



Automatische Erkennung von Trends

Sprachtechnologische Auswertung von Twitter-Daten zu Politikerinnen und Politikern

Bericht über ein Projekt im Master Informationswissenschaft

Studierende: Jennifer Deuschle, Barbara Lenze, Marina Petrovic und Sascha Starker

unter der Leitung von Prof. Dr. Melanie Siegel

Fragestellung

- Erheblicher Aufwand für Meinungsforschung
 - Telefoninterviews
 - Straßeninterviews
- Kann man diesen Prozess automatisieren?
- Kann man den Meinungstrend der Bevölkerung automatisch erfassen?
- Kann Twitter als Quelle für Meinungsforschung dienen?

Ist Twitter vielleicht sogar eine bessere Quelle?

- ▶ Krise der Meinungsforschung:
 - ▶ „Einen Tag vor dem EU-Referendum waren sich die Meinungsforscher einig: Der Brexit fällt aus!“ (<http://www.zeit.de/politik/2016-06/umfragen-brexit-us-praesidentschaftswahl-demoskopie-einfluss>)
 - ▶ „Nur vier von 67 Umfragen sahen zuletzt Donald Trump vorn – für die Meinungsforscher eine Blamage.“ (<http://www.morgenpost.de/politik/article208683451/Wie-Trumps-Wahlsieg-die-Demoskopen-in-die-Krise-stuerzt.html>)
- ▶ Telefonumfragen über Festnetz zwischen 9 und 16 Uhr sind nicht mehr repräsentativ


Idee

- ▶ Analyse von Twitter-Beiträgen zu Politikern
- ▶ Vergleich mit dem ZDF-Politbarometer


Startseite Über uns #steinmeier

← ↻ 7 ❤️ 11 ⋮


← Als Antwort an DerWesten

 **Sakölabo** @Sakoelabo · 2 Std.
.@DerWesten Der Mann, der Kurnaz in Guantanamo sitzen ließ, der bei Drohneinsätzen und NSA alles weglächelt? #Steinmeier geht gar nicht!


← ↻ 1 ❤️ 2 ⋮

 **Klaus Hammer** @klaus_hammer · 2 Std.
Ganz Ehrlich: #Gauck war schon nicht mein Präsident, aber #Steinmeier ist ein Schlag ins Gesicht!
Warum? z.B.: [google.de/search?q=stein...](https://www.google.de/search?q=steinmeier)

← ↻ 2 ❤️ ⋮

 **Oliver Das Gupta** @oliverdasgupta · 3 Std.
Hat sich Frank-Walter #Steinmeier jemals beim Bremer Murat Kurnaz gemeldet wegen #Guantanamo?

← 2 ↻ 2 ❤️ 3 ⋮

 **Julian Röpcke** @JulianRoepcke · 27. Okt.
Russischer Angriff mit 22 getöteten Kindern & 6 Lehrern war „tödlichste Attacke auf Schule seit Kriegsbeginn“ - UN

#Steinmeier schweigt.

← 1 ↻ 19 ❤️ 12 ⋮



Stand der Forschung

- ▶ Twitter seit 2006
- ▶ automatische Sentiment-Analyse seit ca. 15 Jahren
- ▶ automatische Sentiment-Analyse auf Twitter-Daten seit ca. 2009
 - ▶ sehr aktives Feld, regelmäßige Shared Tasks
 - ▶ bezogen auf die englische Sprache
- ▶ Tumasjan et al. (2010): Vorhersage von deutschen Wahlergebnissen anhand der Anzahl an Tweets
 - ▶ Tweets wurden aber für die automatische Klassifikation ins Englische übersetzt
- ▶ Analyse von deutschsprachigen Tweets erst seit ca. 2-3 Jahren
 - ▶ aktuell sehr aktives Feld
 - ▶ annotierter deutscher Twitter-Korpus vorgestellt von Uladzimir Sidarenka auf der LREC 2016

Twitter-Crawler

```
Python 3.5.1 Shell
File Edit Shell Debug Options Window Help
Python 3.5.1 (v3.5.1:37a07cee5969, Dec 6 2015, 01:38:48) [MSC v.1900 32 bit (Intel)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
RESTART: C:\Users\Melanie\Documents\01_Lehre_Darmstadt\06_alte Veranstaltungen\Opinion Mining\2015_WS\twitter_crawler.py
Nach welchem Politiker soll gesucht werden?
Künast
@winnipuh3351 tweeted: @DerSPIEGEL #Künast Erfahrung mit Handwerkern: Drücken sich gern derbe aus, meinen es nicht so. Ich leitete eine gew. Ausbildungsstätte.

@Felix_Huetten tweeted: Vielleicht sollten wir alle wieder mehr miteinander - und weniger übereinander sprechen. #künast #spiegel

@scholl_winfried tweeted: RT @DerSPIEGEL: "Künast, Du ekliges Wesen" - @RenateKuenast liest einige der vielen Hass-Postings vor, die sie täglich erreichen https://t...

@hartig_philipp tweeted: @NdKtweets @hoppenina @RenateKuenast @skoldehoff @DerSPIEGEL @BrittaStuff P.R. ist ok, wenn Künast das auch ohne macht. Das war meine Frage

@afd_bund_2017 tweeted: RT @ShakRiet: Es fehlen hier noch #Künast als Ministerin für vegane Ernährung und #Kipping im neuen Ressort Gendergerechtigkeit. #r2g https://t...
```


