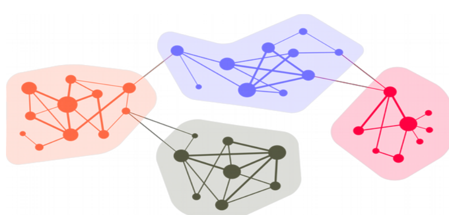
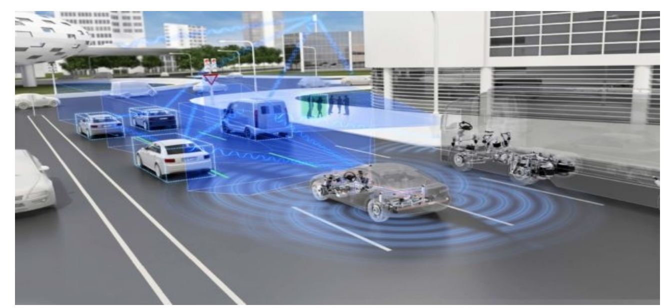
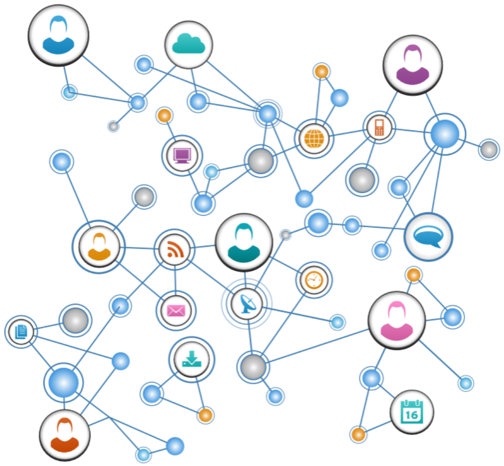
Political Ideology and Affective Attitudes Towards Mobility Innovation: Comparing Survey Responses and Twitter Data

Nikolas Zöller, Ingo Wolf, Tobias Schröder





Debate in social science:



New computational vs. traditional methods

- Replace?

- Complement?

- Strengths, weaknesses, challenges, opportunities

→ Need for more case studies!

Lazer et al. (2009). Life in the network: The coming age of computational social science
Serfass et. al. (2017) Big data in psychological research
Schober et al. (2016) Social media analyses for social measurement

Case Study: Affective Attitudes Towards Mobility Innovation

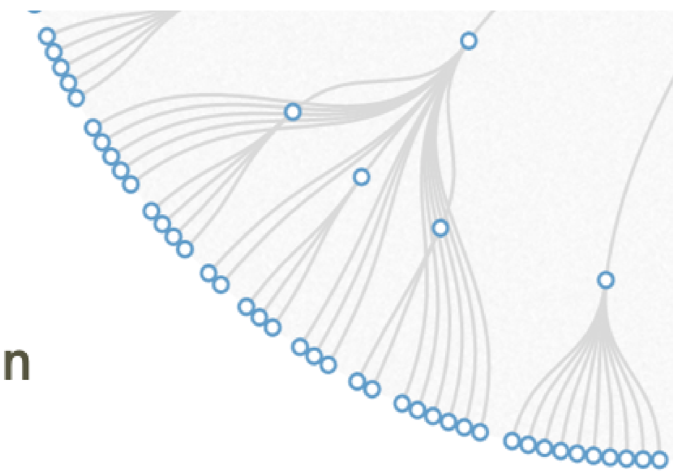
Social media analysis

Classical empirical online survey

FHP:-) University of Applied Sciences Fachhochschule Potsdam

MON FOR SENSE

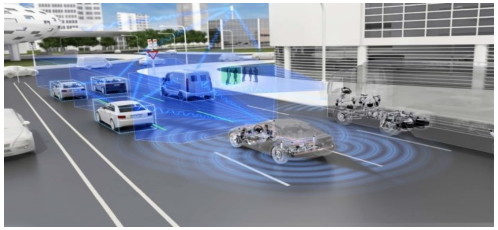
Monitoring und Foresight von Sentiments zu Mobilitätsinnovationen



Freie Universität Berlin

https://monforsense-results.fh-potsdam.de

Examined mobility innovations



Research questions

How do affective attitudes towards different mobility innovations differ?

Do more general political worldviews influence affective attitudes towards mobility innovations?

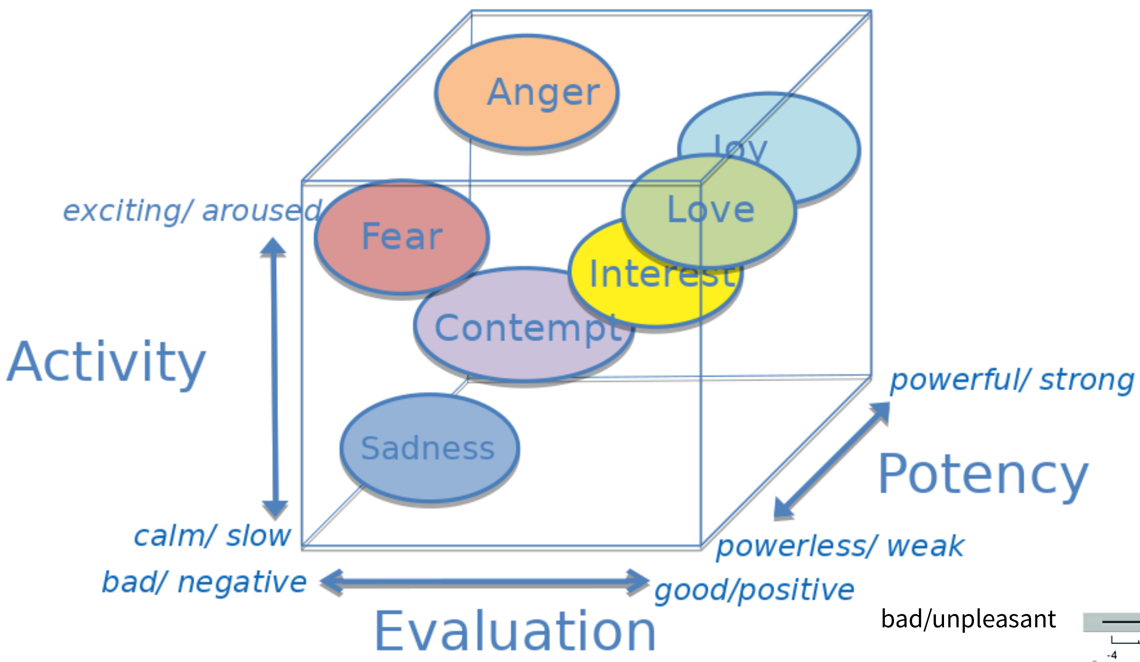
Are Twitter analysis and traditional survey methods suitable to study these questions and how do these methodologies compare?

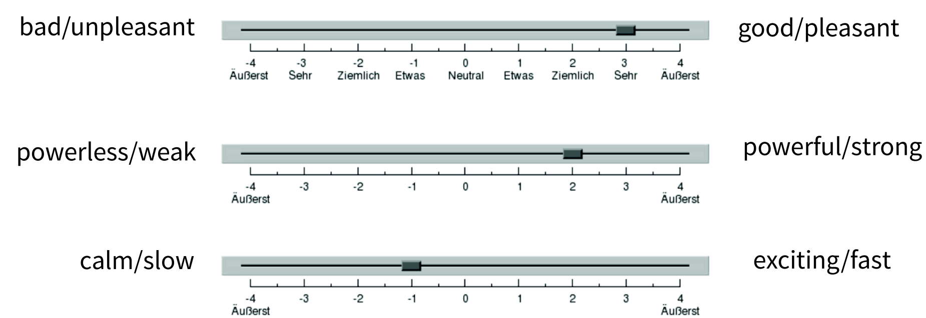
Kahan, D. M. & Braman, D. (2006). Cultural cognition and public policy

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition.

