

Die Hanse Sail 2016 auf Twitter

- Nutzung von Geolokalisation in
Sozialen Netzwerken im
kleinräumigen Maßstab -

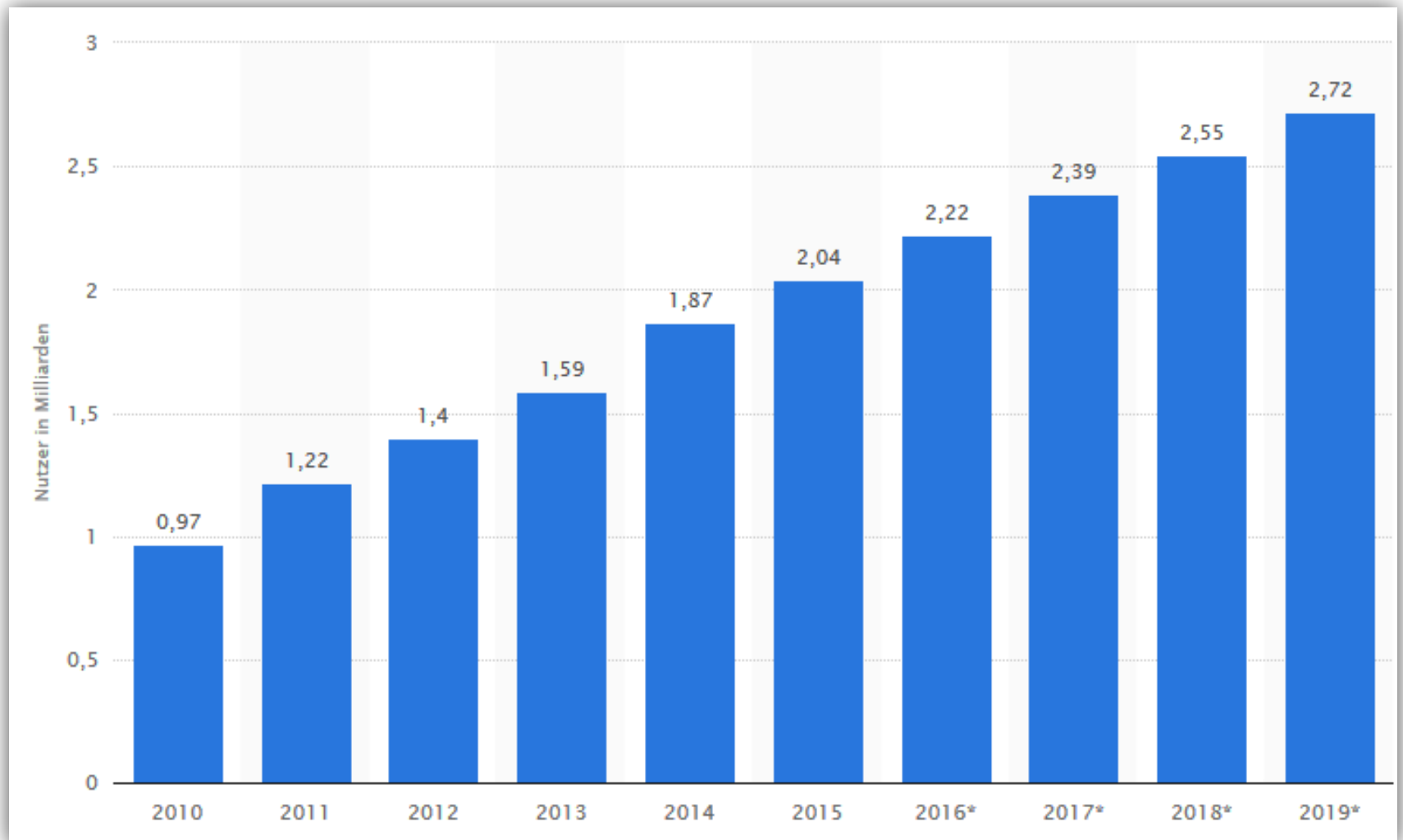
F. Vettermann, C. Seip & R. Bill

Universität Rostock

Agrar- und Umweltwissenschaftliche Fakultät

Professur für Geodäsie und Geoinformatik

- Vorbetrachtungen
- Ziele und Herausforderungen
- Umsetzung
- Ergebnisse
- Zusammenfassung und Ausblick