Sentiment-Analyse

- Fragestellung:
 - Enthält der Tweet eine Meinungsäußerung zum Politiker?
 - Ist diese Meinungsäußerung positiv oder negativ?

Automatische Analyse

Erstellung von Wortlisten

- Wörter, die positive oder negative Meinung ausdrücken, z.B. „gut“ oder „schlecht“

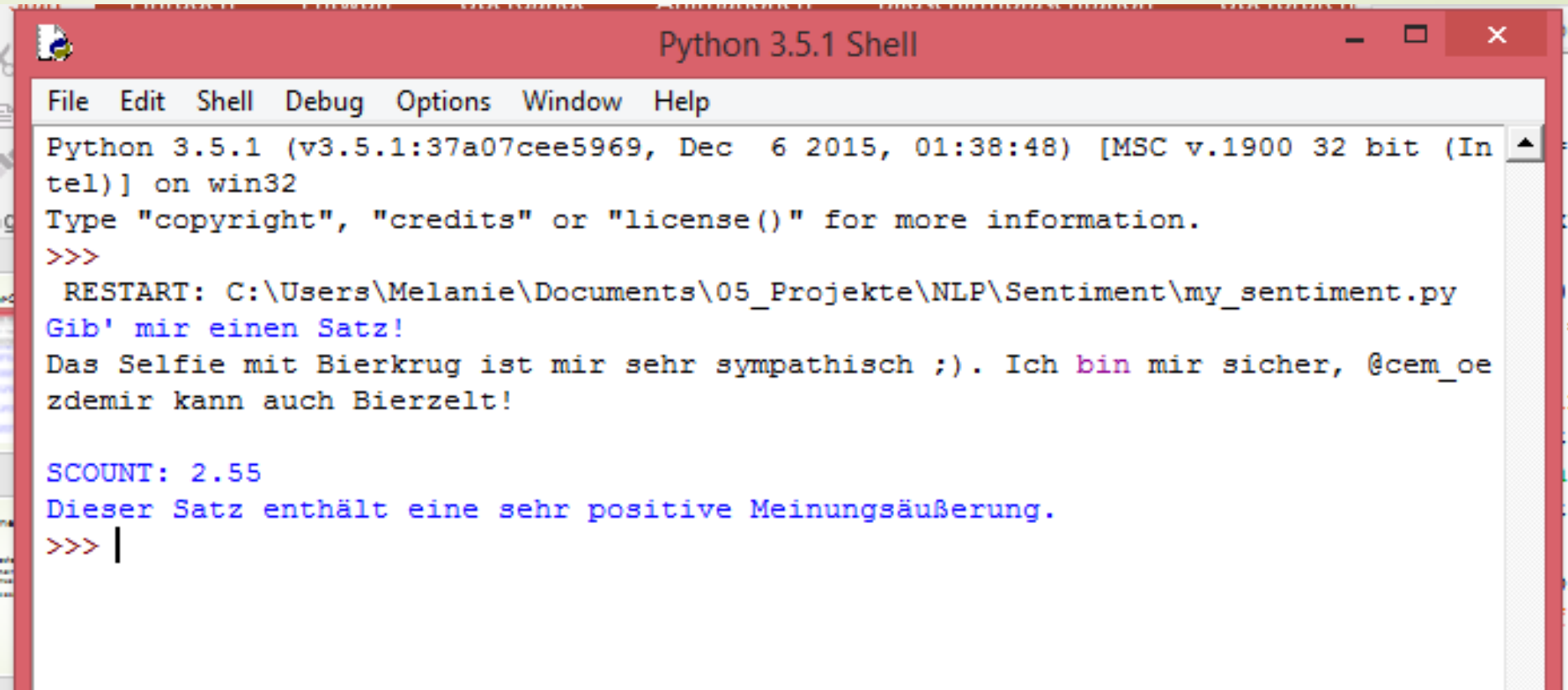
Lemmatisierung

- „Grundform“
- wichtig v.a. bei Verben

Verstärker und Negationen erkennen

- „sehr“ oder „nicht“

Für Twitter-Meinungsäußerungen implementiert



```
Python 3.5.1 Shell
File Edit Shell Debug Options Window Help
Python 3.5.1 (v3.5.1:37a07cee5969, Dec 6 2015, 01:38:48) [MSC v.1900 32 bit (Intel)] on win32
Type "copyright", "credits" or "license()" for more information.
>>>
  RESTART: C:\Users\Melanie\Documents\05_Projekte\NLP\Sentiment\my_sentiment.py
  Gib' mir einen Satz!
  Das Selfie mit Bierkrug ist mir sehr sympathisch ;). Ich bin mir sicher, @cem_oe
  zdemir kann auch Bierzelt!

  SCOUNT: 2.55
  Dieser Satz enthält eine sehr positive Meinungsäußerung.
  >>> |
```

