Automatische Sentimentanalyse zwischen Hotel und Parlament

Manfred Stede
Universität Potsdam

Digital * Humanities im Gespräch / FU Berlin / 6. Dezember 2018

Overview

- Sentiment analysis: Introduction, Terminology
- One system: SO-CAL
- Application: Political text
- Application: Narrative
- Viewpoint: Sentiment / DH

One Bolzano hotel @ tripadvisor.com

Near-Synonyms ?

- Opinion mining
- Sentiment analysis
- Subjectivity analysis

Subjectivity

- *I don't like this wine.*
- *There is a cat on the mat.*
- *I'm dizzy.*
- *Peter adores Barack Obama.*
- *I don't think that Trump can win the election.*
- *Last night I met this really nice musician.*
- *Hooray!*
- *That's probably a dromedar, not a camel.*

Subjectivity

- The linguistic expression of somebody's **opinions**, **sentiments**, emotions, evaluations, beliefs, speculations (Wilson/Wiebe: MPQA guidelines)

- **Private state:** state of a speaker/writer that is not open to objective observation or verification

  Quirk, Greenbaum, Leech, Svartvik (1985). *A Comprehensive Grammar of the English Language*.

- Automatic subjectivity analysis classifies content as **objective** or **subjective**

Subjectivity

- **Sentiment:** an attitude or feeling (not necessarily directed toward something)
- **Opinion:** an evaluation *of* something (necessarily directed)
- => Sentiment analysis and Opinion mining overlap, but there can be sentiment analysis that does not mine opinions (e.g., capture the general mood in the newspapers)
- **In practice**, most automated systems reduce evaluation to **polarity**

Manfred Stede (Uni Potsdam)

Text-level sentiment analysis

- Firstly the good points .... We had a very large room with fantastic bathroom and walk in closet. There was a good breakfast selection and possible to eat outside. There was a pretty garden with an outside bar and it was nice to sit outside after dinner. Location was excellent and a couple of minutes walk to the main square.
- => positive

Manfred Stede (Uni Potsdam)

Text-level sentiment analysis

- Firstly the good points .... We had a very large room with fantastic bathroom and walk in closet. There was a good breakfast selection and possible to eat outside. There was a pretty garden with an outside bar and it was nice to sit outside after dinner. Location was excellent and a couple of minutes walk to the main square.

- => positive

The most basic approach

Lexicon of words with **prior polarity**

- excellent, fantastic, good, nice, ...
- boring, terrible, uncool, ugly, ...

Simple Method:

- Preprocess the text (stemming, lemmatization)
- Count positive and negative words in document

Units for sentiment analysis

- **Text**
  - assume it has one topic and one overall orientation
- **Paragraph**
  - likewise; can compute text orientation afterward
- **Sentence**
  - likewise; can compute para orientation afterward
- **Phrase**
  - can capture things like
    While the breakfast was good, I couldn't stand dinner

Extensions (1): Polar facts

- Firstly the good points .... We had a very large room with fantastic bathroom and walk in closet. There was a good breakfast selection and possible to eat outside. There was a pretty garden with an outside bar and it was nice to sit outside after dinner. Location was excellent and a couple of minutes walk to the main square.

Extensions (2): Aspects

- Firstly the good points .... We had a very large room with fantastic bathroom and walk in closet. There was a good breakfast selection and possible to eat outside. There was a pretty garden with an outside bar and it was nice to sit outside after dinner. Location was excellent and a couple of minutes walk to the main square.

Extensions (2): Aspects

- Firstly the good points .... We had a very large room with fantastic bathroom and walk in closet. There was a good breakfast selection and possible to eat outside. There was a pretty garden with an outside bar and it was nice to sit outside after dinner. Location was excellent and a couple of minutes walk to the main square.

