Annika Hamachers & Andrej Galic

German Police University, Institute of Communication Science

Profiling Malignant Rhetoric.

Linking cognitive linguistics and machine learning algorithms to evaluate the emotionality of populist discourse.

Who we are

German Police University
(Münster, North Rhine-Westphalia)
special higher learning institution for police officers with university status (doctorate)

Joint Project: X-SONAR
"*Extremist Engagement in Social Media Networks: Identifying, Analyzing and Preventing Processes of Radicalization*"

SPONSORED BY THE
Federal Ministry of Education and Research

Emotions as a key to the people

Concentrate/bundle existing ressentiments within society
Intensifying them (vgl. Wodak 2016)

wide range of feelings, which include nostalgia, fear, helplessness, hatred, vindictiveness, ecstasy, melancholy, anger, fear, indignation, envy, spite and resentment (Minogue 1969, Taggart 2000, etc.)

fear of social decline / loss of social status (vgl. Manow 2019)

contributing to the forming of collective identities (We vs. The Others) (vgl. Demertzis 2006, Wodak 2016)

SPONSORED BY THE
Federal Ministry of Education and Research

Emotions as a key to the people

- intensifying emotions can lead to „vigilante terrorism“ (vgl. Quent 2016); Chemnitz demonstrations 2018 (chivvy on migrants)

  vigilante justice out of the majority against marginalized groups (Quent 2015, Quent 2016a, Quent 2016b)

- only a few are violent, but often initiated by many others who advocate/support that violence (Krüpper/Zick/Krause 2015: 42f.)
- "populist protests are a platform zu transform anger to hate and to intensify it and instrumentalize it for political demands (Quent 2017: 59)
- "***growing anger of the silent majority?***"

  when justifiable anger is combined with prejudices, anger becomes hate that again reinforces violent potential (vgl. Quent 2017: 58)

SPONSORED BY THE
Federal Ministry of Education and Research

Emotions as a key to the people

- intensifying emotions can lead to „vigilante terrorism“ (vgl. Quent 2016); Chemnitz demonstrations 2018 (chivvy on migrants)

  vigilante justice out of the majority against marginalized groups (Quent 2015, Quent 2016a, Quent 2016b)

Emotions as a key to the people

- intensifying emotions can lead to „vigilante terrorism“ (vgl. Quent 2016); Chemnitz demonstrations 2018 (chivvy on migrants)

  vigilante justice out of the majority against marginalized groups (Quent 2015, Quent 2016a, Quent 2016b)

- only a few are violent, but often initiated by many others who advocate/support that violence (Krüpper/Zick/Krause 2015: 42f.)
- "populist protests are a platform zu transform anger to hate and to intensify it and instrumentalize it for political demands (Quent 2017: 59)
- "***growing anger of the silent majority?***"

  when justifiable anger is combined with prejudices, anger becomes hate that again reinforces violent potential (vgl. Quent 2017: 58)

SPONSORED BY THE
Federal Ministry of Education and Research

Research Question

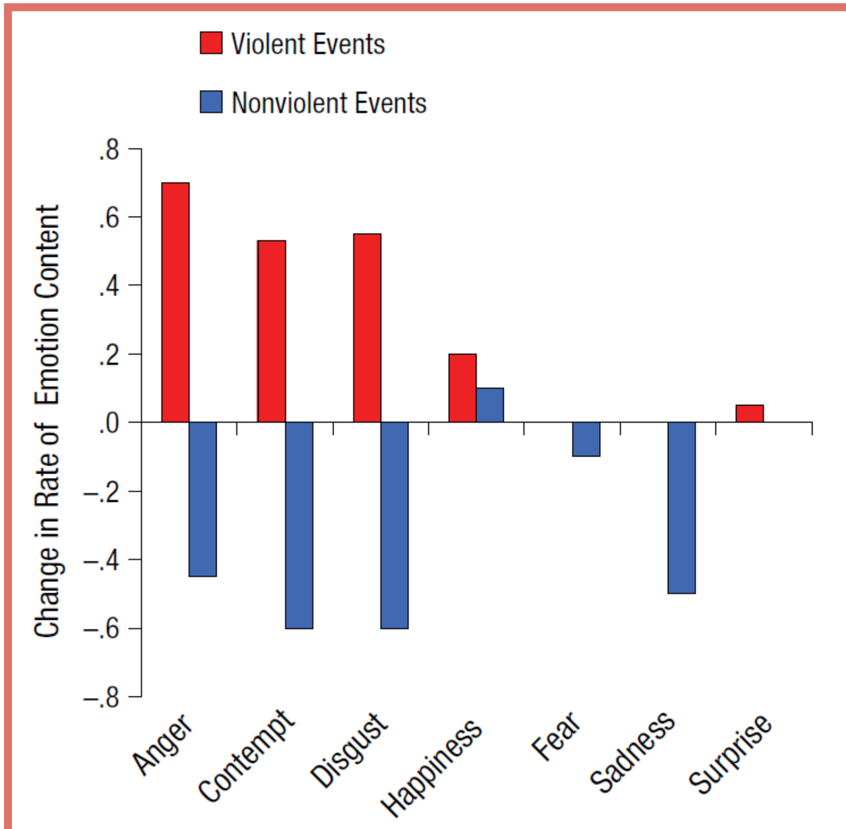
Does the emotionalization of populist language actually trigger violent action?