Funktioniert auch für viele Tweets

- Das Selfie mit Bierkrug ist mir sehr sympathisch ;). Ich bin mir sicher, @cem_oezdemir kann auch Bierzelt!
- @welt bleiben Sie standhaft , Herr Özdemir
- @welt özdemir will die islamisierung Deutschlands,nach eigener Aussage.
- Nach den Tiraden von Horst Seehofer , würde ich keinenfalls da auftreten, allein aus Sicherheitsgründen!
- Empörung über EU-Kommissar: Oettinger verteidigt "Schlitzaugen"-Rede Jetzt wird es peinlich...
- Da können einem schonmal die Ressentiments durchsaloppieren... #oettinger

Viele Twitter-Beiträge sehen aber so aus

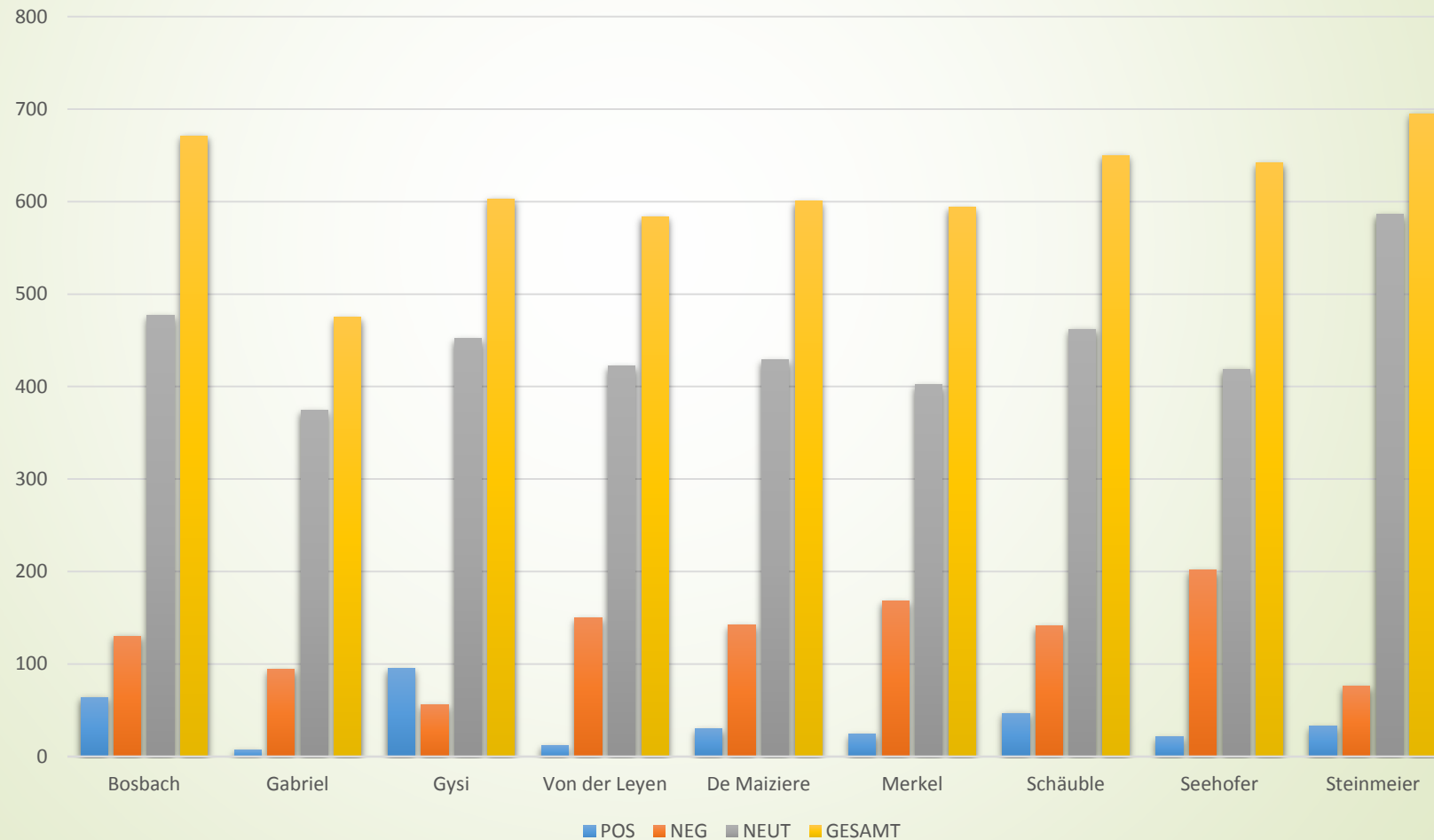
- ▶ Oezdemir: "Der nächste Verkehrsminister darf nicht mehr von der CSU aus Bayern kommen." Dem ist nichts mehr hinzuzufügen. #Urwahl
- ▶ @welt Oh da bewegt sich Seehofer auf sehr dünnem Eis , dasieht nach ein Verpflechtung der Vergangenheit aus.
- ▶ Es ist enorm wichtig, die Gemeinsamkeiten zu benennen. Von Seehofer bis Stiefelnazis. Selbe Argumente, selbe Ziele, selbe Scheiße
- ▶ Dann stellt #Seehofer wieder Forderungen beim #CSU_Partaitag und knickt vor #Merkel in #Berlin ein!
- ▶ #Oettinger wird niemals eine Anästhesistin brauchen. Der Mann ist echt absolut schmerzbehaftet.