Extensions (3): Fine-grained analysis

While the breakfast was good, I couldn't stand dinner

Extensions (3): Fine-grained analysis

[While the breakfast was very good,]$_{opin1}$
[I couldn't stand dinner]$_{opin2}$

- Words with prior polarity
- Intensifiers/diminishers
- Source
- Target

Extensions (4): Entity-level sentiment

Roger Federer won the match against Nadal, who had been fervently supported by the audience.

Extensions (4): Entity-level sentiment

Roger Federer won the match against Nadal, who had been fervently supported by the audience.

What is good or bad for whom?

*Roger Federer **won** the match against Nadal, who had been fervently supported by the audience.*

Extensions (4): Entity-level sentiment

Roger Federer won the match against Nadal, who had been fervently supported by the audience.

What is good or bad for whom?

Roger Federer ***won*** *the match against Nadal, who had been fervently supported by the audience.*

Extensions (4): Entity-level sentiment

Roger Federer won the match against Nadal, who had been fervently supported by the audience.

What is good or bad for whom?

Roger Federer ***won*** *the match against Nadal, who had been fervently* ***supported by*** *the audience.*

I can‘t say that I enjoyed my stay at the Belvedere Hotel. Other reviewers said it‘s a great place, but my impression was otherwise. Neither was the food particularly good, nor did we consider the location very convenient. Just a standard place to live for a day, that‘s it.

I can‘t say that I enjoyed my stay at the Belvedere Hotel. Other reviewers said it‘s a great place, but my impression was otherwise. Neither was the food particularly good, nor did we consider the location very convenient. Just a standard place to live for a day, that‘s it.

Extensions (5): Contextual polarity

I can‘t say that I enjoyed my stay at the Belvedere Hotel. Other reviewers said it‘s a great place, but my impression was otherwise. Neither was the food particularly good, nor did we consider the location very convenient. Just a standard place to live for a day, that‘s it.

Overview

- Sentiment analysis: Introduction, Terminology
- **One system: SO-CAL**
- Application: Political text
- Application: Narrative
- Viewpoint: Sentiment / DH

SO-CAL

- **S**emantic **O**rientation **CAL**culator
- Text-level polarity analysis (graded)

- **Selling points:**
  - Use crowdsourcing in building a lexicon
  - Rule-based approach to contextual polarity (with some new ideas)
  - Achieves good level of domain-neutrality

M. Taboada / J. Brooke / M. Tofiloski / K. Voll / M. Stede: Lexical Methods for Sentiment Analysis. *Computational Linguistics* 37(2), 2011

Size

2252 adjectives
1142 nouns
903 verbs
745 adverbs

Words collected from 500 movie and product reviews (8 categories, balanced for pos and neg)

Manually ranked on -5 .. 5 scale: prior polarity and strength

Word	SO Value
monstrosity	−5
hate (noun and verb)	−4
disgust	−3
sham	−3
fabricate	−2
delay (noun and verb)	−1
determination	1
inspire	2
inspiration	2
endear	3
relish (verb)	4
masterpiece	5

Lexical ambiguity

- **Sense** ambiguity: sometimes resolved via PoS
  - *plot*: neutral noun, negative verb
  - *novel*: neutral noun, positive adjective

- **Connotation** ambiguity: „resolved“ by averaging
  - *The teacher inspired her students to pursue their dreams.*
  - *This movie was inspired by true events.*

Derivations

- Some nouns **derived automatically** from verb dictionary, but strength can change
  - *exaggerate*: -1
  - *exaggeration*: -2
  - also: *complicate* / *complication*, etc

  - (hypothesis: general trend?)