Affective Meaning

Concepts have, apart from a denotative meaning, a connotation along the three dimensions: Evaluation, Potency and Activity (EPA)





Osgood, C. E. (1962). Studies on the generality of affective meaning systems
Heise, D. R. (2010). Surveying cultures: Discovering shared conceptions and sentiments

Online Survey

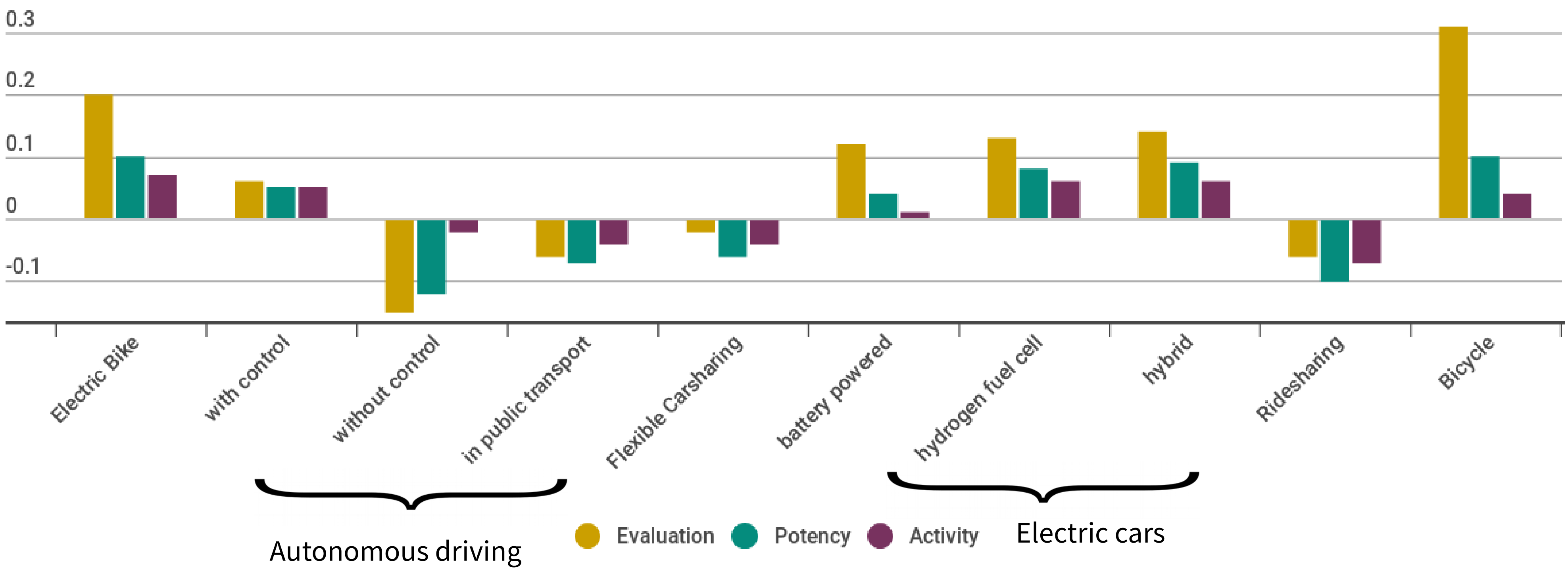
- N = 6047 participants recruited through commercial online panel
- Quotas for: age, sex, education, income, home town according to census data for Germany

  → (quasi-)representative

- Participants were asked for their political party preference
- Participants rated different transport options (electric cars, carsharing, etc.) on Evaluation, Potency, Activity - Scales