Manuelle Annotation von Twitter-Beiträgen zu Politikern

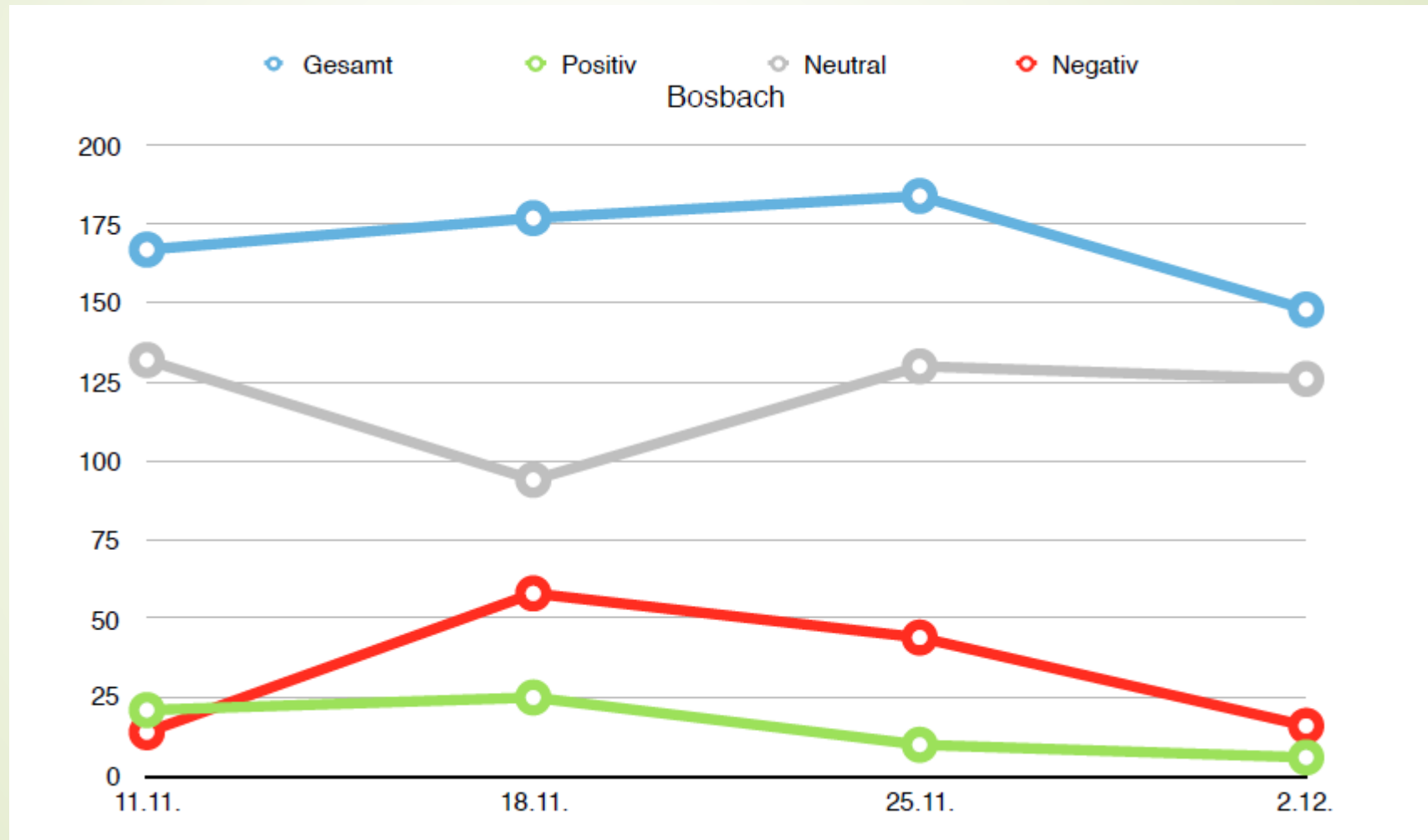
- 800 Tweets pro Politiker, randomisiert ausgewählt
- manuelle Bewertung nach positiv, negativ, neutral, nicht relevant
- Diskussion der Ergebnisse
- Problem Inter-Annotator Agreement