SPONSORED BY THE

Federal Ministry of Education and Research

The 'ACONDI' Hypothesis

Matsumoto and colleagues (2012; 2015) could demonstrate a drastic increase in expressions of ***A**nger*, ***CON**tempt*, and ***DI**sgust* in the political rhetoric precendent to violent events (as compared to non-violent campaign outcomes).

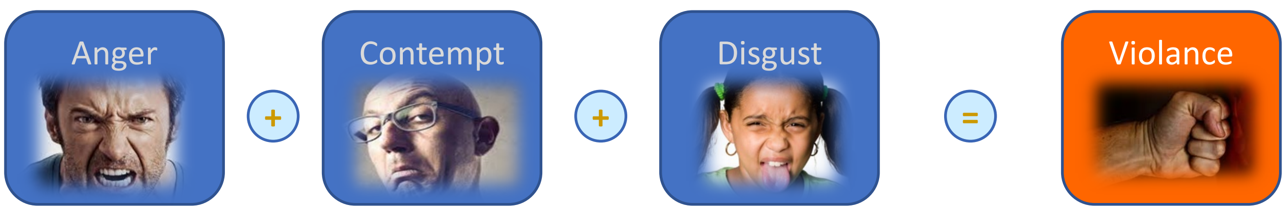


Source: Matsumoto et al. 2012

The ‘ACONDI’ Hypothesis

[T]he combination of the emotions of anger, contempt, and disgust (ANCODI) produce[s] a more volatile mix than any one of these emotions alone, and thus their presence in speeches and behavior predicts intergroup hostility and political violence […].

These emotions function through the **ability of anger to motivate action**, of **contempt to motivate devaluation** of others, and of **disgust to motivate the elimination** of others.



Hypotheses

H1: Right-wing populist discourses convey the emotions anger, disgust, and contempt to a greater amount than non-populist discourses.

H2a: An increase in expressions of anger, disgust, and contempt by right-wing populists correlates with/precedes an increase in *violent action* (protests).

H2b: An increase in expressions of anger, disgust, and contempt by right-wing populists correlates with/precedes an increase in *violent language* expressed on social media.

SPONSORED BY THE
Federal Ministry
of Education
and Research

Dictionary-Based Word Count Analyses

The rational:

The **words** we use in daily life **reflect** what we are paying attention to, what we are thinking about, what we are trying to avoid, how we are feeling, and **how we are organizing and analysing our worlds**.

(Pennebaker, 2010)

→ focused analyses of the words used in a given text allow us to draw conclusions about the personality, thoughts, feelings, and intentions of the author

→ WC techniques already successfully applied to analyses of radical contents (e.g. Chalothorn & Ellmann, 2013; Cohen, Kaati, & Shresta, 2016; Pennebaker & Chung, 2008;)

→ ***LIWC*** (Linguistic Inquiry and Word Count) as 'gold standard' in dictionary-based text analysis with a scientific foundation/validation

SPONSORED BY THE
Federal Ministry
of Education
and Research

The LIWC

contains >7000 terms, coded for 68 categories

emotions_LIWC.txt | LIWC_German.txt

```
%
01	Pronoun
02	I
03	We
04	Self
05	You
06	Other
07	Negate
08	Assent
09	Article
10	Preps
11	Numbers
12	Affect
13	Positiveemotion
14	Positivefeeling
15	Optimism
16	Negativeemotion
17	Anxiety
18	Anger
19	Sad
20	Cognitivemechanism
21	Cause
22	Insight
23	Discrepancy
24	Inhibition
25	Tentative
26	Certain
31	Social
32	Communication
33	Otherreference
34	Friends
35	Family
36	Humans
37	Time
```

emotions_LIWC.txt | LIWC_German.txt

```
gegolten	38
gegoren	38
gegossen	38
gegraben	38
gegriffen	38
gegrinst*	12	13
gegrübelt*	20	22
gegruebelt*	20	22
gehabt	38
gehaenselt	12	16	31	32	38
gehaessig*	12	16	18
gehalt*	39	47	49	56
gehalten	38
gehaltserhoehung*	47	49
gehaltserhöhung*	47	49
gehaltsscheck*	47	49	56
gehangen	38
gehänselt	12	16	31	32	38
gehässig*	12	16	18
gehasst*	12	16	18
gehaßt*	12	16	18
gehauen	38
gehe	39	46
geheißen	38
gehemmt*	12	16	20	24
gehen	46
gehindert*	12	16	20	24
gehirn*	60	61
gehoben	38
gehofft*	12	13	15	20	23	25
geholfen	38
geholt	38
```