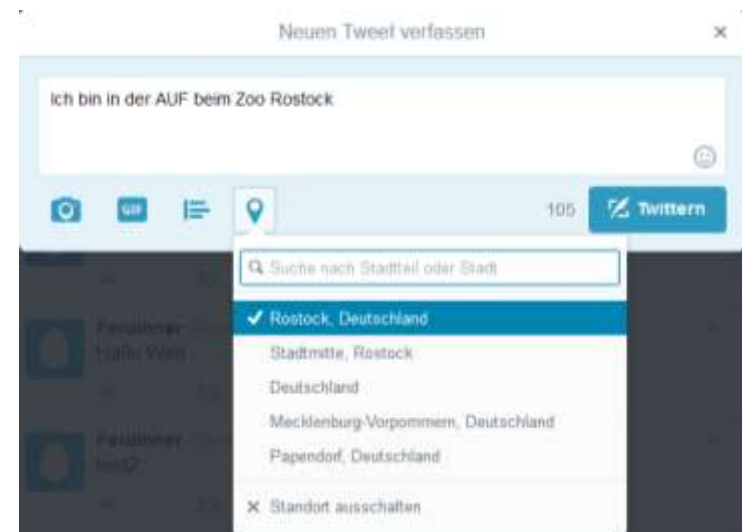
Prognostizierte Anzahl der Nutzer von Sozialen Netzwerken (STATISTA 2015)

● Twitter

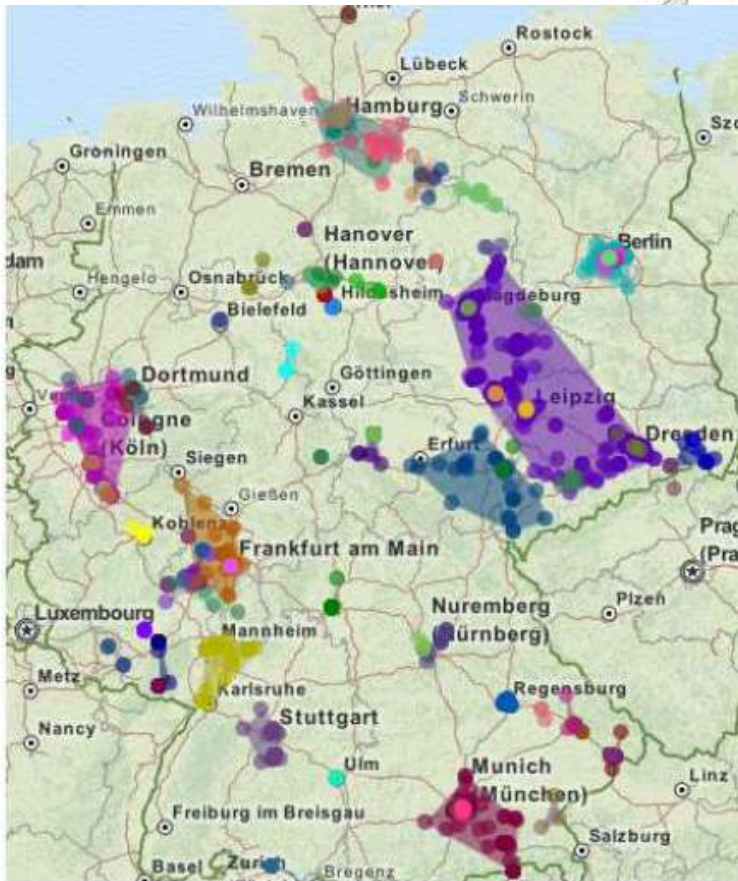


- 2006 durch Jack Dorsey gegründet
- Auf 140 Zeichen beschränkter Mikroblogging-Dienst
- “To give everyone the power to create and share ideas and information instantly, without barriers“
(TWITTER INC. 2016)
- 320 Mio. Monthly Active Users; 500 Mio. Nachrichten pro Tag