Auswertung: Vergleich der Politiker miteinander

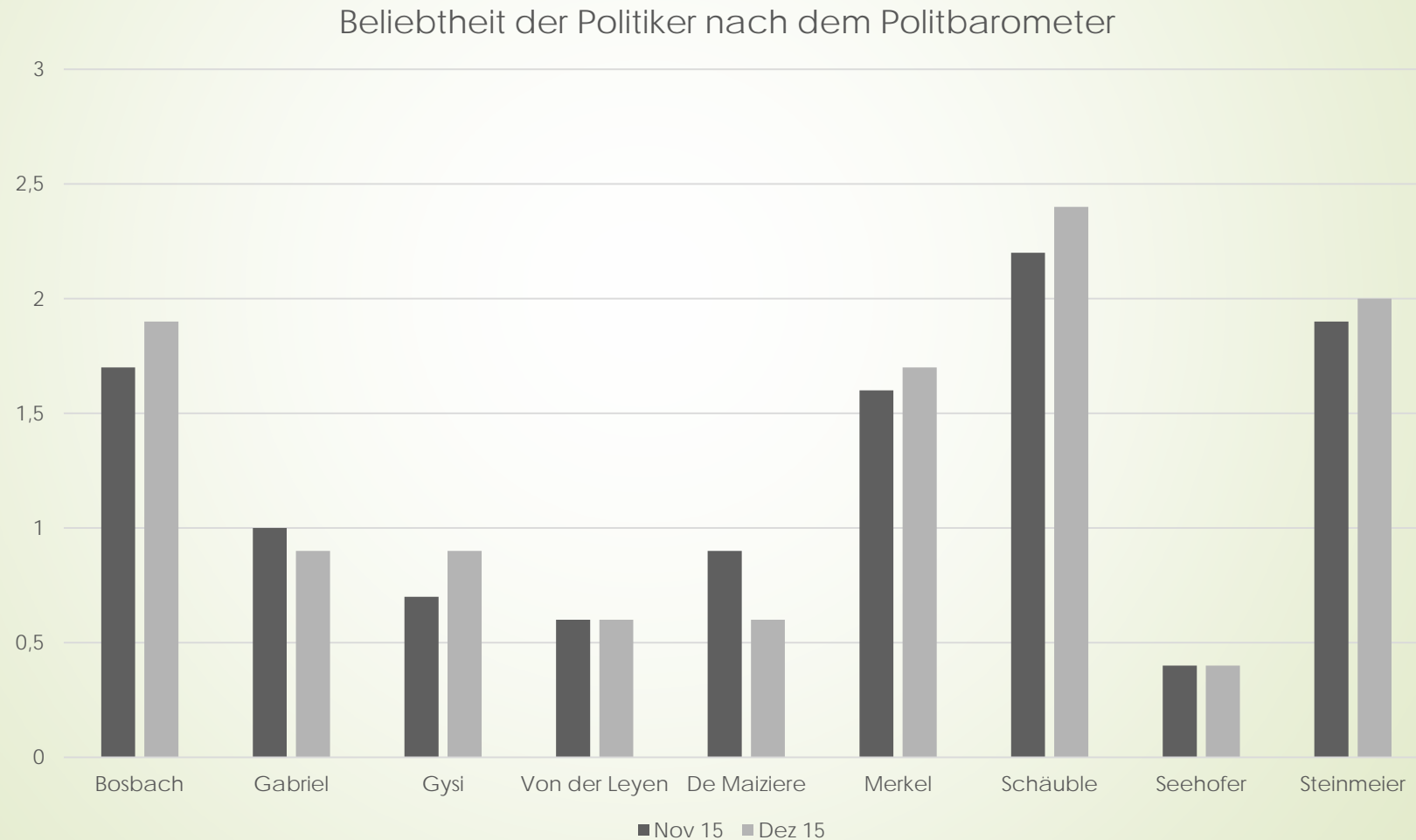
Tweets zu Politikern im Herbst 2015



Auswertung: Verlauf über vier Wochen



Vergleich mit dem Politbarometer: Veränderungen im November und Dezember 2015



z.B. Gysi

- Politbarometer: + 0,2 Punkte
- Twitter: starker Anstieg der positiven Tweets, geringer Anstieg der negativen

