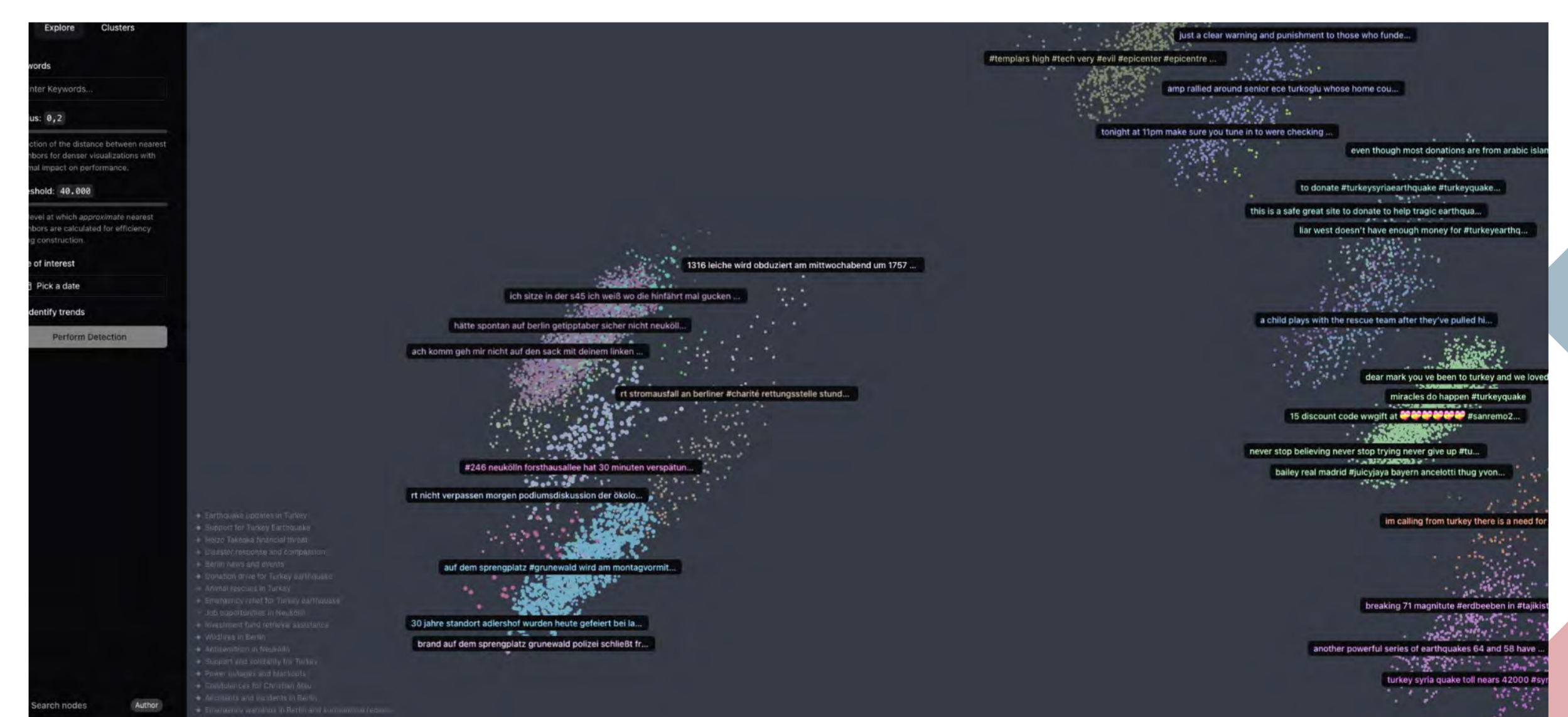
Schematische Darstellung der AdvCSE-Methode.

<sup>1</sup> T. Gao, X. Yao, and D. Chen. „Simcse: Simple contrastive learning of sentence embeddings.“ arXiv:2104.08821