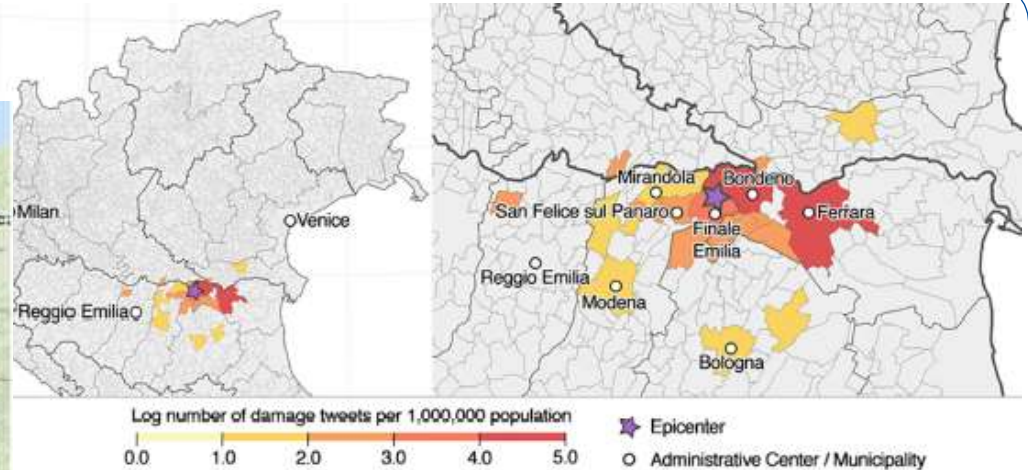
 **JULIAN ROBLING@** @julianrobling · 8 Min.
@just_maykeyy - Hi Maike, danke Dir für's Folgen hier auf Twitter :) Viele Grüße von München nach **Rostock** :) Jules



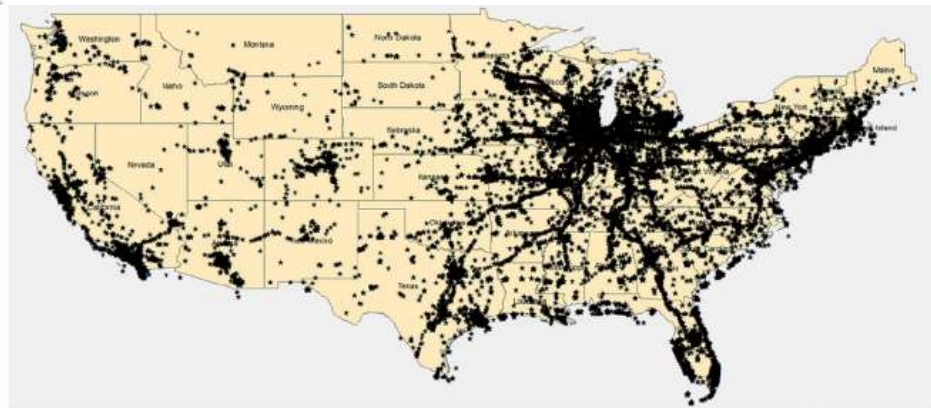
● Aktuelle Anwendungen



Spatio-temporale Verteilung von Tweets zur Elbe-Flut 2013 (FUCHS et al. 2013)



Abschätzung des Schadens beim Erdbeben 2012 in Norditalien via Twitter (CRESCI et al. 2015)



Aufenthaltort in Chicago beheimateter Twitter-Nutzer (LUO et al. 2016)