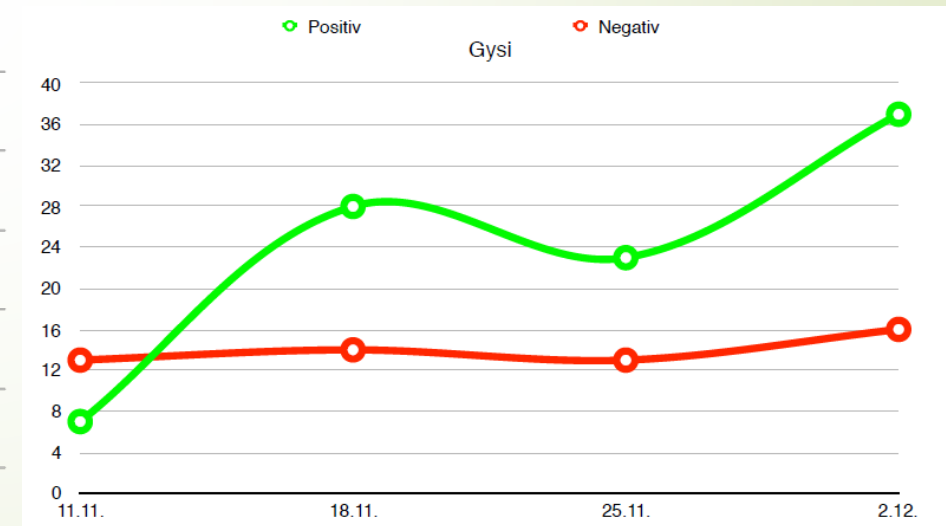
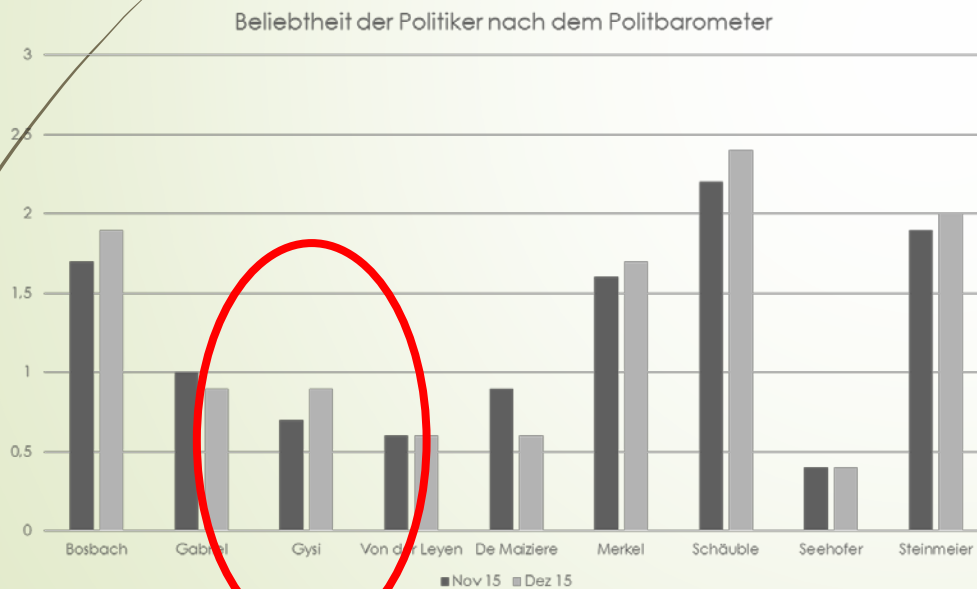