Derivations (2)

- Adverb dictionary built from adjectives, by *–ly* matching
- sometimes value needs to be corrected: *essential* / *essentially*

Word	SO Value
excruciatingly	-5
inexcusably	-3
foolishly	-2
satisfactorily	1
purposefully	2
hilariously	4

Context: Intensification

- **amplifiers** (*very*) **downtoners** (*slightly*)
- Polanyi/Zaenen 06, Kennedy/ Inkpen 06:
  add and **subtract** values
- **BUT:** degree of intensification should depend more on the word intensified
  => **multiply**
- 177 intensifiers in the lexicon

Intensifier	Modifier %
slightly	-50%
somewhat	-30%
pretty	-10%
really	+15%
very	+25%
extraordinarily	+50%
(the) most	+100%

really very good$_3$
3 x (100 + 25%) x (100 + 15%)
= 4.3

Context: Negation

- *The acting was not very good.*
- Some negators appear at long distance
  - *Nobody gives a good performance in this movie*.
- Strategy: Look backwards until a clause boundary (punctuation or connective) is reached
  - *I don't think this will be a problem*.

Manfred Stede (Uni Potsdam)

Context (2): Negation - value change

- One approach: **polarity flipping** (e.g., Choi/Cardie 08)
- Problems
  - *excellent*: +5
  - *not excellent*: -5 ??
  - *atrocious*: -5
- => Use **polarity shift** (+/-4) rather than flip
  - *The food is not terrific* (5 – 4 = 1) *but not terrible* (-5 + 4 = -1) *either*.
  - *It's not a spectacular* (5 - 4 = 1) *film*.

Manfred Stede (Uni Potsdam)

Context (3): Irrealis blocking

- *For kids, this movie **could** be one of the **best** of the holiday season.*
- *I thought this movie **would** be as **good** as the Grinch, but unfortunately it wasn't.*
- **Implementation:** ignore polar words in the scope of an irrealis marker (scope: heuristic)
  - modals
  - conditionality
  - NPIs (any, anything, ..)
  - questions
  - material in quotes
  - certain verbs (doubt, expect, ...)

- *This **should** have been a **great** movie.* (3 -> 0)

Example output

- *The food was wonderful.* => 5.0
- *The food was particularly good.* => 3.9
- *The food was good.* => 3.0
- *The food was not bad.* => 2.0
- *The food was OK.* => 2.0
- *The food was not particularly good.* => -0.1
- *I didn't really enjoy the food.* => -0.4
- *I didn't enjoy the food.* => -1.0
- *The food was bad.* => -3.0

Evaluation: Lexicon complexity

- Use not/recommended value of the review: >0 / <0
- 3 variants of the approach
  - Simple: only 2/-2 values and 1/-1 intensification (Polanyi/Zaenen 06)
  - Only-Adj: use only adjectives
  - One-Word: don't use multi-word expressions
  - Full: complete system as described

Dictionary	Percent correct by corpus				
	Epinions 1	Epinions 2	Movie	Camera	Overall
Simple	76.75	76.50	69.79*	78.71	75.11*
Only-Adj	72.25*	74.50	76.63	71.98*	73.93*
One-Word	80.75	80.00	75.68	79.54	78.23
Full	80.25	80.00	76.37	80.16	78.74

*Statistically significant using the chi-square test, $p < 0.05$.

Evaluation: Domain (in)dependence

Subcorpus	Epinions 1			Epinions 2		
	Pos-F	Neg-F	Accuracy	Pos-F	Neg-F	Accuracy
Books	0.69	0.74	0.72	0.69	0.77	0.74
Cars	0.90	0.89	0.90	0.80	0.75	0.78
Computers	0.94	0.94	0.94	0.90	0.89	0.90
Cookware	0.74	0.58	0.68	0.79	0.76	0.78
Hotels	0.76	0.67	0.72	0.80	0.70	0.76
Movies	0.84	0.84	0.84	0.76	0.79	0.78
Music	0.82	0.82	0.82	0.83	0.81	0.82
Phones	0.81	0.78	0.80	0.85	0.83	0.84
Total	0.81	0.79	0.80	0.81	0.79	0.80