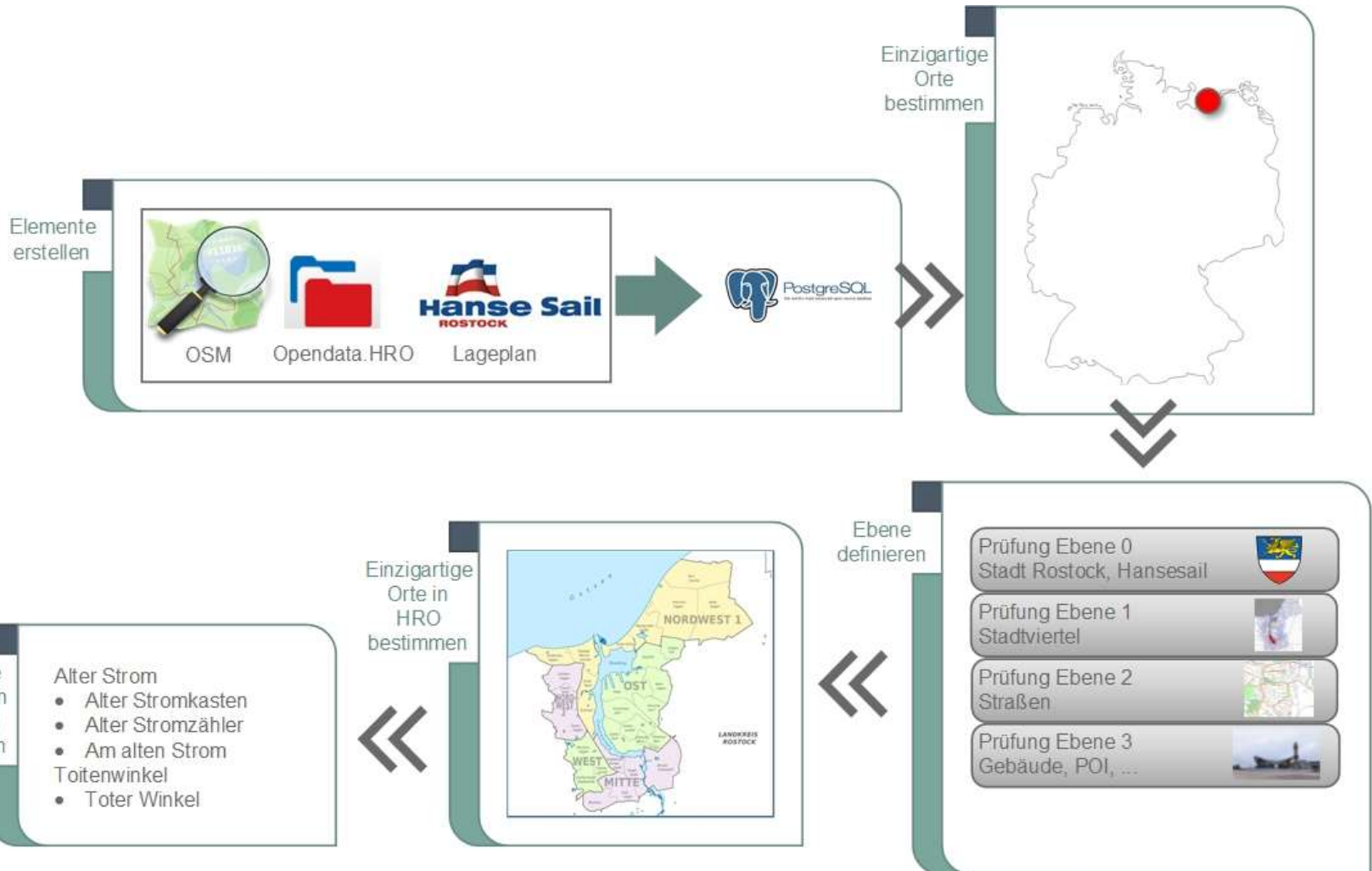
- Twitter kann in den folgenden Bereichen unterstützende Informationen liefern:
 - Hazardforschung
 - Wetterinformationen
 - Werbung und Marketing
 - Planung
 - Stadtgeographie

→ Alle diese Bereiche benötigen genaue Informationen über den **Aufenthaltort**, den **Bezugspunkt** oder die **Herkunft** der jeweiligen Person

- Methoden zur Lokalisationsbestimmung umfassen:
 - Ableitung aus **Geotag**
 - Datenbasis umfasst nur 2 % aller Nachrichten
 - Ableitung aus **Nutzerinformation** wie Wohnort, Zeitzone, ...
 - Bezug auf Regionen oder Städte (HECHT et al. 2011)
 - Ableitung aus **Text**
 - Location Indicative Words (LIWs); Fehler: 479 km - 209 km (EISENSTEIN et al. 2010, HAN et al. 2014)
 - **Kombinierte** Ansätze
 - Nutzung von Links, Text, Nutzerprofil, Zeitzonen, etc.; Fehler: 30 km – 4,5 km (SCHULZ et al. 2013, LAYLAVI et al. 2016)