Survey results

Average sentiments of survey participants towards different mobility innovations

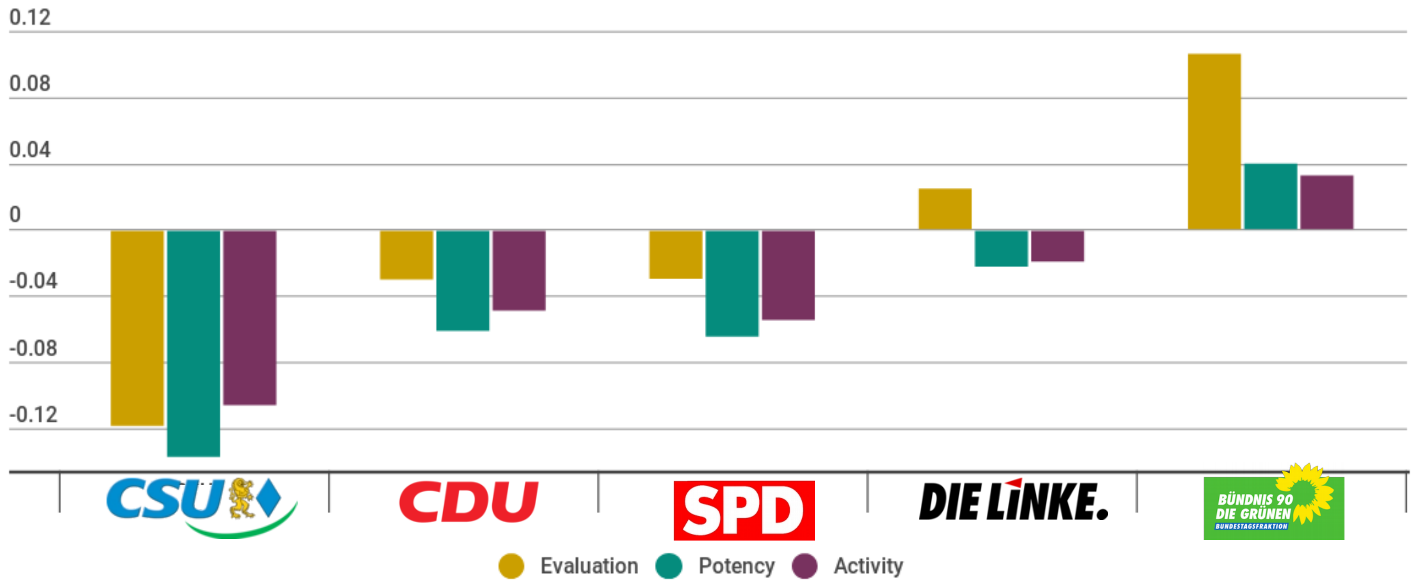


More survey results https://monforsense-results.fh-potsdam.de

Survey results

Average sentiments of participants with shared political party preference

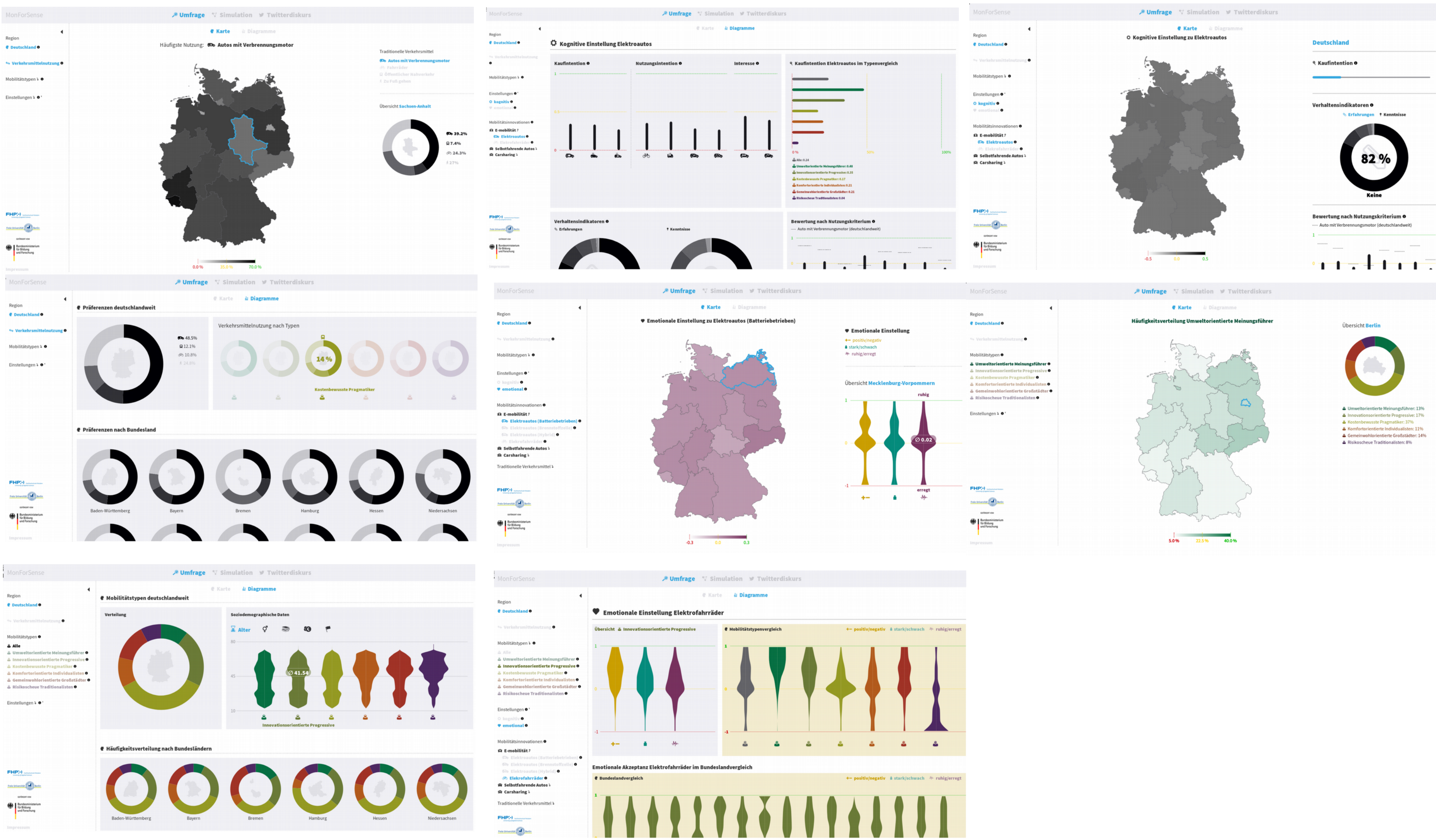
Example: Flexible Carsharing



Survey in May 2016

More survey results https://monforsense-results.fh-potsdam.de

Survey results



More survey results https://monforsense-results.fh-potsdam.de

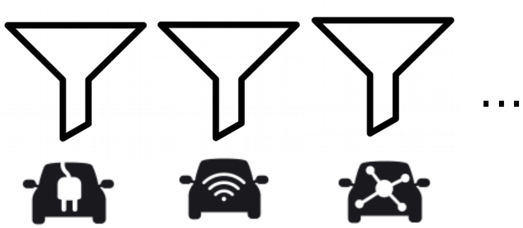
Twitter analysis

Data collection

- Collected German Tweets for 1 year via streaming API (Jan 2016 – Jan 2017)

- **Filter 1**: stopword list with 400 most frequent German words, language classification Software

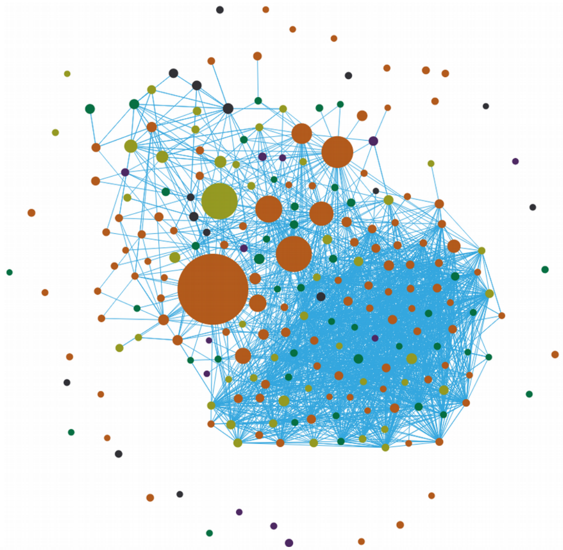
  → 370 million German tweets

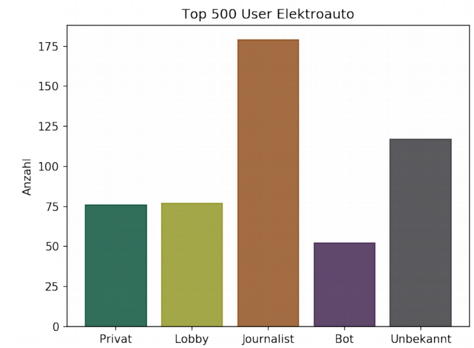


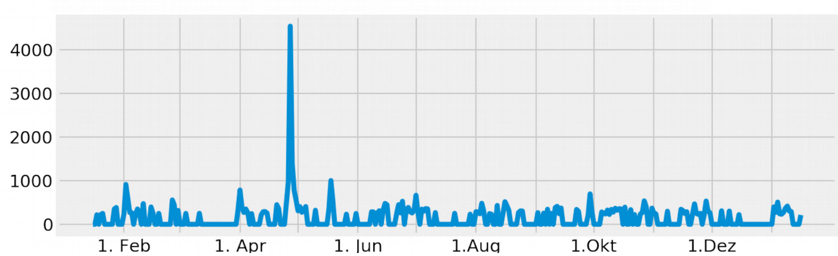
- **Filter 2**: topical key words, i.e. 'electric car', 'autonomous driving', 'e-bike'...

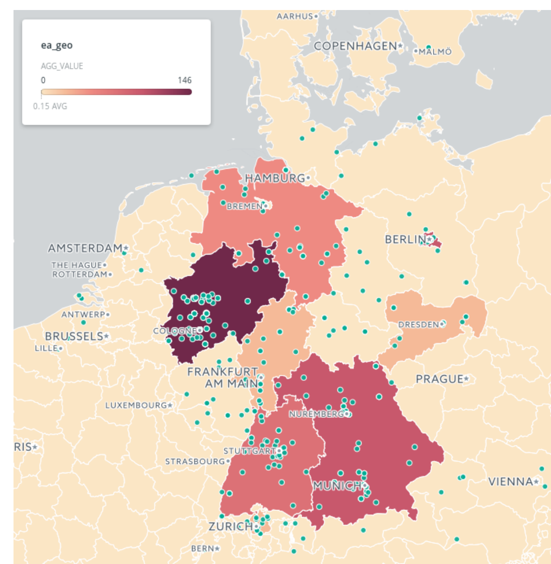
Scheffler, T. (2014). A German Twitter snapshot.

Twitter data analysis







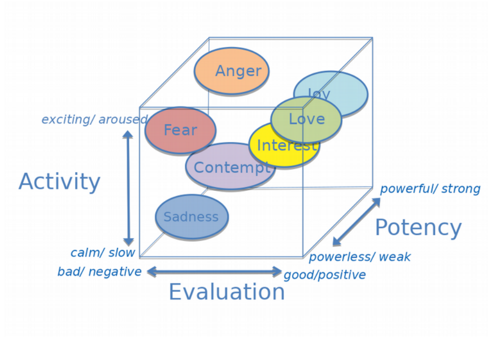


https://monforsense-results.fh-potsdam.de

Sentiment analysis

Problem: How to automatically classify a tweet in the dimensions Evaluation, Potency and Activity?

