Technische und rechtliche Fragen der Korpuserstellung

Maik Stührenberg (Bielefeld)

In vielen Bereichen linguistischer Forschung werden heute Korpora benötigt, also Sammlungen linguistischer Daten, die für eine bestimmte Forschungsfrage hin ausgewählt, annotiert und analysiert werden.

Die Erstellung von Korpora ist zeitaufwendig und benötigt darüber hinaus ein hohes Maß an Personaleinsatz, denn nur aus hochgütigen Daten und Annotationen lassen sich konkrete Schlüsse ziehen.

Im besten Fall lassen sich qualitativ hochwertige Korpora zur Untersuchung von mehr als einer Fragestellung einsetzen.

Das kann dann gegeben sein, wenn sowohl einige technische als auch rechtliche Fragestellungen bei der Erstellung des Korpus beachtet wurden.

In der Präsentation sollen diese technischen und rechtlichen Fragen beleuchtet werden.

Dazu gehören Fragen des Urheberrechts genau so wie die Auswahl geeigneter standardisierter Annotationsformate.