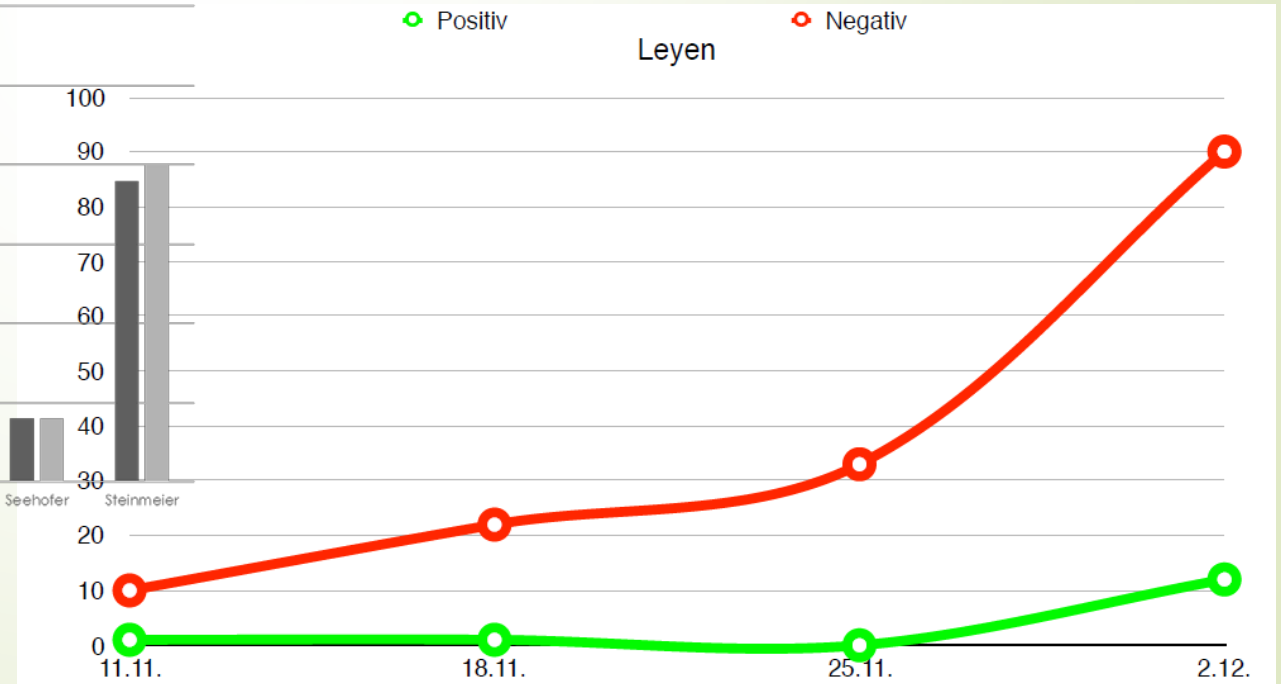
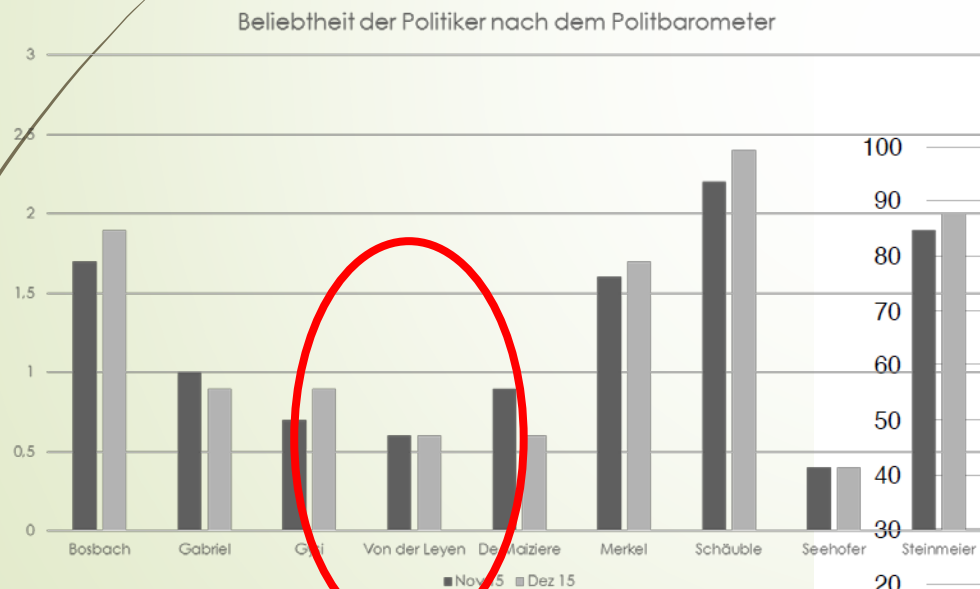