SPONSORED BY THE
Federal Ministry of Education and Research

The LIWC

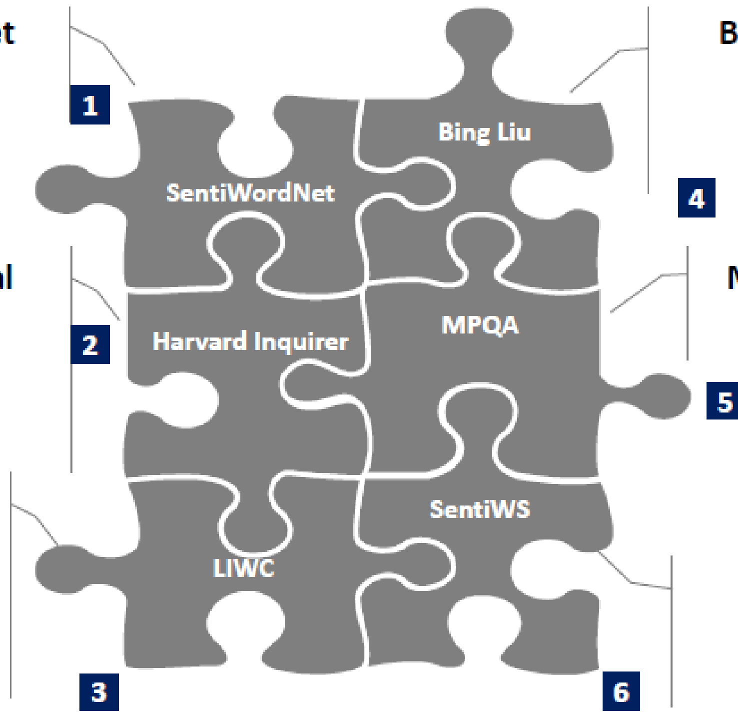
CF	CG	CH	CI	CJ	CK	CL	CM	CN	
Anxiety LIWC	Anger LIWC	Sad LIWC	Social LIWC	Friends LIWC	Family LIWC	Past LIWC	Present LIWC	Future LIWC	A
abschreck*	aerger*	versäum*	abgelehnt	arbeitskolleg	adoptivkind*	angeregt*	versäum*	zielstrebig*	a
aengst*	aersche	abgestumpft	abgerufen	begleiter*	angehoerige	steuert	abbreche	bald*	a
angst*	aggress*	allein	abgesagt	begleitperso	angehörige*	beeinflusst	abbrich*	demnächst	r
ängst*	androh*	alleine	ablehn*	bekannte*	bruder*	entschloss*	abfaehrst	morgen	s
aufgeregt*	anekel*	aufgab*	abrief*	bekanntscha	brüder*	abbrach*	abfaehrt	übermorgen	v
aufreg*	angedroht*	aufgebe*	abruf*	brieffreund*	brueder*	abflog	abfahre	uebermorge	z
bang*	angeekelt*	aufgegeben*	absag*	busenfreund	cousin*	abfuhr*	abfährst	werde	b
befuercht*	angekotzt*	aufgib*	adoptivkind*	exfreund*	ehefrau	abgab*	abfährt	werden	e
beklemm*	ankotz*	bedauer*	aerger*	feundin*	ehegatte*	abgebrochen	abfliege	werdet	a
beklommen*	ärger*	bedaure*	aeussere	freund*	ehemaenner	abgefahren	abfliegst	wird	a
besorg*	arsch*	bedrückend*	aeussern	freunde*	ehemann*	abgeflogen	abfliegt	wirst	a
beunruhig*	ärsche	bedrueckend	aeusserte*	freundin*	ehemänner*	abgegeben	abgebe	zukuenftig*	a
demuetig*	aufgelehnt*	beklommen*	aeusserung*	gaeste*	ehepartner*	abgelaufen	abgib*	zukunft*	a
demütig*	auflehn*	bekuemmert	andeute	gast*	einzelkind*	abgelehnt	ablaeuf*	zukünftig*	a
erniedrig*	aufstand	bekümmert*	andeuten	gäste*	eltern*	abgeliefert	abläuf*	NA	a
erpress*	ausfallend*	bemitleid*	andeutest	gefaehrte*	enkel*	abgenomme	ablaufe	NA	a
erschrak*	ausflipp*	benachteilig*	andeutete*	gefaehrtin*	ex*	abgepackt	ableb*	NA	a
erschrecke	ausgenutzt*	bereu*	andeutung*	gefährte*	familie*	abgereist	abliefere	NA	a

Dictionary-Based Word Count Analyses

SentiWordNet

Harvard General Inquirer

Linguistic Inquiry and Word Count (LIWC)



Bing Liu's Opinion Lexicon

MPQA Subjectivity Lexicon

SentiWS

http://sentiment.christopherpotts.net/lexicons.html

SPONSORED BY THE

Federal Ministry of Education and Research

Sampling Populism

- Thin ideology? (host ideology)
  People centrism, restoration of sovereignty, anti-elitism
  (following Canovan, Mudde, Taggart, Schmitt, Puhle, Werz, etc.)