Comparative Evaluation on Tweets

Polarity-Changing Factors	System Scores HL Macro $F_1{+/-}$	HL Micro F_1	TBD Macro $F_1{+/-}$	TBD Micro F_1	MST Macro $F_1{+/-}$	MST Micro F_1	JRK Macro $F_1{+/-}$	JRK Micro F_1	KLCH Macro $F_1{+/-}$	KLCH Micro F_1
PotTS										
All	0.615	0.685	0.593	0.671	0.606	0.675	0.339	0.467	0.468	0.651
–Negation	0.622	**0.691**	0.596	0.672	**0.641**	0.7	0.357	0.473	0.298	0.463
–Intensification	NA	NA	0.595	0.672	NA	NA	0.339	0.467	NA	NA
–Other Modifiers	NA	NA	0.613	0.684	NA	NA	NA	NA	NA	NA

U. Sidaranka: Sentiment analysis on German Twitter. Forthcoming dissertation, Univ. Potsdam

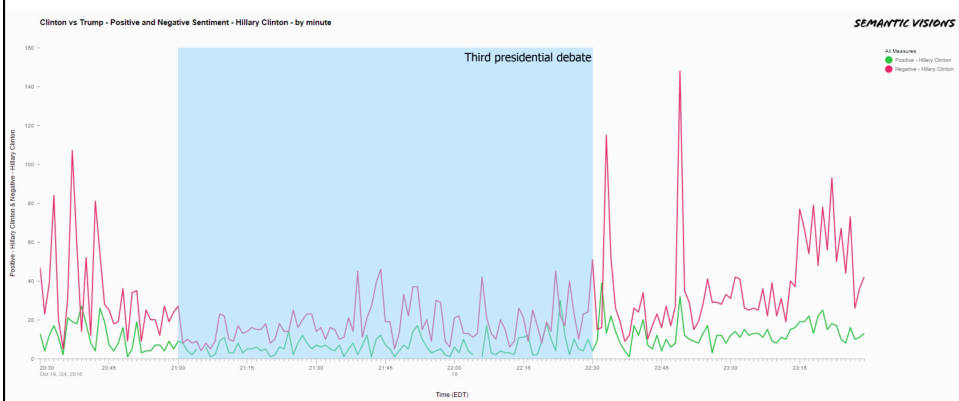
Overview

- Sentiment analysis: Introduction, Terminology
- One system: SO-CAL
- **Application: Political text**
- Application: Narrative
- Viewpoint: Sentiment / DH

(1) „Tweet sentiment viz“

https://www.csc2.ncsu.edu/facu/lty/healey/tweet_viztweet_app/

(2) „Semantic visions“



https://semantic-visions.com

(3) 2008 Primaries: Twitter

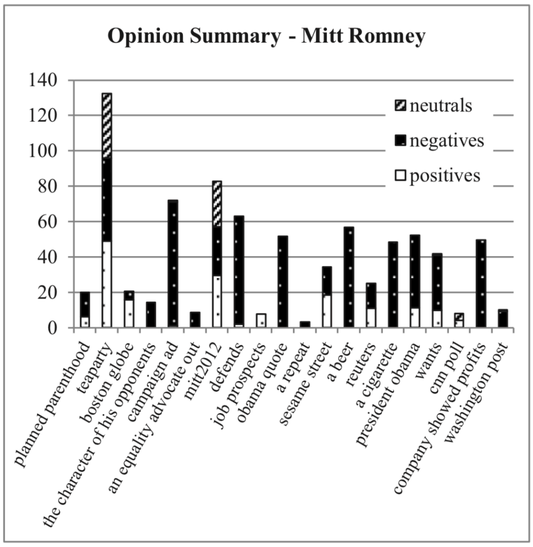
7000 Tweets / candidate

Aspect extraction

- correlation between noun phrases and cand. names via PMI

Sentiment

- use prior polarities from existing lexicon (+ domain-specific adjectives); count
- context: directly preceding negations
- no evaluation



M. Ringsquandl, D. Petkovic: Analyzing political sentiment on Twitter. Proc. of the AAAI Spring Symposium on Analyzing Microtext, 2013