Tweet



E,P,A

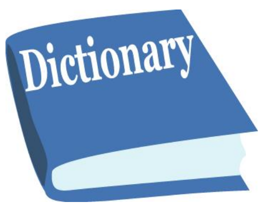
Example:

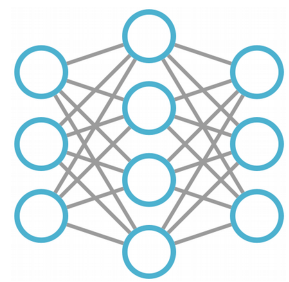
Software-Fehler: Selbstfahrendes #Google-Auto rammt Bus in Kalifornien! https://t.co /8hS6Mortnk
@B***o 01.03.2016 07:44:22 h
Sentiments: -0.48 0.08 0.15

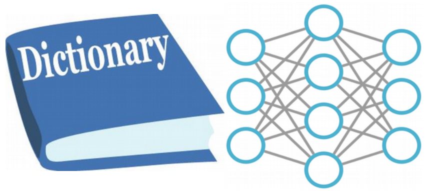
E P A

“Software failure: Self driving #Google-Car crashes into bus in California!”

Sentiment analysis methods







Lexicon methods

Dictionary lookup
+ rule based

Machine/Deep learning

- handcrafted features
  (e.g. n of words, n of #, n of !, ...)
- or automatically extracted features
  through deep learning

Hybrid approach

Example:

failure – E: -0.53 P: 0.25 A: 0.23
to crash – E: -0.40 P: 0.28 A: 0.36

"Software failure: Self driving #Google-Car crashes into bus in California!"

Medhat et. al. (2014), Sentiment analysis algorithms and applications: A survey

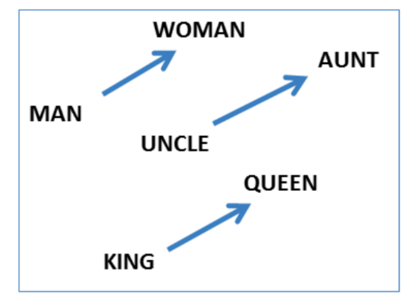
Word vectors and dictionary expansion

Base: 2753 labelled words + 500 mobility related labelled words

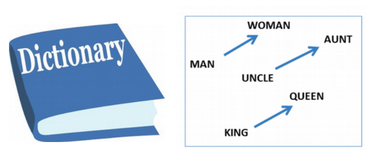
Goal: Dictionary Extension

1 **Train word vectors**

370 million German tweets → Word2vec word vectors Dim n=200



2 **Dictionary Extension** for vocabulary of word vectors



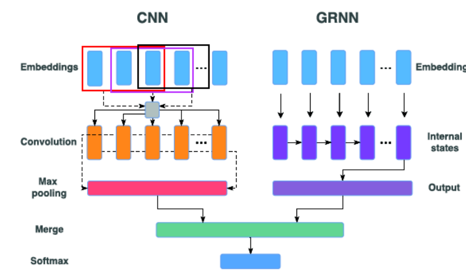
→ Support Vector Regression → EPA label prediction for words

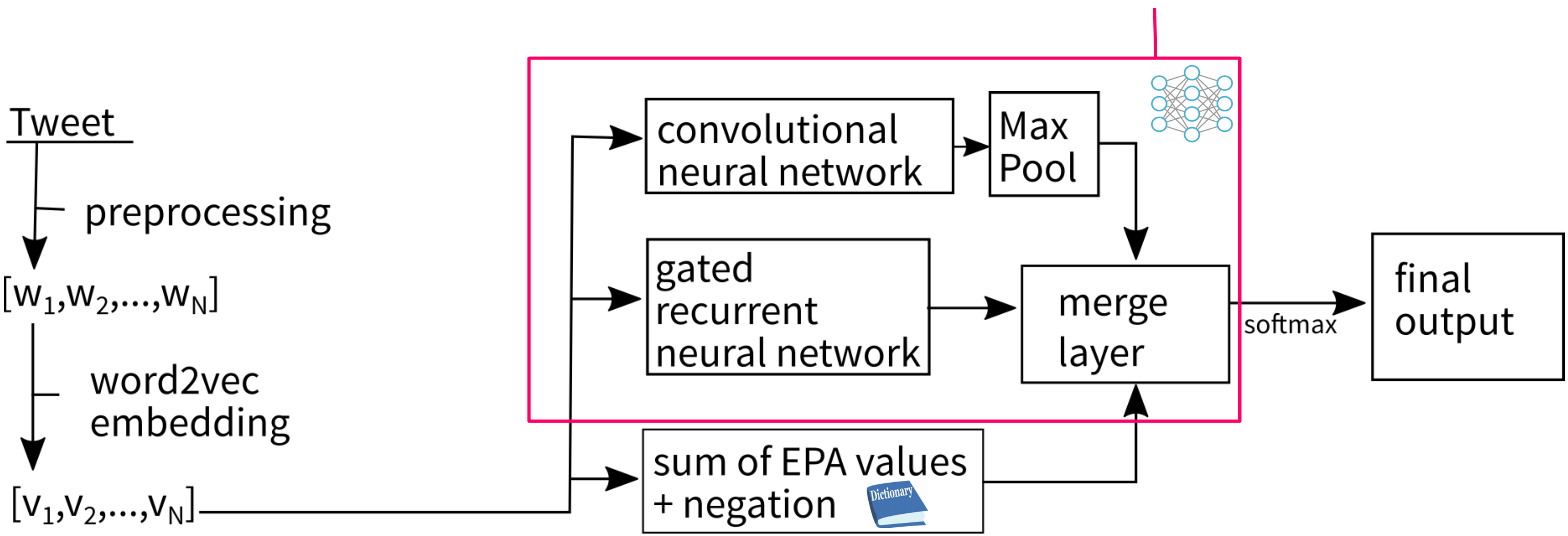
Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space
Alhothali, A. & Hoey, J. (2017). Semi-supervised affective meaning lexicon expansion using semantic and distributed word representations

Model architecture:

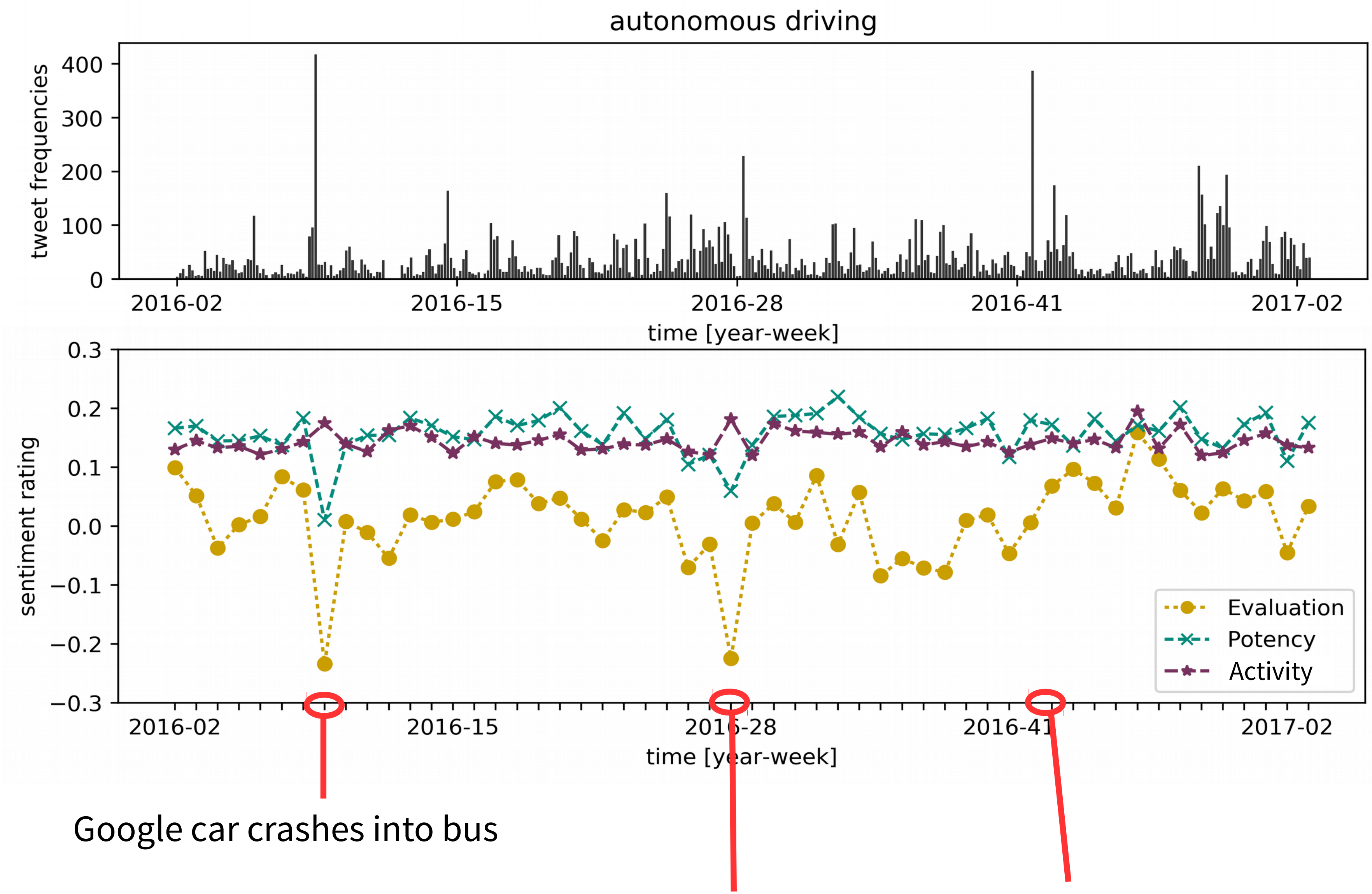
Training data: 1500 labelled tweets. Labelled by 10 raters each.
Data augmentation

Base model: Finki (Stojanovski et al, 2016)
- Performed well on SemEval-2016 Task 4: Twitter Sentiment analysis





Example time series for autonomous driving

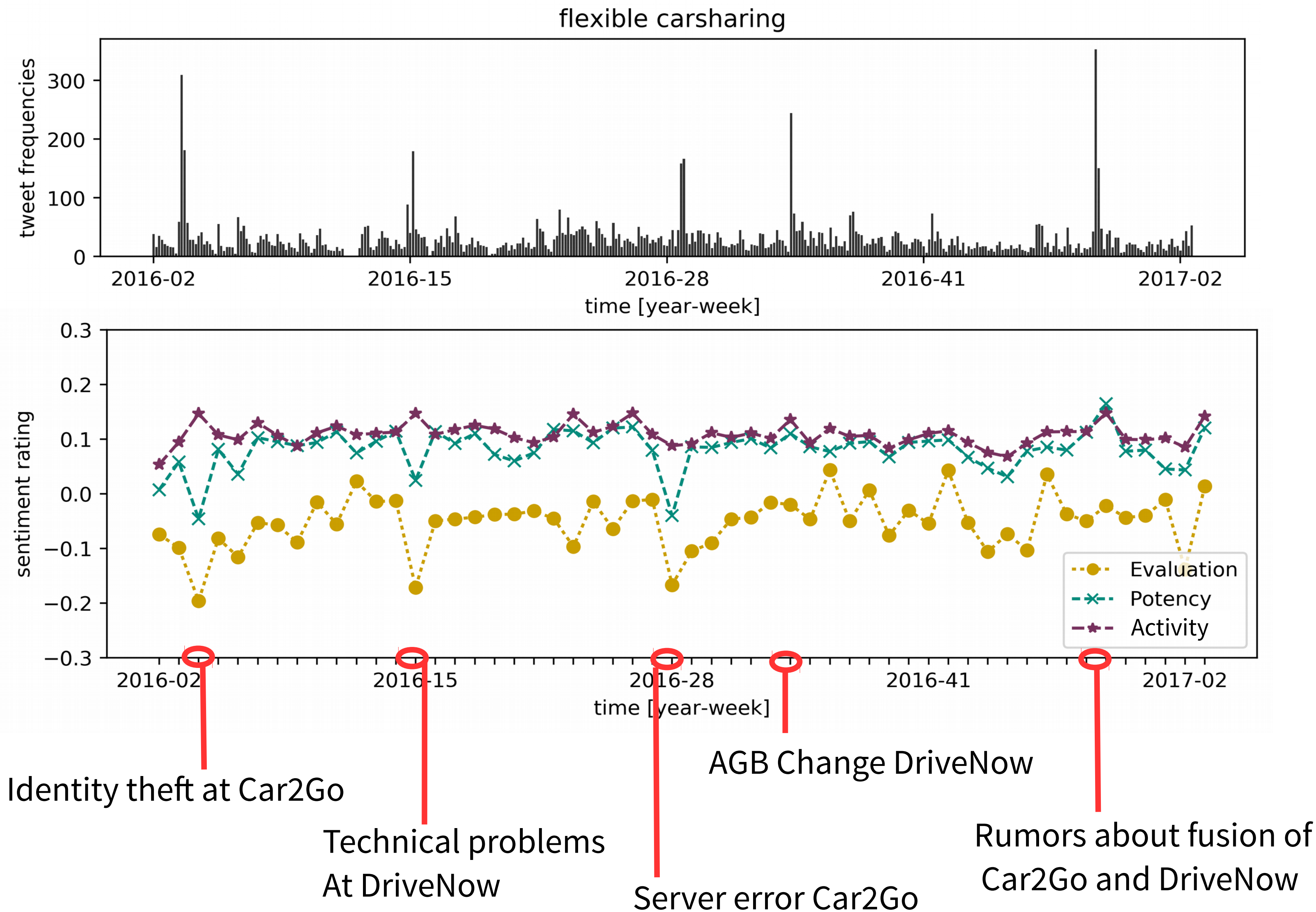


Google car crashes into bus

Tesla car deadly accident

News that Tesla will include autonomous technology in all their cars

Example Time series for flexible carsharing



Comparison survey and twitter data

Average E,P,A values/relative rankings

	survey [E,P,A]	Ranking [E,P,A]	Twitter [E,P,A]	Ranking [E,P,A]	n tweets
e-bike	[0.20, 0.10, 0.07]	[2,2,1]	[0.03, 0.12, 0.14]	[1,4,2]	20,033
autonomous driving av.	[-0.05, -0.05, -0.02]	[5,4,4]	[0.00, 0.16, 0.14]	[3,1,1]	11,378
- with control	[0.06, 0.05, 0.05]	-			
- without control	[-0.15, -0.12, -0.06]	-			
- in public transport	[-0.06, -0.07, -0.04]	-			
flexible carsharing	[-0.02, -0.06, -0.04]	[4,5,5]	[-0.06, 0.10, 0.11]	[5,5,4]	10,953
electric car av.	[0.13, 0.07, 0.05]	[3,3,2]	[-0.01, 0.14, 0.11]	[4,2,5]	58,268
battery powered car	[0.12, 0.04, 0.01]	-			
hydrogen fuel cell car	[0.13, 0.08, 0.06]	-			
hybrid car	[0.14, 0.09, 0.06]	-			
ridesharing	[-0.06, -0.10, -0.07]	[6,6,6]	[-0.07, 0.09, 0.11]	[6,6,6]	9,011
bicycle	[0.31, 0.10, 0.04]	[1,1,3]	[0.03, 0.12, 0.12]	[2,3,3]	172,189

Kendall's ranking coefficient:

$\tau_E = 0.60 \quad \tau_P = 0.33 \quad \tau_A = 0.33$

Political party networks on Twitter

- Twitter handles of politicians in the German parliament
- Follower networks → Identified party affiliated twitter users
- Filtered tweets with key words, i.e. “electric car”, “autonomous driving”,...

