SLANG Lab @ UMass NLP

Socio-statistical **Lang**uage Analysis

Natural language processing *and*
computational social science

slanglab.cs.umass.edu nlp.cs.umass.edu

Candidate Friday, March 2022

Talk to my PhD students! →

Brendan O'Connor

Tessa Masis

Ankita Gupta

and lab alumni! →

Katie Keith (→AI2, Williams)

Su Lin Blodgett (→MSR)

Abe Handler (→CU Boulder)

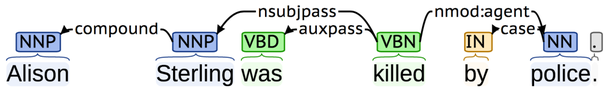
Social Factors
(location, gender, race…)

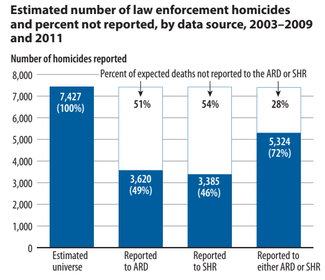
How do social factors affect NLP?

How can NLP help learn about society?

NLP
(models, linguistics, data…)

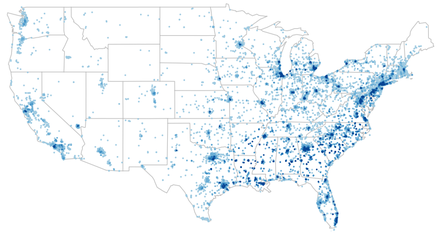
Social event extraction: Tracking police fatalities from news reports





Social media dialects: Racial disparities in NLP

(w/ Lisa Green, Linguistics)



```
>>> s = 'Bored af den my phone finna die!!!'
>>> langid.classify(s)
('da', 0.9999968001354156)
```

	AA-Aligned	White-Aligned
langid.py	13.2%	7.6%
Twitter	24.4%	17.6%

Disparate error rates

- **Many other projects, too!**
  - Computational linguistics of African-American English *(w/ Kristine Yu, Lisa Green, Linguistics)*
  - Misinformation tracking regarding Asian-American and Latin communities *(FACT-CHAMP; w/ Jonathan Ong, Communications)*
  - Corpus statistical inference: categories, entities, syntactic dependencies, events
  - Subjective knowledge and viewpoint modeling
  - Crowdsourced coreference *(w/ Mohit Iyyer, CS)*
  - Visualizing news corpora *(w/ Narges Mahyar, CS)*