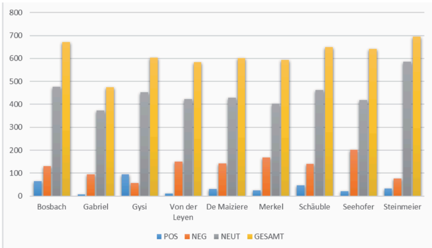
(4) ZDF Politbarometer

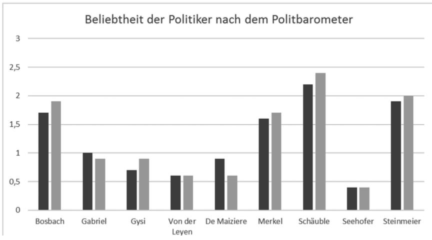
- Student project
- Try **automatic** analysis (manually-extended lexicon + rules) with context analysis => too difficult!
- **Manual** analysis of 8000 Tweets => not easy

- *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.*
- *#Oettinger wird niemals eine Anästhesistin brauchen. Der Mann ist echt absolut schmerzbefreit.*

M. Siegel et al.: Automatische Erkennung von politischen Trends mit Twitter – brauchen wir Meinungsumfragen noch? *Information, Wissenschaft & Praxis* 68(1), 2017

(4) ZDF Politbarometer



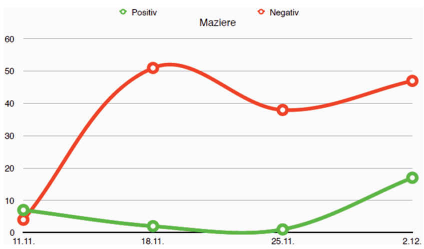


M. Siegel et al.: Automatische Erkennung von politischen Trends mit Twitter – brauchen wir Meinungsumfragen noch? *Information, Wissenschaft & Praxis* 68(1), 2017

(4) ZDF Politbarometer

Interior minister De Maizière on cancelling the football match D - NL (Nov 2015):

„Teile dieser Antworten würden die Bevölkerung verunsichern"



M. Siegel et al.: Automatische Erkennung von politischen Trends mit Twitter – brauchen wir Meinungsumfragen noch? *Information, Wissenschaft & Praxis* 68(1), 2017

(5) Crisis perception in UNRWA reports

- **United Nations Relief for Palestine Refugees (UNRWA)**
  - financed largely by voluntary contributions
    => resource mobilization is important, especially in times of unexpected demand for action
  - Data set: annual reports 1951-2016

- **Hypothesis:** *Under conditions of policy crisis, international bureaucracies are expected to signal increased budgetary stress to principals, donors and/or the public through documents or speech produced by the bureaucracy.*

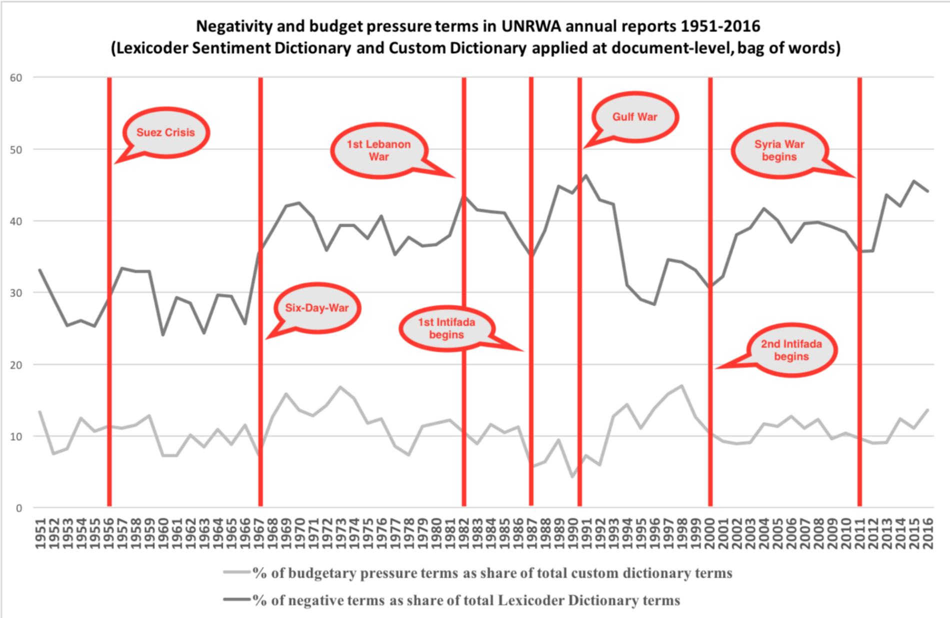
R. Patz et al.: International Public Administrations and the Perception of Crisis: The Case of UNRWA and Palestine Refugees. Panel "*Regional and Global Crisis Management*" at ECPR General Conference, Oslo 2017