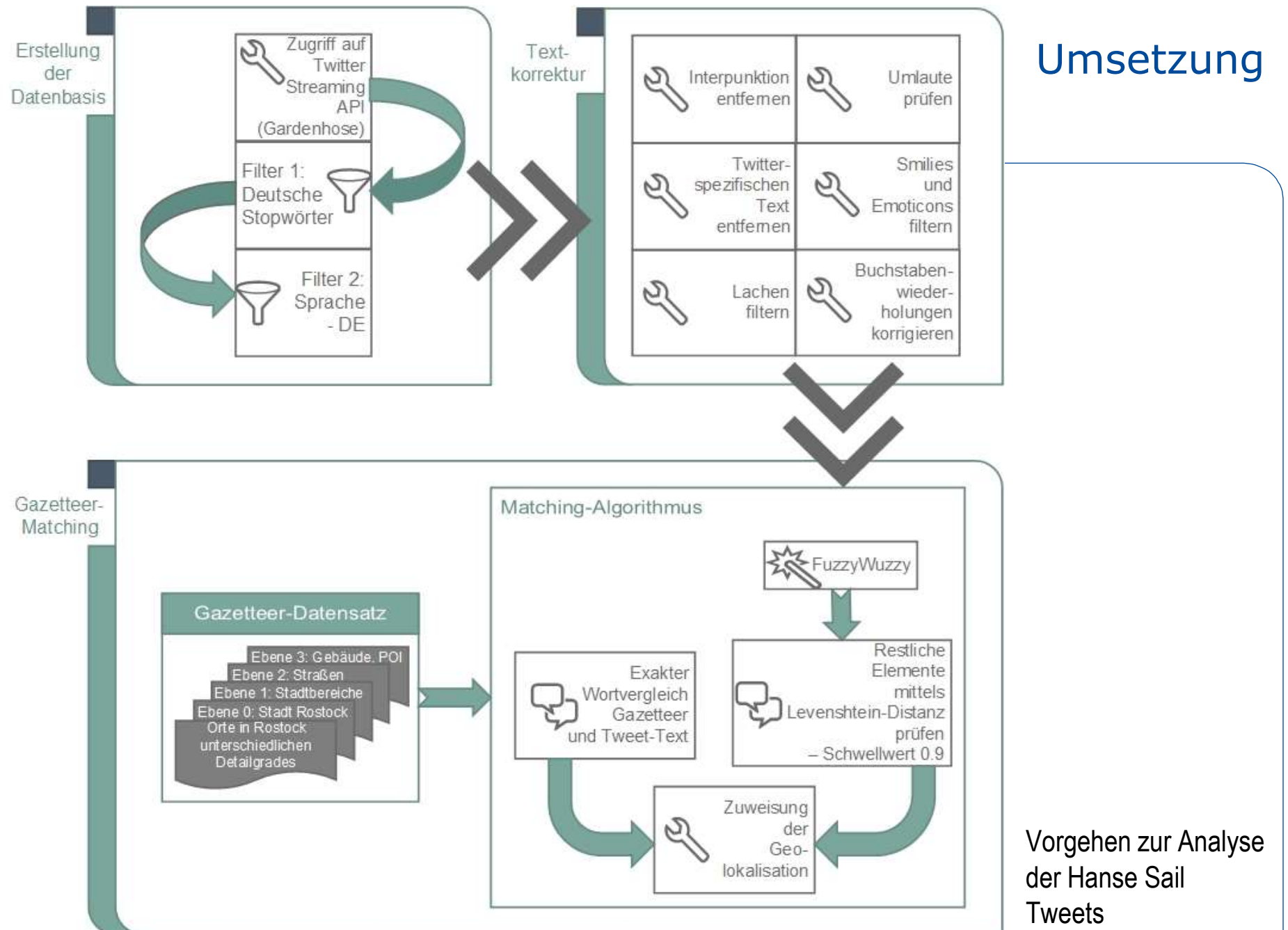
1. Sammeln einer möglichst großen Anzahl geolokalisierbarer Tweets im Zeitraum der Hanse Sail.
 - Geringes Nachrichtenaufkommen
 - Geringer Anteil an Lokalisationsinformation (< 2%)
 - Ungenaue / Fehlerhafte Ortsangaben der Nutzerherkunft→ Eine Erweiterung der Datenbasis durch die Analyse des Nachrichtentextes selbst ist notwendig.
2. Mindestens ortsteilgenaue Verortung der Nachrichten.
 - Unstrukturierte Sprache
 - Aufbau eines Gazetteers
 - Algorithmen zum Text-Vergleich (Levenshtein-Distanz, Jaro-Winkler-Distanz, ...)
3. Detektion von Hot-Spots und einzelnen, interessanten Orten.
 - Graphische Aufbereitung

- Python zum Zugriff auf die Twitter Streaming API und zum Schreiben in eine PostgreSQL-DB
 - Twython
 - psycopg2
- Sammeln aller deutschsprachigen Nachrichten zwischen 10.08.2016 und 16.08.2016 (Methodisches Vorgehen nach SCHEFFLER o.J.)
 - Stop-Word-Liste
 - Sprachfilter Deutsch
- Aufbau eines Gazetteers
- Textkorrektur
- Gazetteer-Matching
 - FuzzyWuzzy
 - Python-Levenshtein