Political party networks on Twitter

Numbers of party affiliated twitter handles and topical tweets

	SPD	CDU	CSU	DIE LINKE.	BÜNDNIS 90 DIE GRÜNEN
N party-affiliated ids	124,266	181,827	37,927	152,352	176,436
N collected topical tweets					
e-bike	646	937	212	774	1,586
autonomous driving	651	763	299	365	1,206
flexible carsharing	491	613	339	376	1,268
electric car	2,995	5,692	1,180	2,894	9,281
ridesharing	227	290	71	408	815
bicycle	5,942	8,624	2,642	10,850	16,042

Comparison survey responses and Twitter data

Survey data: | Twitter data:

E-bike:

Survey data – Ranking:

	E	P	A
1.	BÜNDNIS 90 DIE GRÜNEN	BÜNDNIS 90 DIE GRÜNEN	BÜNDNIS 90 DIE GRÜNEN
2.	CDU	CDU	CDU
3.	SPD	SPD	DIE LINKE.
4.	DIE LINKE.	DIE LINKE.	SPD
5.	CSU	CSU	CSU

Twitter data – Ranking:

	E	P	A
1.	SPD	BÜNDNIS 90 DIE GRÜNEN	CSU
2.	BÜNDNIS 90 DIE GRÜNEN	CDU	DIE LINKE.
3.	CDU	SPD	SPD
4.	DIE LINKE.	CSU	BÜNDNIS 90 DIE GRÜNEN
5.	CSU	DIE LINKE.	CDU

Kendall rank correlation

$\tau_E = 0.60$ $\tau_P = 0.80$

$\tau_A = -0.60$

Ridesharing:

Survey data – Ranking:

	E	P	A
1.	BÜNDNIS 90 DIE GRÜNEN	BÜNDNIS 90 DIE GRÜNEN	BÜNDNIS 90 DIE GRÜNEN
2.	DIE LINKE.	DIE LINKE.	DIE LINKE.
3.	CDU	CDU	CDU
4.	SPD	SPD	SPD
5.	CSU	CSU	CSU

Twitter data – Ranking:

	E	P	A
1.	BÜNDNIS 90 DIE GRÜNEN	BÜNDNIS 90 DIE GRÜNEN	CDU
2.	CDU	DIE LINKE.	SPD
3.	CSU	CSU	CSU
4.	DIE LINKE.	SPD	DIE LINKE.
5.	SPD	CDU	BÜNDNIS 90 DIE GRÜNEN

Kendall rank correlation

$\tau_E = 0.40$ $\tau_P = 0.40$

$\tau_A = -0.40$

Comparison survey responses and Twitter data

Rank correlations of twitter data and survey data

Rankings with respect to party affiliation

	τ Evaluation	τ Potency	τ Activity
e-bike	0.60	0.80	-0.60
autonomous driving	-0.20	0.40	0.20
flexible carsharing	0.20	0.40	0.40
electric car	0.20	0.40	0.20
ride-sharing	0.40	0.40	-0.40
bicycle	0.80	0.80	-0.60

Overall correlation:

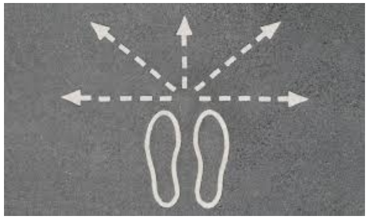
$\tau_E = 0.33$ $\tau_P = 0.53$ $\tau_A = -0.13$

$p_E < 0.05$, $p_P < 0.001$, $p_A = 0.37$

Moderate alignment between results in the dimensions Evaluation and Potency, no correlation between Twitter analysis and survey responses in the dimension Activity

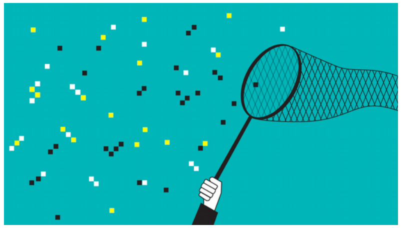
Challenges and Limitations

Sentiment Analysis



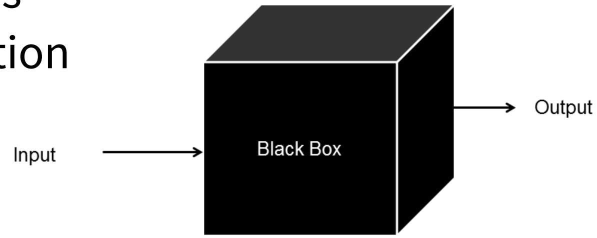
1. Immense researcher's degree of freedom
   A lot of different algorithmic architectures to pick from

2. Published models often depend on very specific data sets and domains
   → not always clear in how far they generalize



3. Training data is crucial
   → quality
   → quantity

4. Deep learning techniques are to some extend black boxes
   → difficult to know what exact feature drove a classification process



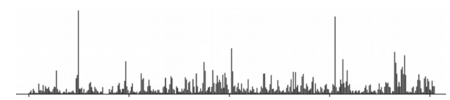
Marcus, G. (2018). Deep learning: A critical appraisal

Challenges and Limitations

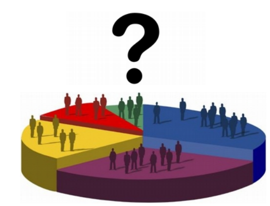
Difficulties in comparing Twitter analysis with survey responses

1. Twitter activity is event driven - reflective of discourse in wider media ecosystem
   → fluctuations in the results of sentiment analysis

2. Twitter analysis was done over a longer period while
   the survey data was collected within one week



3. Sociodemographics of Twitter users are unknown
   → probably not representative of the general population
   → previous studies have found gender and age bias



4. Content-based sentiment analysis disregards importance of
   communicative intent and impression management

Mislove et. al. (2011) Understanding the demographics of twitter users

Conclusions:

Both survey and twitter data show distinctions between the EPA-profiles towards mobility innovations for different party-affiliated groups.

We found moderate correlations between results of analysis of survey data and sentiment analysis of twitter data (but only in the dimensions evaluation (E) and potency(P)).

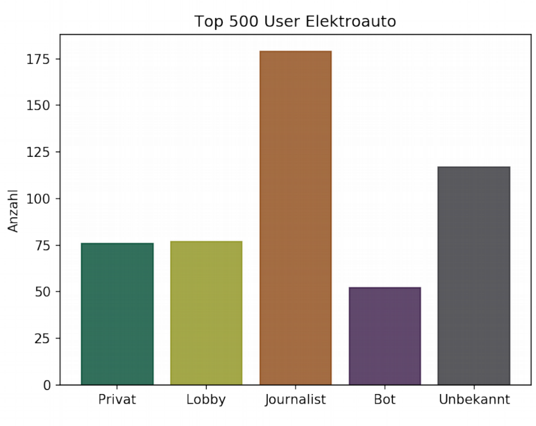
There are a lot of challenges in comparing Twitter data with traditional survey responses, nevertheless sentiment analysis shows great promise for social psychological attitude research, not in replacing survey methods but by complementing them.