- Main focus on
  authoritarian populism (vgl. Zürn 2018)
  with extremist characteristics

  illiberal (constraining minority rights)
  anti-pluralistic (charismatic leadership)
  anti-multilateral (national sovereignty)

  Cleavage: Communitarianism vs. Cosmopolitanism

SPONSORED BY THE
Federal Ministry
of Education
and Research

Sampling Populism

The “**intellectual”** network of the **New Right** in Germany

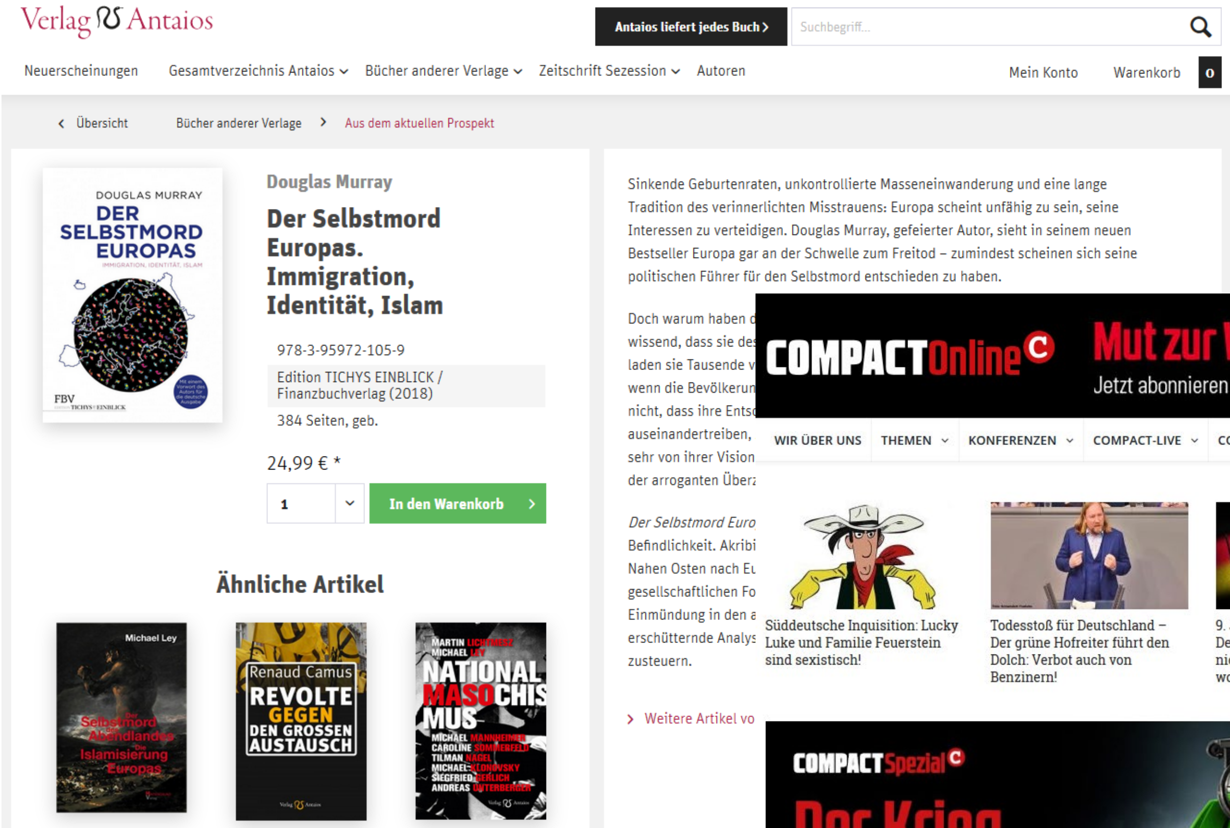
„New Right“ Education

- ‘***Institut für Staatspolitik’*** (think-tank)
- Publishing company ‘***Antaios Verlag***’
- Newspaper (‘***Junge Freiheit’***)
- Magazines (‘***Compact***’, ‘***Sezession’***)

close linkage to the party „***AfD***“
11,5% in the German Bundestag (91 of 701 seats)

SPONSORED BY THE
Federal Ministry
of Education
and Research

Sampling Populism



SPONSORED BY THE
Federal Ministry of Education and Research

Sampling Populism

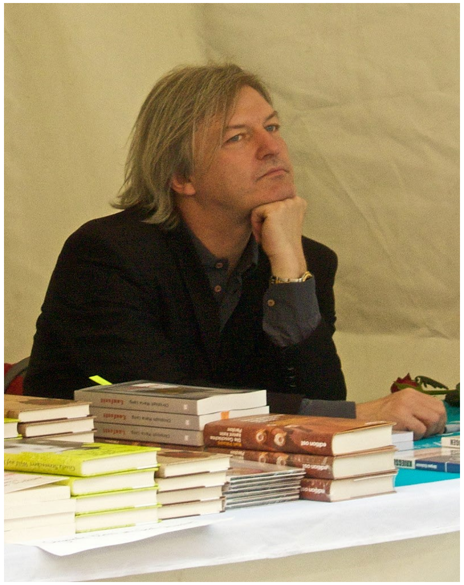
- self entitled "true press"
- self positioning against mainstream media