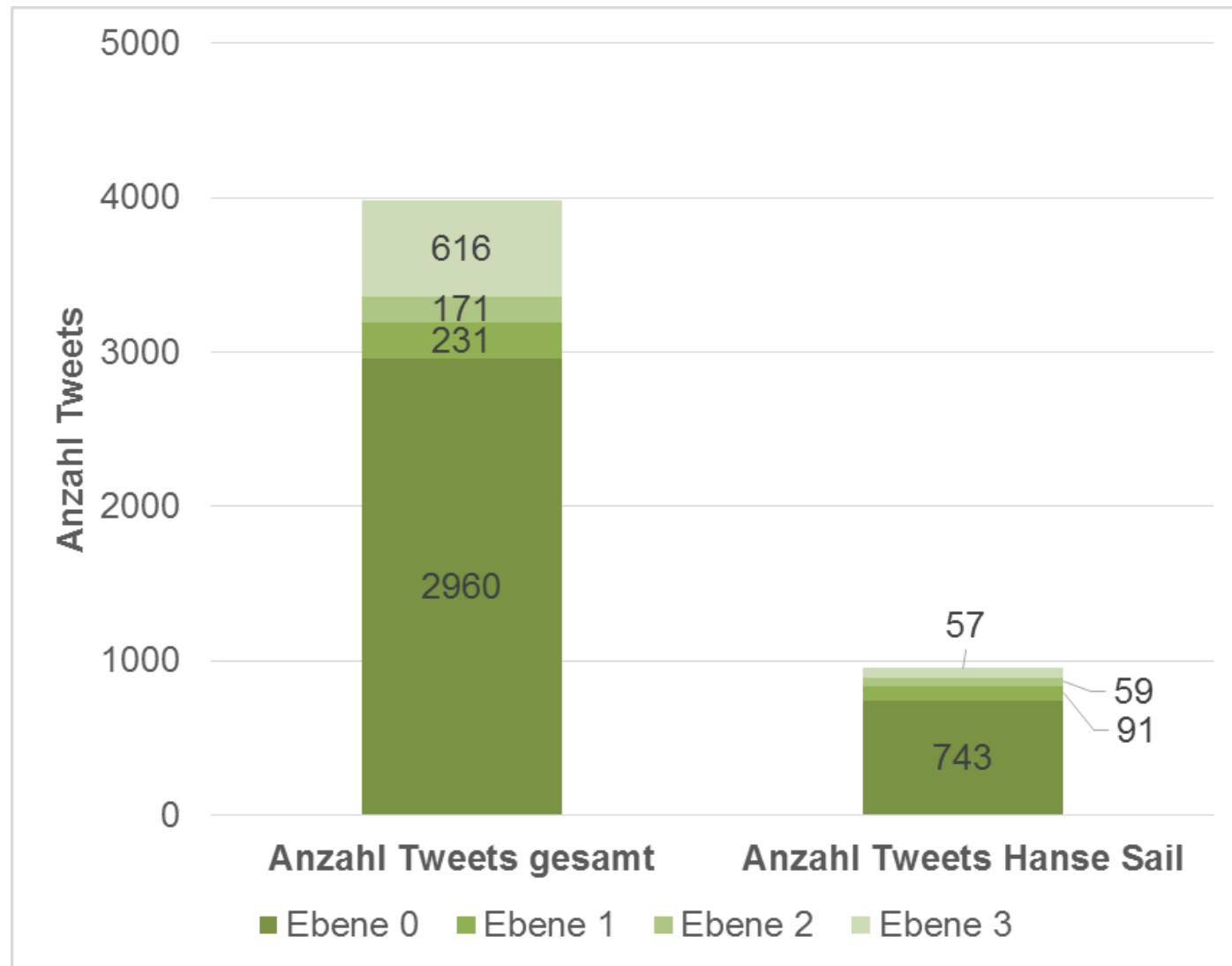


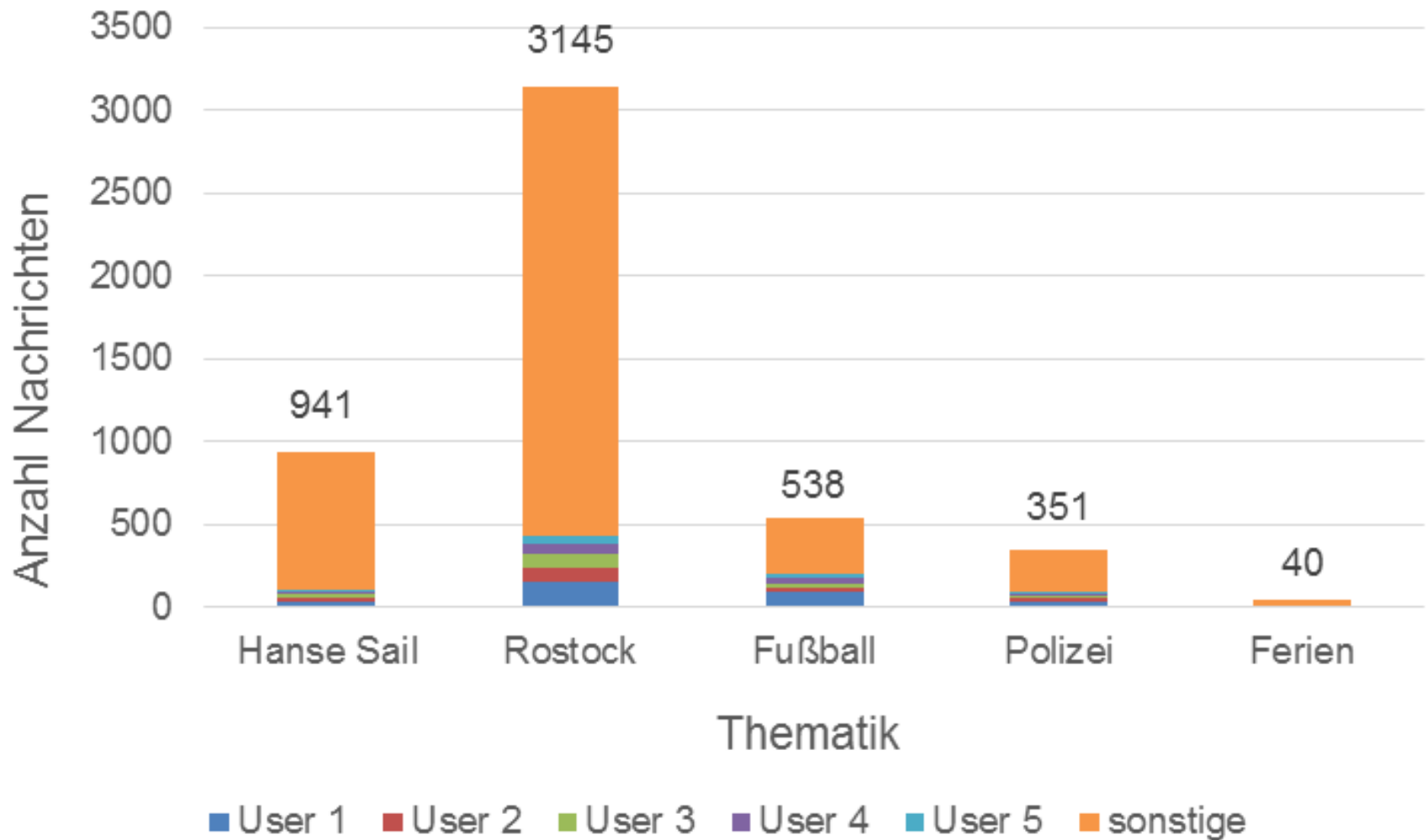
Aufbau des Gazetteers

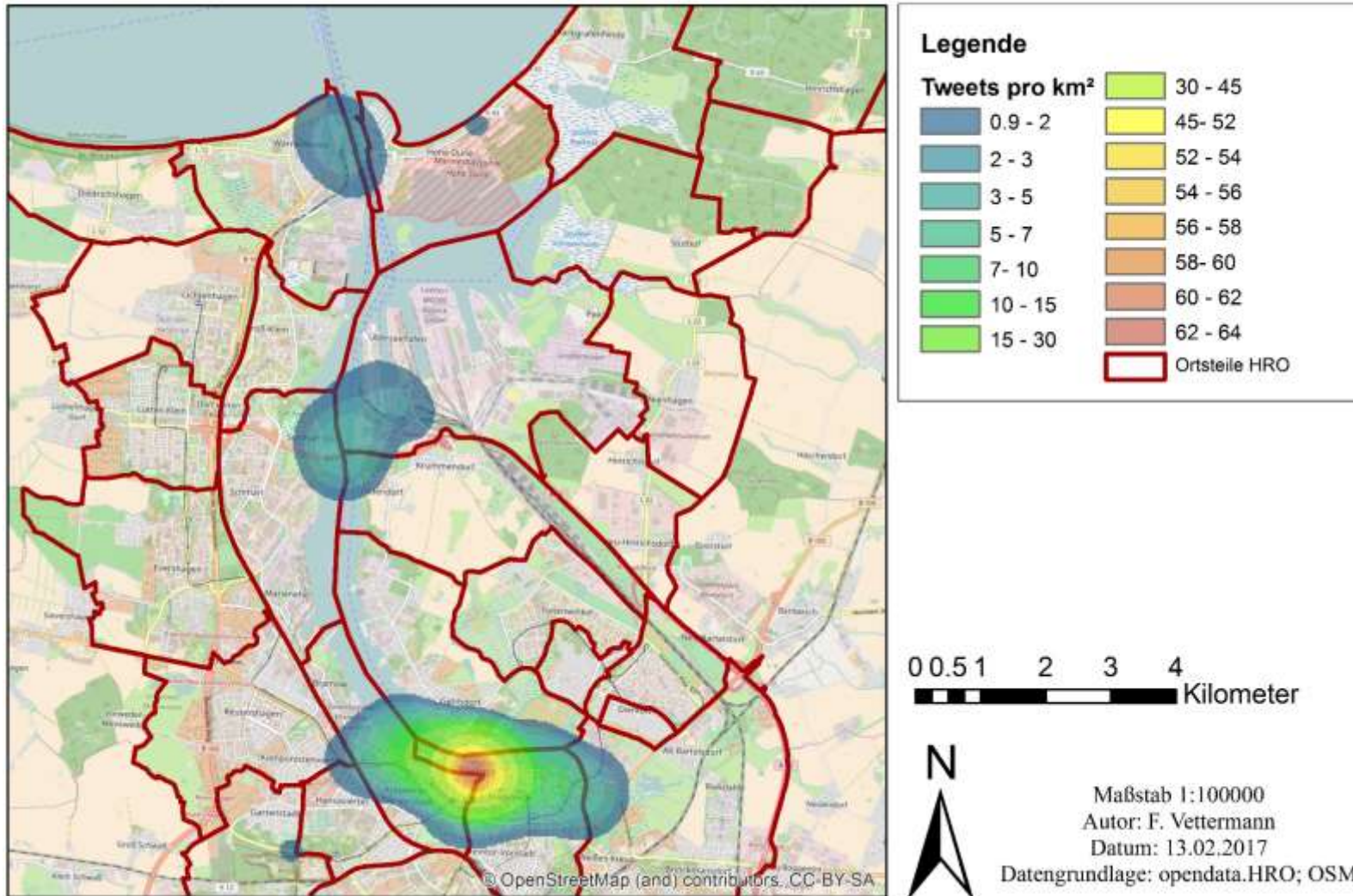
Umsetzung



Vorgehen zur Analyse der Hanse Sail Tweets







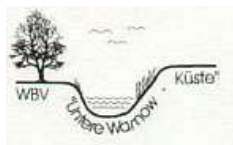
- Erstellung eines Modells zur kleinräumigen Ortserkennung
 - Textanalyse
 - Identifikation von Hot Spots
 - Identifikation besonderer Schiffe, Bühnen, etc.
 - Kein Zusammenhang zwischen Besucherzahl und Nachrichten erkennbar
 - Hochgenaue Ortslokation aus deutschsprachigen Nachrichten möglich
 - Ausreichende Nachrichtenanzahl in ländlichen Räumen und kurzen Zeiträumen
 - Schwächen im Gazetteer
- Soziale Medien können gute Ergänzung im Rahmen der Stadt- und Veranstaltungsplanung sowie Stadtgeographie darstellen

- Exakte Fehlerbestimmung
- Einbeziehung der Nutzerangaben
- Bestimmung der Besucherherkunft
- Zeitliche Aufschlüsselung
- Sentimentanalyse
- Erweiterung des Gazetteers
- Live-Darstellung

Wir danken allen beteiligten Projektpartnern sowie insbesondere dem BMBF zur Förderung des Projektes KOGGE.



Unsere Partner:



Staatliches Amt für Landwirtschaft und Umwelt
Mittleres Mecklenburg



Neuen Tweet verfassen



Vielen Dank für Ihre Aufmerksamkeit!



GIF



80



Twittern