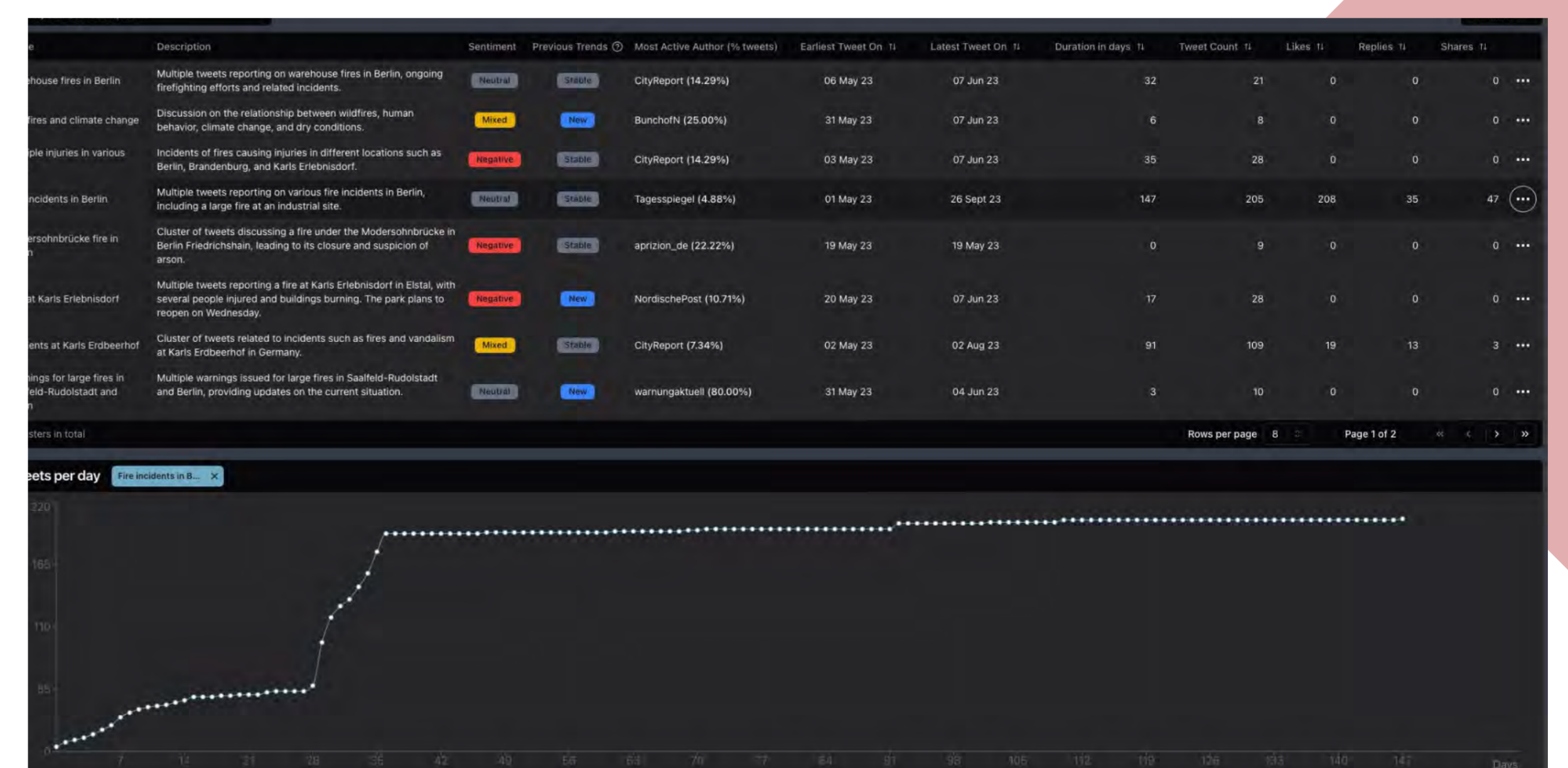
<sup>2</sup> Y. Zhang, R. Zhang, S. Mensah, X. Liu, and Y. Mao. „Unsupervised sentence representation via contrastive learning with mixing negatives“. AAAI 2022

## Der Embedding-Visualizer

Visualisierung der Einbettungen zur frühzeitigen Erkennung aufkommender oder bestehender Trends in sozialen Medien. Dies befähigt Nachrichtenagenturen, Regierungsbehörden oder Unternehmen, schnell und gezielt auf relevante Entwicklungen und Krisen zu reagieren. Durch seine Analysefunktionen trägt der Visualizer dazu bei, Trends deutlich zu identifizieren, zu differenzieren und zu analysieren. Auf diese Weise kann der Beginn eines Trends sowie seine Bestandteile und Unterschiede zu anderen Themenclustern besser verstanden werden.



Übersicht aller aktuellen Tweet-Cluster



Analyse einzelner Cluster: Detaillierte Trenduntersuchung und Vergleiche