Götz Kubitschek

one of the most important protagonists of the Neue Rechte
publisher, journalist and right-wing political activist; founder of „Ifs“

SPONSORED BY THE
Federal Ministry of Education and Research

Sampling Populism



Jürgen Elsässer

Chief editor of „Compact Magazin“
changing from radical left (antiimperialistic) to positions of the New Right (PEGIDA, AfD)

SPONSORED BY THE
Federal Ministry of Education and Research

Sampling Populism

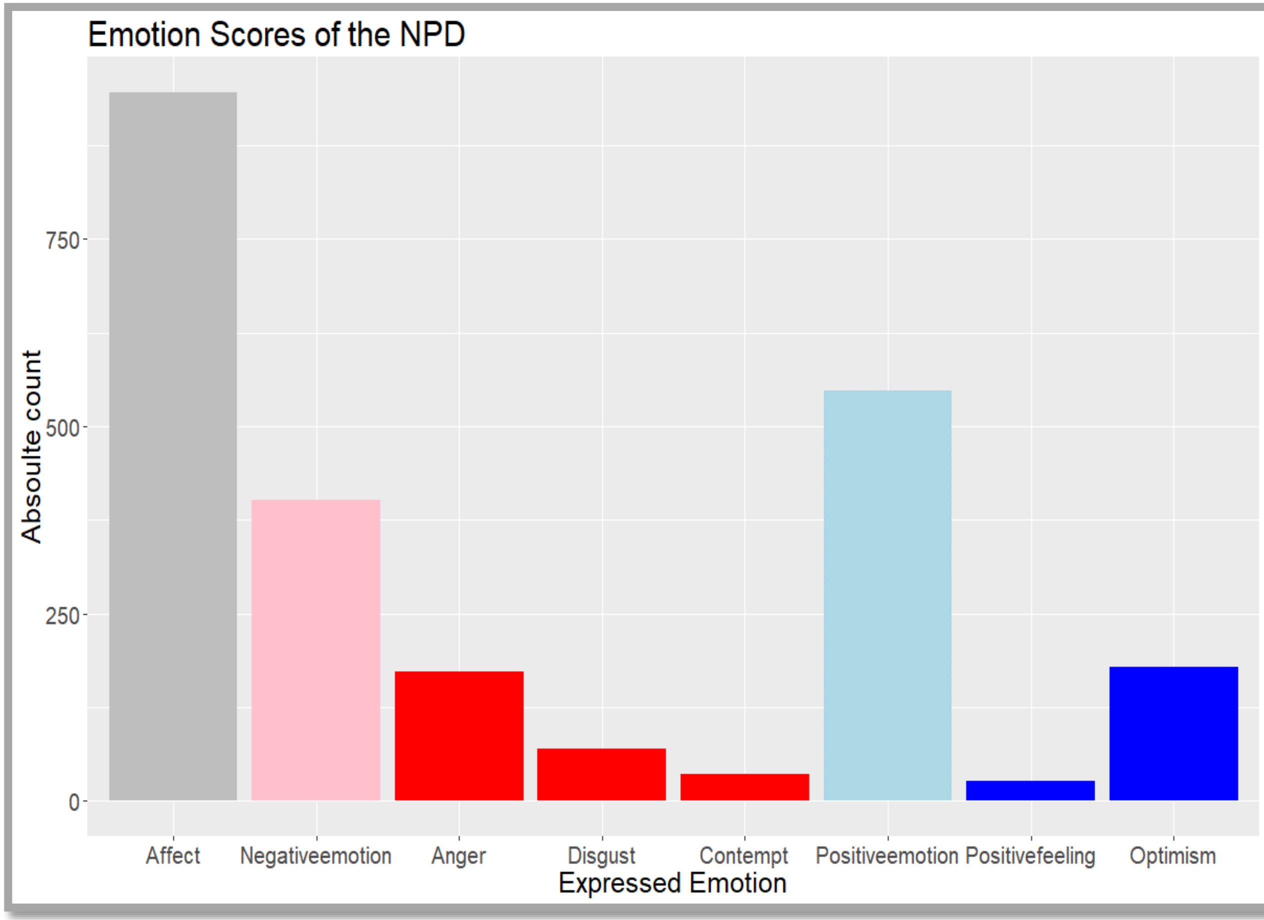
SONAR

plus interconnected ***Facebook fanpages*** (by ‘like’)

