

Statistische Analyse der Schweizer Politlandschaft im Nachrichtendienst Twitter

Diese Arbeit nimmt sich dem Thema der Datenanalyse des Nachrichtendienstes Twitter an. Ziel der Arbeit ist, Personen der Schweizer Politszene zu erfassen und ihr Verhalten zu analysieren. Aus diesem Grund wurden ca. 500 Personen, die aktiv in der Schweizer Politik tätig sind, erfasst und deren Verhalten aufgezeichnet. Um Daten verarbeiten zu können, müssen sie zuerst aufgezeichnet werden. Zu Beginn der Arbeit wird aufgezeigt, wie diese Aufzeichnung möglich ist und wie die Verfasser dieser Arbeit die Aufzeichnung umsetzen. Für die späteren Analysen müssen die Daten ebenfalls aufbereitet werden. Diese Aufbereitung wird bereits in die Aufzeichnung integriert und die Daten werden so aufbereitet, dass sie den späteren Analysen optimal zur Verfügung stehen.

Als Erstes widmet sich die Arbeit dem Thema "Topic Modelling". Dieses Verfahren ermöglicht es, aus Texten, anhand eines Vokabulars, vorkommende Themen zu bestimmen. Es wird versucht, Themen zu lokalisieren, über die in der Schweizer Politszene gesprochen wird. Aus den Analysen folgern wir, dass es enorm schwierig ist, aus Twitterdaten Themen herauszufiltern. Dies liegt sehr wahrscheinlich daran, dass Tweets relativ kurz sind und vor allem Kurzkonversationen darstellen.

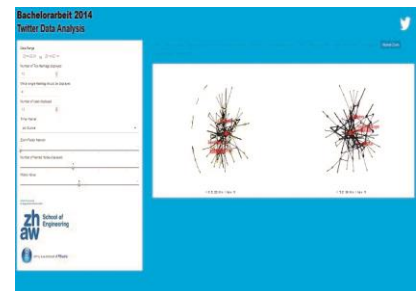
Eine weitere Analyse der Arbeit stellt die Netzwerkanalyse dar. Es wird versucht, die ca. 500 erfassten Personen in einem Netzwerk darzustellen. Dieses Vorgehen zeigt klare Muster des Kommunikationsverhaltens und zeigt auf, wie sich Communities über einen Hashtag aufbauen. Darüber hinaus wird eine Rangliste erstellt, die die wichtigsten Benutzer im Netzwerk lokalisiert. Des Weiteren werden Hashtags, die unsere ausgewählte und beobachtete Zielgruppe benutzt, genauer beobachtet. Es werden verschiedene Muster erkannt, die den zeitlichen Verlauf von Hashtags aufzeigen und so deren Eigenschaft genauer beschreiben. Weiter werden andere Analysen aufgezeigt, die Hashtags im Netzwerk hervorheben.

Die Netzwerk- und Hashtaganalysen könnten in Zukunft noch weiter verfeinert und automatisiert werden. Das Topic Modelling müsste in einem grösseren Rahmen angewendet werden. Die hier erfassten Daten haben in diesem Bereich nicht das gewünschte Resultat gezeigt. Es wäre jedoch unter Umständen möglich, mit einer anderen Flughöhe und anderen Verfahren bessere Ergebnisse zu erzielen. Diese hätten jedoch den Rahmen dieser Arbeit gesprengt und werden als weiterführende Arbeit erwähnt.

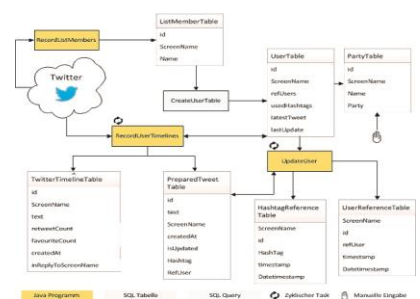


Diplomierende
Kilian Kessler
Pascal Stäheli

Dozierende
Christoph Hofer
Oliver Dürr



Im Verlauf der Arbeit wurde die abgebildete Shiny-Applikation in R programmiert. Diese Applikation erlaubt es dem Benutzer diverse Analysen auf den selbst aufgezeichneten Daten durchzuführen.



In diesem Bild wird das Vorgehen während der Aufzeichnung sichtbar. Die Daten wurden live aufgezeichnet und in eine Datenbank abgelegt. Von dort greift Shiny auf die Daten zu.