(5) Crisis perception in UNRWA reports

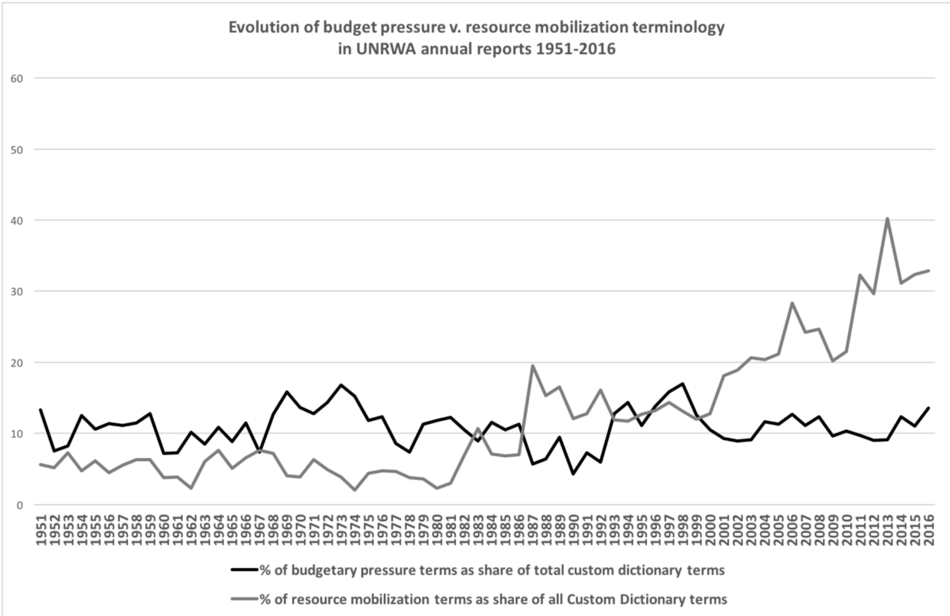
- **Assumption:** budgetary pressure coincides with sentiment: negative polarity

- **Lexicoder Sentiment Dictionary** (Young/Soroka 2012)
  - designed to capture the sentiment of political texts
  - 1710 positive and 2858 negative words
- Compute negativity of report: neg words / all sentiment words

- **Custom-built lexicon for matters of budget**
  - 18 budgetary pressure terms
  - 11 resource mobilization-related terms
  - 9 terms of financial matters, in particular income
  - 4 expenditure-related terms
- Compute share of budgetary pressure terms among all budget terms in report



R. Patz et al.: International Public Administrations and the Perception of Crisis: The Case of UNRWA and Palestine Refugees. Panel "*Regional and Global Crisis Management*" at ECPR General Conference, Oslo 2017



R. Patz et al.: International Public Administrations and the Perception of Crisis: The Case of UNRWA and Palestine Refugees. Panel "*Regional and Global Crisis Management*" at ECPR General Conference, Oslo 2017

(5) Authors‘ conclusions

- **Interpretation** has to be done with caution, but:
  – Sentiment detection is able to detect shifts in negativity over time,
  – which can be interpreted as shifts of crisis perception. (Correlation with events)
  – Diverging results in late 1980s: policy crisis increases, budgetary stress decreases
  => due to successful resource mobilization?