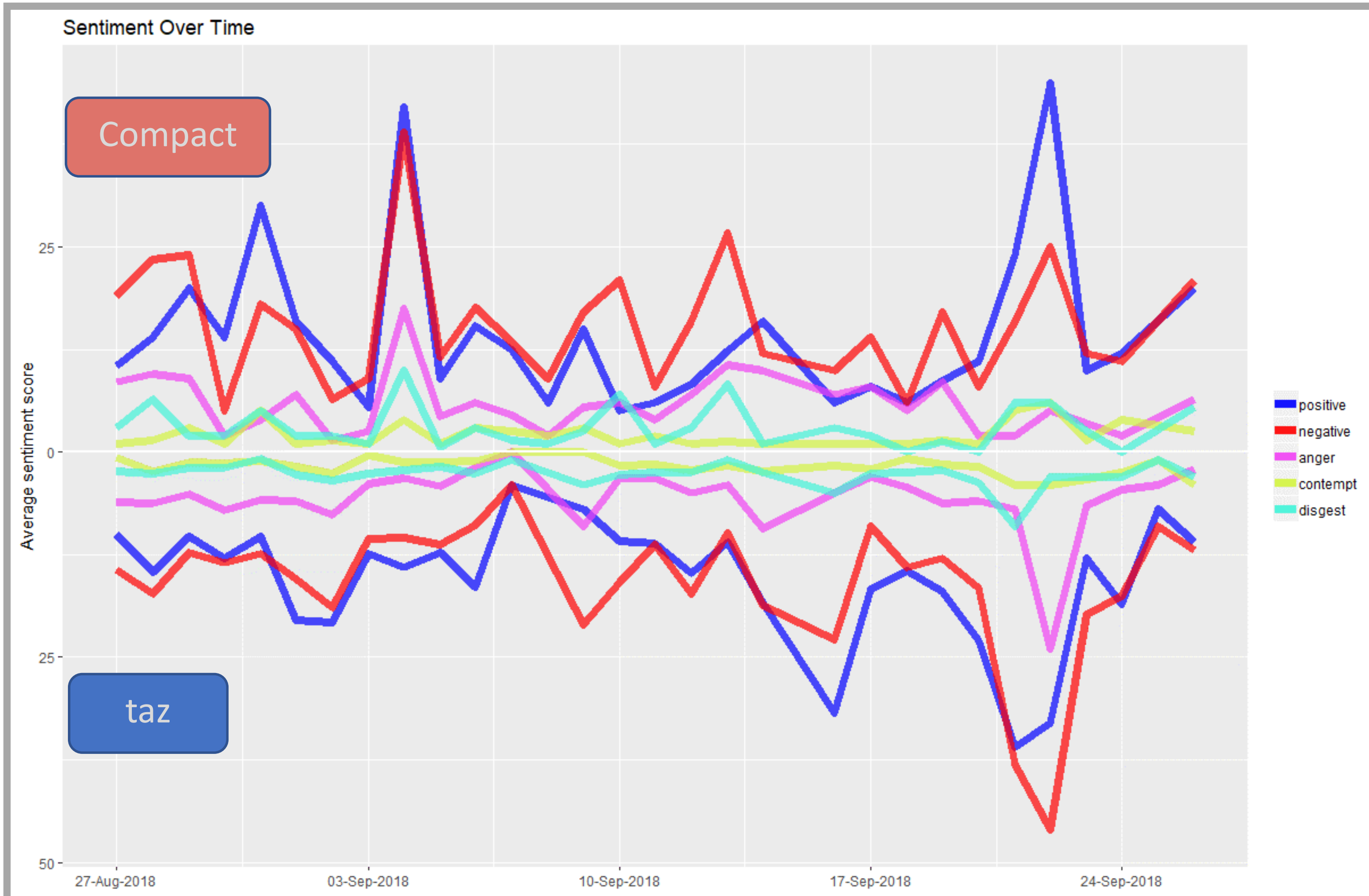
SPONSORED BY THE
Federal Ministry of Education and Research

Applying the LIWC & NRC

The different assumedly populist outcomes were explored with help of the R package ‘**quanteda**’ (Benoit et al., 2018).



Applying the LIWC & NRC



SPONSORED BY THE

Federal Ministry of Education and Research

Shortcomings of WC techniques

- existing emotion dictionaries tend to differ a lot
- existing emotion dictionaries do not withstand validity checks

SPONSORED BY THE

Federal Ministry
of Education
and Research

Shortcomings of WC techniques

- existing emotion dictionaries tend to differ a lot
- existing emotion dictionaries do not withstand validity checks

```
> get_nrc_sentiment("So ein verdammter Scheiss, ich hasse Juden",  language = 'german')
  anger anticipation disgust fear joy sadness surprise trust negative positive
1     0            0       0    0   0       0        0     0        0        0
> get_nrc_sentiment("Wir sollten gleich morgen den Bundestag anzünden!",  language = 'german')
  anger anticipation disgust fear joy sadness surprise trust negative positive
1     0            2       0    0   0       0        0     0        0        0
> get_nrc_sentiment("Merkel muss weg!",  language = 'german')
  anger anticipation disgust fear joy sadness surprise trust negative positive
1     0            0       0    0   0       0        0     0        0        0
> get_nrc_sentiment("Merkel macht mich wütend!",  language = 'german')
  anger anticipation disgust fear joy sadness surprise trust negative positive
1     0            0       0    0   0       0        0     0        0        0
> get_nrc_sentiment("Als guter Muslim solltest du in den Jihad ziehen",  language = 'german')
  anger anticipation disgust fear joy sadness surprise trust negative positive
1     0            0       0    0   0       0        0     0        0        1
```