Abbildung 15: Positiv/Negativ Twitterverlauf Gregor Gysi

z.B. von der Leyen

Hintergrund: Auftritt im Morgenmagazin am 2.12.15, Einsatz von Tornados in Syrien angekündigt

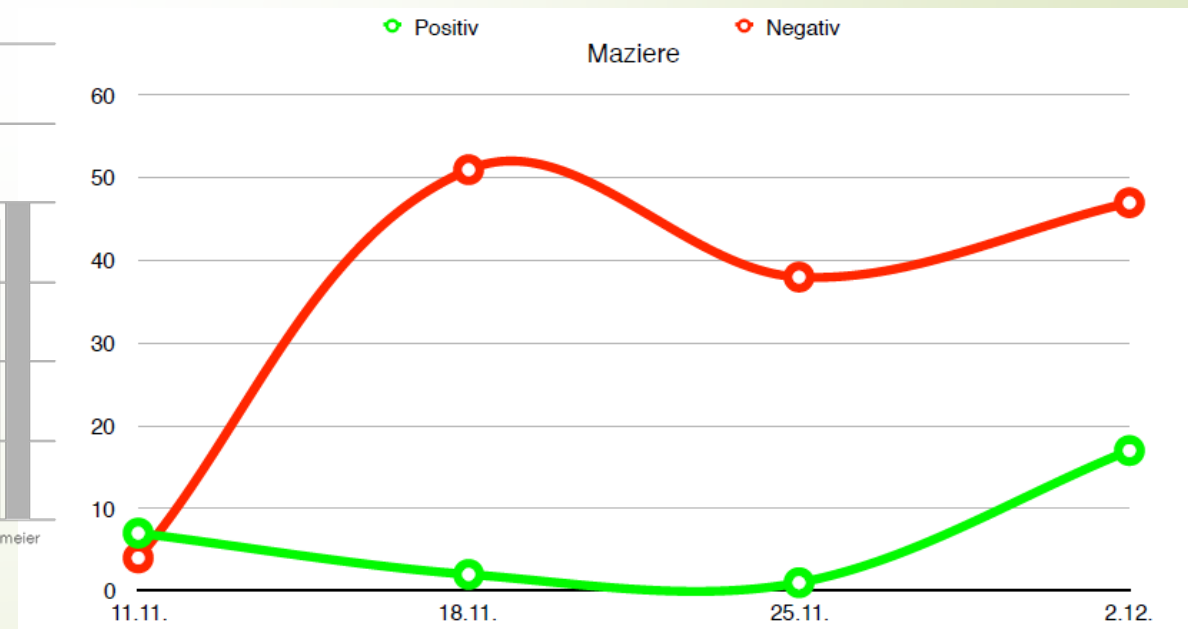
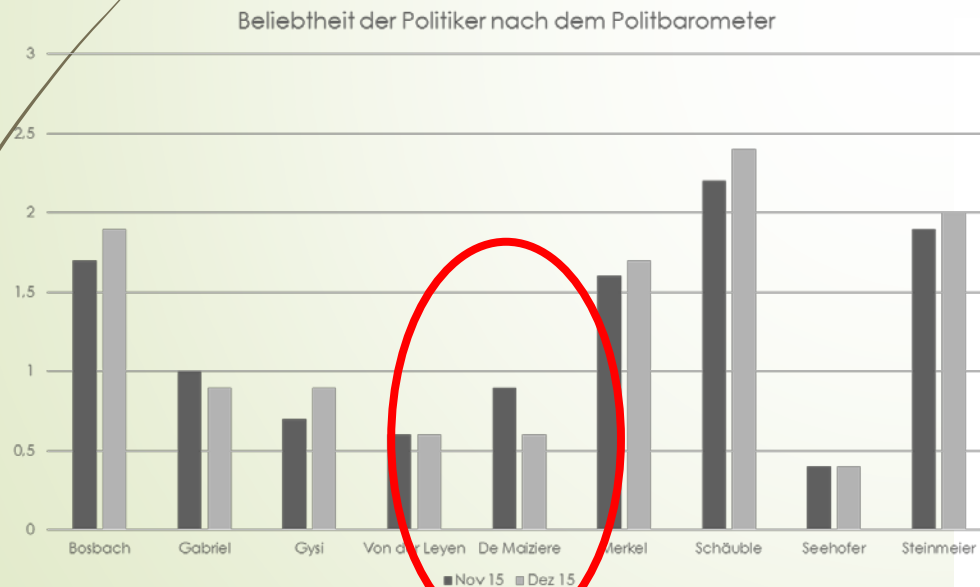
- Politbarometer: gleichbleibend
- Twitter: Anstieg der negativen Tweets zu einem Zeitpunkt, leichter Anstieg der positiven Tweets



z.B. De Maiziere

Hintergrund:
„Ein Teil dieser Antworten würde die Bevölkerung verunsichern“

- Politbarometer: - 0,3 Punkte
- Twitter: erheblicher Anstieg der negativen Tweets, Anstieg der positiven Tweets



Ergebnisse des Vergleichs

- ▶ Deutliche Übereinstimmung Twitter und Politbarometer erkennbar
- ▶ Trotz unterschiedlicher Vorgehensweisen bei der Erhebung der Daten!
- ▶ Twitter-Analyse ist teilweise unmittelbarer, reagiert schneller auf Wendungen aufgrund politischer Ereignisse

Grenzen des Ansatzes

- automatische Klassifizierung ist notwendig, aber komplex
- gefälschte, angeblich unabhängige Meinungsäußerungen
- Social Bots
- Wie aussagekräftig ist ein Vergleich mit einer Meinungsforschung, die in der Krise ist?

Vielen Dank für die Aufmerksamkeit!

23

melanie.siegel@h-da.de