In particular, sentiment analysis of social media data allows to examine the dynamics of attitude formation
→ analysis of time series
→ it can be used to identify emotionally charged events that influence attitude formation

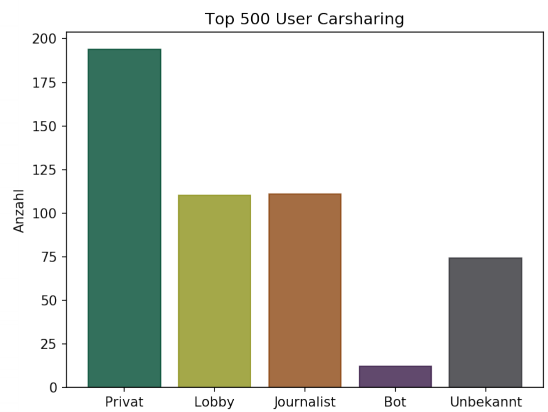
Contact: Nikolas Zöller zoeller@fh-potsdam.de

Types of Twitter users

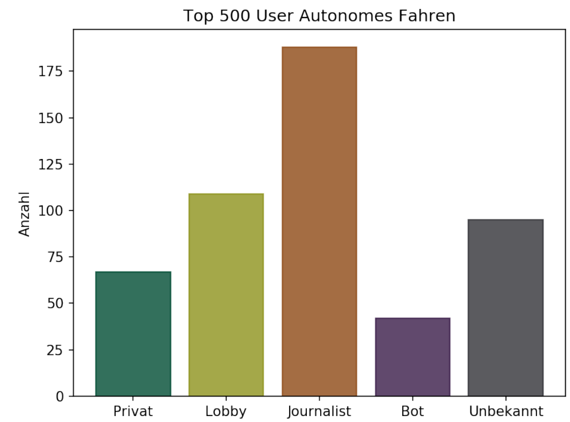
Electric car



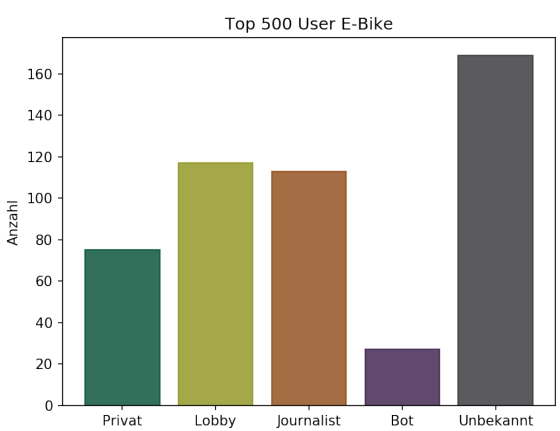
Flexible carsharing



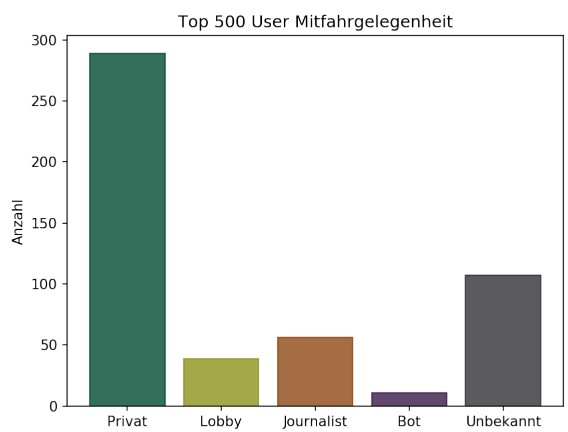
Autonomous driving



Electric bicycle



Ridesharing



- Privater User
- Blogger/Journalist
- Corporate User/ Lobbyist
- Bot
- Unbekannt

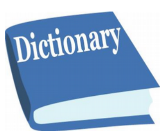
Model details

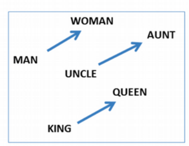
Preprocessing

- Remove URLs, Tabs, Newline Characters
- Replace Umlaute (i.e. ä → ae)
- Remove Punctuation except emoticons, !,?
- Everything to lower case
- Check if word is in word2vec vocabular; if not spell check, if still not → remove
- Write to array and pad sequences

Automatic dictionary expansion with support vector regression

- 2753 labelled words
- 500 mobility related labelled words rated by 30 people each





→

Support Vector Regression

→

EPA label prediction for words

Metrics evaluated on test split and model parameters

	Kendall τ	F1-binary	MAE	C	γ
E	0.61	0.86	0.19	75	0.001
P	0.43	0.87	0.23	11	0.002
A	0.47	0.84	0.15	11	0.002

Training of the model

- Labelled tweets are split into training and validation set
- Upsamling of training set to avoid class bias
- Loss function: $$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i|$$
- Optimizer: rmsprob

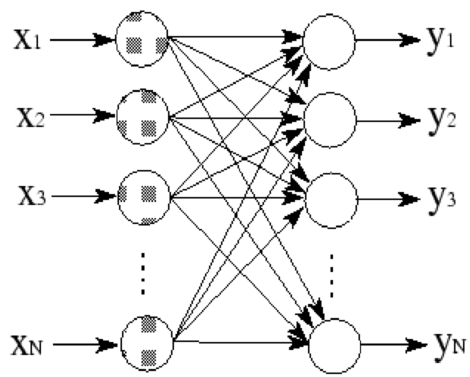
$$w := w - \frac{\eta}{\sqrt{v(w,t)}}\nabla Q_i(w)$$

$$v(w,t) := \gamma v(w,t-1) + (1-\gamma)(\nabla Q_i(w))^2$$

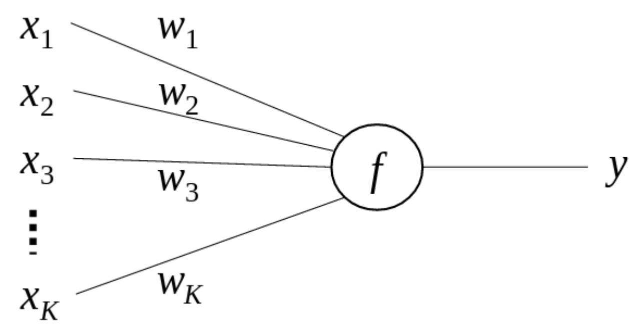
Evaluation metrics on test split:

Accuracies=[0.68, 0.67,0.69] $MAE_w = [0.18, 0.175, 0.14]$

Simple neural net



Single neuron:



activation function f: $y = f(u)$

$$u = \sum_{i=0}^{K} w_i x_i$$

Dot product with row vectors of the weight matrix W

Softmax activation: $$y_j = \frac{\exp(u_j)}{\sum_{j'=1}^{V} \exp(u_{j'})}$$

Normalized, values between 0 and 1
→ probability distribution over possible labels