Shortcomings of WC techniques

- existing emotion dictionaries tend to differ a lot
- existing emotion dictionaries do not withstand validity checks

→ reason: existing dictionaries are not ‘domain-specific‘ for populism

→ future aim: develope strategies to enhance the existing dictionaries

SPONSORED BY THE
Federal Ministry of Education and Research

How to create do*main-specific* dictionaries

word frequency analyses
(terms, n-Grams, hashtags, tf-idf)

expert coding

the emotion dictionary

unsupervised topic modeling

supervised topic modeling

target material
(posts, tweets, online articles)

additional material
(manifestos, books, cover texts)

Unsupervised Topic Modeling (Text *Clustering*)

Unsupervised machine learning:

- data are not labelled in advance

→the algorithms has to learn associations totally ‘on its own’

- output is a *clustering* of text

+ can be done ‘quick’n’dirty’, good for broad topic exploration

⊖ no possibility to include theoretical or prior empirical knowledge of the researcher

SPONSORED BY THE
Federal Ministry
of Education
and Research

Unsupervised Topic Modeling (Text *Clustering*)

Latent Dirichlet Allocation (LDA)(Blei et al. 2003)

- topic modeling generally applies a statistical model of the distribution of words in a text
- LDA assumes specific words to appear more frequently to appear in certain document as compared to other documents, if it is about a particular topic, compared to other documents
- analysis performed in **KNIME** (Berthold et al., 2008)

Supervised Topic Modeling (Text *Classification*)

Unsupervised machine learning:

- data are split into a 'labelled' training in advance

→the algorithms has to figure out the rules underlying that labeling

- output is a *classification probability* of text

+ labelling (usually) is a very tedious task (some hundred labelled documents necessary)

− labelling can reflect the knowledge of the researcher

Supervised Topic Modeling (Text *Classification*)

Unsupervised machine learning:

- data are split into a ‘labelled’ training in advance

→the algorithms has to figure out the rules underlying that labeling

- output is a *classification probability* of text

```
> head(class_emo_n,25)
      ANGER              DISGUST            FEAR               JOY                SADNESS            CONTEMPT           BEST_FIT
 [1,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "2.78695866252273" NA
 [2,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "2.78695866252273" NA
 [3,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "2.78695866252273" NA
 [4,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "2.78695866252273" NA
 [5,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "2.78695866252273" NA
 [6,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "7.34083555412327" "surprise"
 [7,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "7.34083555412328" "2.78695866252273" "sadness"
 [8,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "2.78695866252273" NA
 [9,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "2.78695866252273" NA
[10,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "2.78695866252273" NA
[11,] "1.46871776464786" "3.09234031207392" "2.06783599555953" "1.02547755260094" "1.7277074477352"  "2.78695866252273" NA
```

Supervised Topic Modeling (Text *Classification*)

Unsupervised machine learning:

- data are split into a 'labelled' training in advance

→the algorithms has to figure out the rules underlying that labeling

- output is a *classification probability* of text

+ labelling (usually) is a very tedious task
(some hundred labelled documents necessary)

– labelling can reflect the knowledge of the researcher

SPONSORED BY THE
Federal Ministry
of Education
and Research

SONAR

Using the dictionaries to create labelled data

emotions_LIWC.txt LIWC_German.txt

```
gegolten	38
gegoren 38
gegossen	38
gegraben	38
gegriffen	38
gegrinst*	12	13
gegrübelt*	20	22
gegruebelt*	20	22
gehabt	38
gehaenselt	12	16	31	32	38
gehaessig*	12	16	18
gehalt*	39	47	49	56
gehalten	38
gehaltserhoehung*	47	49
gehaltserhöhung*	47	49
gehaltsscheck*	47	49	56
gehangen	38
gehänselt	12	16	31	32	38
gehässig*	12	16	18
gehasst*	12	16	18
gehaßt*	12	16	18
gehauen	38
gehe	39	46
geheißen	38
gehemmt*	12	16	20	24
gehen	46
gehindert*	12	16	20	24
gehirn*	60	61
gehoben	38
gehofft*	12	13	15	20	23	25
geholfen	38
geholt	38
```

	row.names	text	emotion	larity
1	27	schulleiterin limit „ deutsche kinder geflüchtetenklassen integrieren " vorsitzende	anger	utral
2	258	aggressive bürger null toleranz statt schlagstöcke dortmund aggressiven bürgern über	anger	gative
3	367	richtig apothekerin erteilt kopftuch klare absage genau vorfälle irritation wut sor	anger	gative
4	515	besteht hoffnung spanien schiebt ceutagrenzstürmer ab mittwoch heer asylsuchender ü	anger	gative
5	578	aggressive migranten landfriedensbruch straftatbestand mehr wilden westen massensch	anger	gative
6	719	mordfall susanna immer schlimmere details kommen ans tageslicht tragische fall 14jä	anger	gative
7	744	herzlich willkommen beim livestream afd-großdemonstration brandenburger tor	anger	sitive