(6) Parliamentary debate

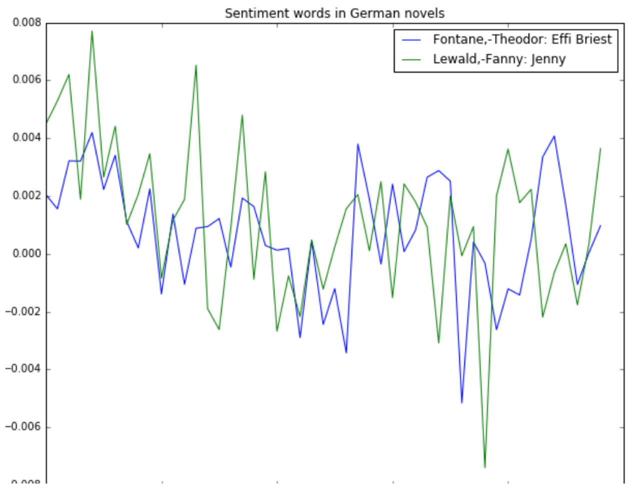
- **House of Commons (GB)**
  – Motion
  – Speeches
  – Vote
- Motion often contains sentiment (polarity)
  => polarity in speeches needs to be interpreted relative to the motion polarity!
- **Corpus: 1997 - 2017**
  – 1251 motion-speech units from 129 debates
  – unit has max. 5 utterances (avg. 1049 words)
  – motion also gets a government/opposition label
  – speech gets both a text-based label and a vote-based label (concurrence: 92.8%)

Gavin Abercrombie and Riza Batista-Navarro: A Sentiment-labelled Corpus of Hansard Parliamentary Debate Speeches . Proc. of Language Ressources and Evaluation Conference (LREC), 2018

Overview

- Sentiment analysis: Introduction, Terminology
- One system: SO-CAL
- Application: Political text
- **Application: Narrative**
- Viewpoint: Sentiment / DH

(1) Sentiment in novels



Happy ends?

Fotis Jannidis: Bedeutungsanalyse und distant reading. Ringvorlesung ‚Digital Humanities' (BBAW), 16.1.2018