Word2Vec Embeddings

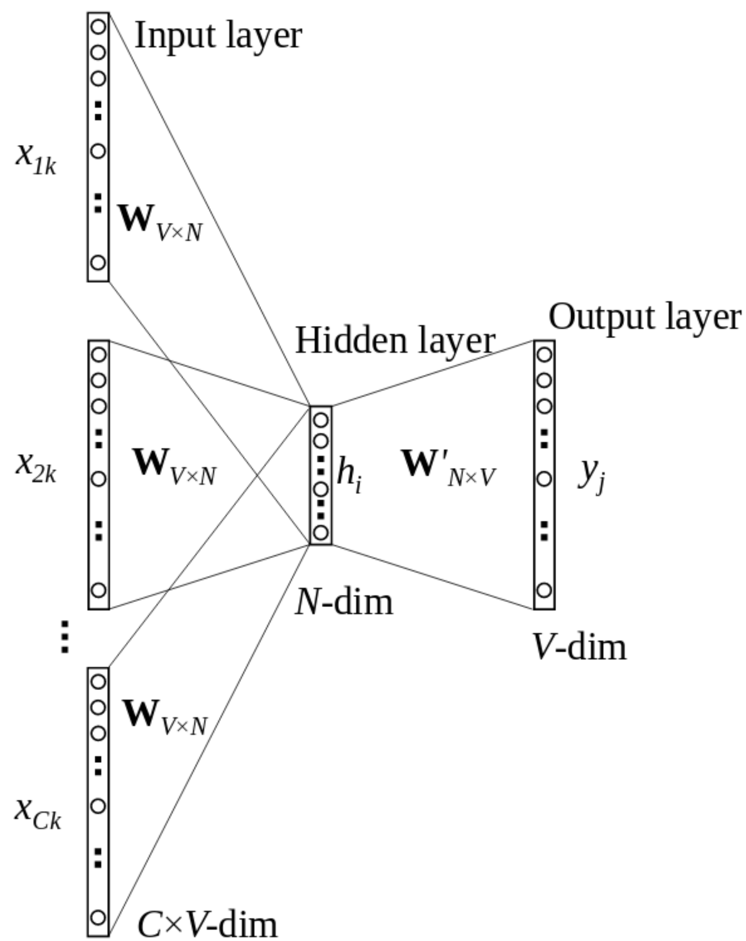
What ? Model to represent words as vectors in n-dimensional vector space

How ? 2 – layered neural net that is trained with a huge corpus of sentences/tweets

Why ? Linguistic context is 'encoded' into the word vectors.
i.e. words that occur often with other words will be close to each other in the vector space.

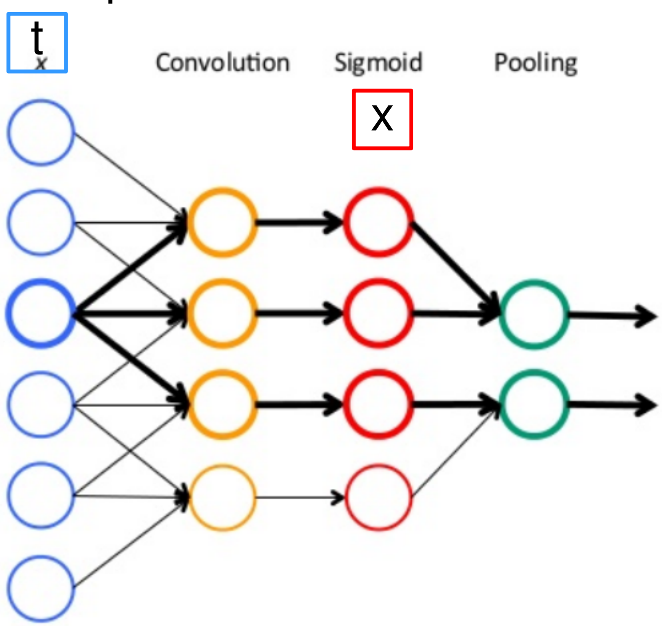
Word2vec Model

X_n : words after preprocessing

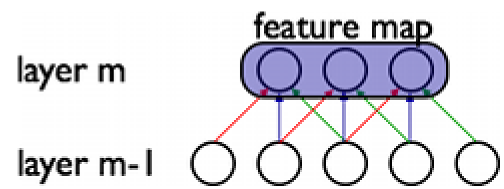


Convolutional Neural Network

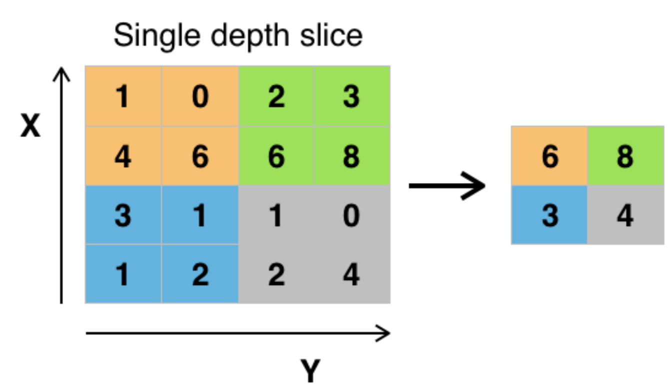
Principle:



$$x_i = f(W_c \cdot t_{i:i+h-1} + b_c)$$



Maximum pooling:

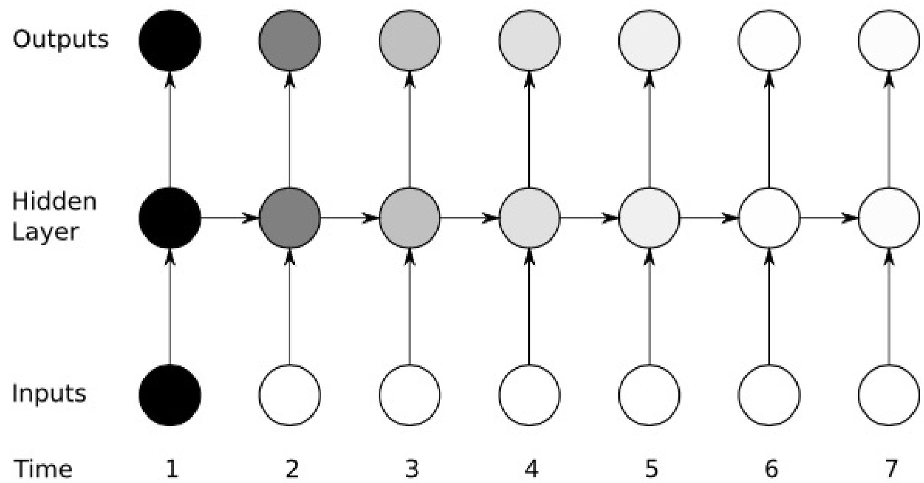


Sigmoid funcion:

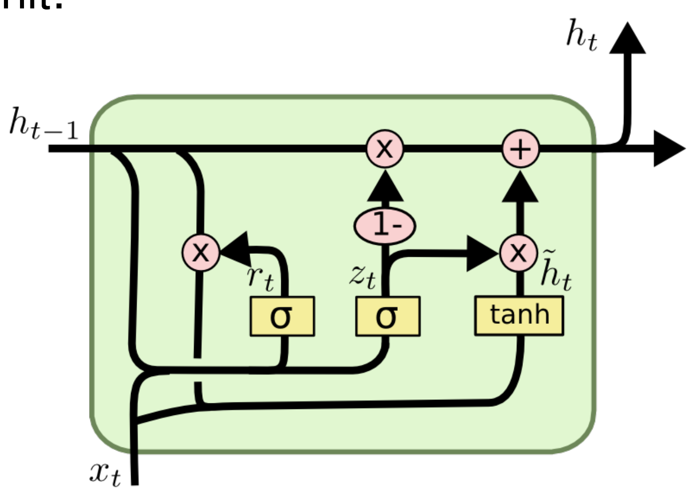
$$S(t) = \frac{1}{1 + e^{-t}}$$

Gated recurrent neural network

RNN:
Vanishing/
Exploding
gradient
problem



GRUnit:



$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$