SPONSORED BY THE
Federal Ministry
of Education
and Research

Supervised Topic Modeling (Text *Classification*)

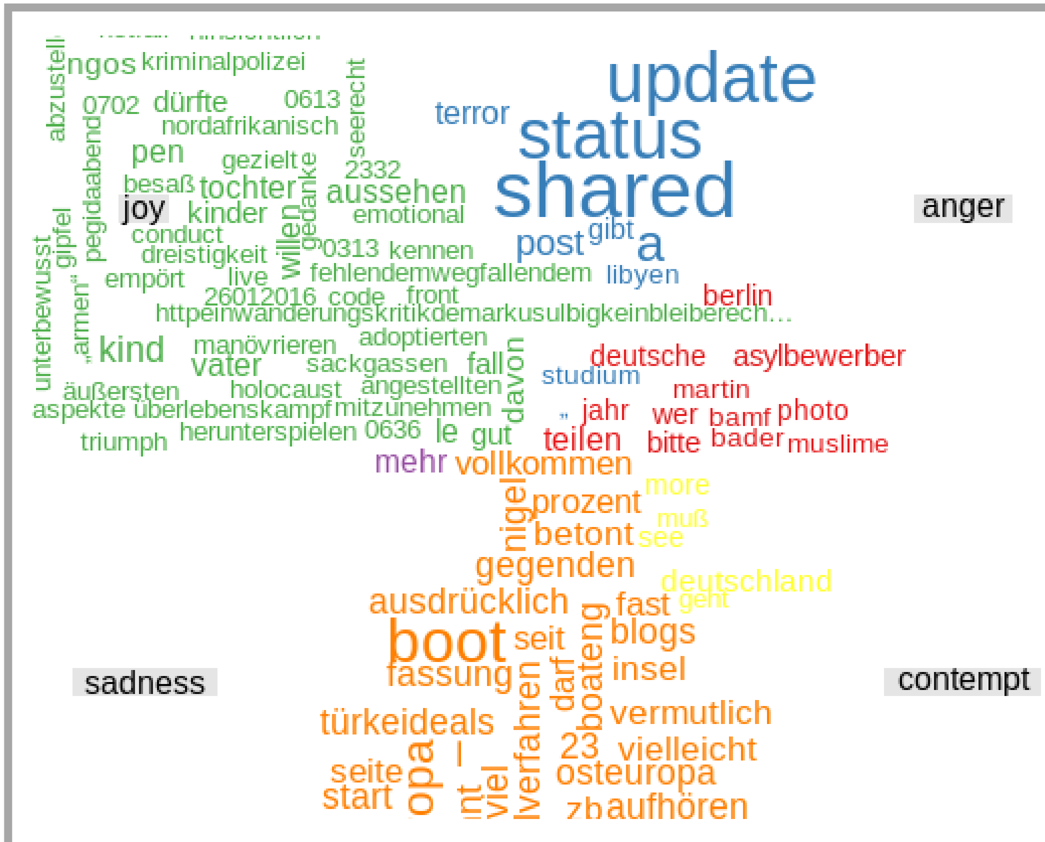
Naïve Bayes Classification

- NB treats the words of a text as feature of underlying concepts, organizes them as vectors, and returns the words/features which it found the most crucial for making the classification decisons
- analysis performed in R with the ‘sentiment’ package

Supervised Topic Modeling (Text *Classification*)

Naïve Bayes Classification

- NB treats the words of a text as feature of underlying concepts, organizes them as vectors, and returns the words/features which it found the most crucial for making the classification decisons
- analysis performed in R with the ‘sentiment’ package



Discussion

Future steps:

- try other aggregates of the data corpus
- try other supervised algorithms that make less simplistic assumptions (Hvitfeldt, 2019), e.g.
  - support vector machines
  - neural networks
  - random trees
- replace tools (especially the ‘sentiment’ package)

General concerns/theoretical shortcomings:

- Problems of data validity
  - limits of textual analysis (excludes photos, memes, info graphs etc.)
  - irony, sarcasm, stance, negation
  - embeddedness of online material→ ‘recursiveness’ of social media discourses
- definitory exclusion: populism ≠ extremism

SPONSORED BY THE
Federal Ministry of Education and Research

Keeping up with traditions…

Thank you for your attention!

~

Questions, comments, remarks?

SPONSORED BY THE
Federal Ministry of Education and Research