(2) Sentiment in Lessing‘s dramas

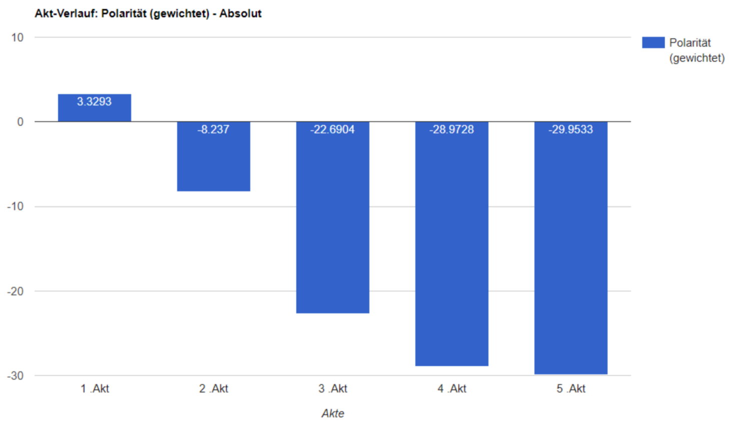


Figure 3. Polarity progression for Emilia Galotti per Act

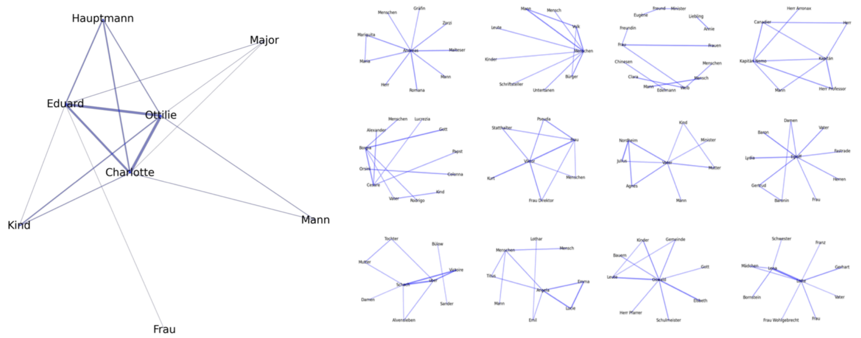
T. Schmidt, M. Burghardt: **An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing.** *Proc of Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, Santa Fe, 2018.

(2) Sentiment in Lessing‘s dramas

- **Manual annotation** of 200 snippets
  – agreement: kappa 0.4 => quite difficult
- **Automatic analysis** using different German sentiment dictionaries
  – accuracy: 70% (much lower than in other tasks)
- **Problems**
  – Orthography (e.g., *betriegen, bös*)
  – Vocabulary (e.g., *verdrießlich, pfui*)
  – Diachronic change (e.g., *Freier*)

T. Schmidt, M. Burghardt: **An Evaluation of Lexicon-based Sentiment Analysis Techniques for the Plays of Gotthold Ephraim Lessing.** *Proc of Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, Santa Fe, 2018.

(3) Character constellations



Fotis Jannidis: Bedeutungsanalyse und distant reading. Ringvorlesung ‚Digital Humanities' (BBAW), 16.1.2018

Overview

- Sentiment analysis: Introduction, Terminology
- One system: SO-CAL
- Application: Political text
- Application: Narrative
- **Viewpoint: Sentiment / DH**

Manfred Stede (Uni Potsdam)

Summary

Methods

- Lexicon-based vs machine learning
- Word counting
- Graded judgement
- Context: negation, irrealis, ...
- Example: SO-CAL
- Polar facts
- Aspect-based analysis
- Fine grain: source, target, opinion
- Entity level: sentiment „flow“

Applications

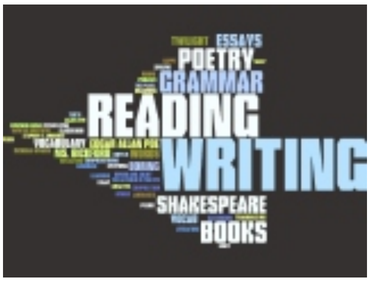
- Product reviews
- Political text
  - Word matching for Twitter analysis
  - Word matching for crisis perception in annual reports
  - Parliamentary debate
- Narrative
  - Polarity development in drama
  - Happy end?
  - Character constellations

Viewpoint: Sentiment analysis in DH

- Sentiment analysis has great potential for interesting and fruitful applications in the Humanities.

Viewpoint: Sentiment analysis in DH

- Using quite simple methods, for certain texts and purposes one can achieve good results.
- But not for all texts and purposes.
- The „wordle trap“
  - quality evaluation?
  - dependence on domain and genre
    - (e.g., *small*)



Manfred Stede (Uni Potsdam)

Viewpoint: Sentiment analysis in DH

https://bigdatablog.de/2016/01/05/digital-humanities-ueber-das-potenzial-von-big-data-in-den-geistes-und-kulturwissenschaften/

Viewpoint: Sentiment analysis in DH

- „Read close (too)!“
- Context matters – but in what ways, exactly?
- Domain and genre matter – but where do we draw the lines?
- Annotation guidelines for difficult sentiment problems => let‘s debate!
- Annotated corpora => let‘s debate!
- => Machine learning

Manfred Stede (Uni Potsdam)

Thank you!