Applications

Computational Linguistics: Jordan Boyd-Graber
University of Maryland
COMPUTATIONAL SOCIAL SCIENCE

Alternatives / Complements to Traditional Social Science

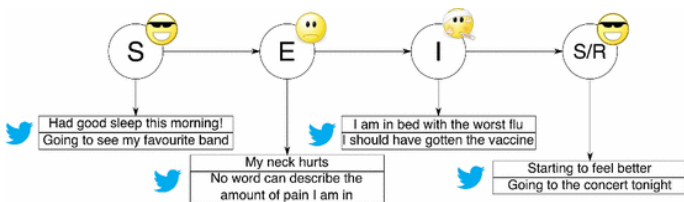|	JUN 21-25	B+	Suffolk University	500 LV	Wolf	49%	36%	Wagner	Wolf +13
al	JUN 23-24	B	YouGov	1,000 A	Approve	45%	49%	Disapprove	Disapprove +4
	JUN 22-24	B	YouGov	1,000 A	Approve	39%	50%	Disapprove	Disapprove +11
	JUN 20-24	C+	Rasmussen Reports/Pulse Opinion Research	1,500 LV	Approve	46%	52%	Disapprove	Disapprove +6
	JUN 18-24	B	Gallup	1,500 A	Approve	41%	55%	Disapprove	Disapprove +14
	JUN 20-21	B	YouGov	1,000 A	Approve	38%	51%	Disapprove	Disapprove +13
	JUN 16-19	B	Public Opinion Strategies	800 A	Approve	41%	47%	Disapprove	Disapprove +6
	JUN 22-24	B-	Morning Consult	1,989 RV	Democrat	44%	37%	Republican	Democrat +7
J-11	JUN 22-25	A+	Monmouth University	339 LV	Sherrill	45%	40%	Webber	Sherrill +5
J-11	JUN 22-25	A+	Monmouth University	406 V	Sherrill	40%	38%	Webber	Sherrill +2
1.	JUN 21-25	B+	Suffolk University	500 LV	Casey	47%	32%	Barletta	Casey +15
a.	JUN 21-25	A-	Quinnipiac University	1,153 RV	Kaine	54%	36%	Stewart	Kaine +18

- Polls
- Surveys
- Voting patterns

Social Media

- People share more readily
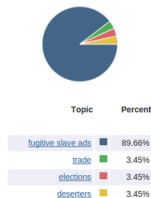- Self-censorship on official channels
- Faster feedback

Social Media

- People share more readily
- Self-censorship on official channels
- Faster feedback

Topic Models

TOPIC PROPORTIONS

Topic	Percent	
fugitive slave ads	■	89.66%
trade	■	3.45%
elections	■	3.45%
deserters	■	3.45%

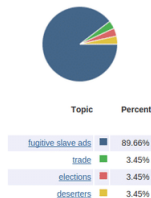*From the Wed., Dec. 25, 1861 issue*

—Ranaway from the subscriber, on the 3d inst., my slave woman PARTHENA. Had on a dark brown and white calico dress. She is of a ginger-bread color; medium size; the right fore-finger shortened and crooked, from a whitlow. I think she is harbored somewhere in or near Duvall's addition. For her delivery to me I will pay $10.

G. W. H. TYLER. de 6—ts

- Open-ended responses
- Questions you "forgot to ask"
- Help with interpretation

Topic Models

TOPIC PROPORTIONS

Topic	Percent	
fugitive slave ads	■	89.66%
trade	■	3.45%
elections	■	3.45%
deserters	■	3.45%

From the Wed., Dec. 25, 1861 issue

—Ranaway from the subscriber, on the 3d inst., my slave woman PARTHENA. Had on a dark brown and white calico dress. She is of a ginger-bread color; medium size; the right fore-finger shortened and crooked, from a whitlow. I think she is harbored somewhere in or near Duvall's addition. For her delivery to me I will pay $10.

G. W. H. TYLER. de 6—ts

- Open-ended responses
- Questions you "forgot to ask"
- Help with interpretation
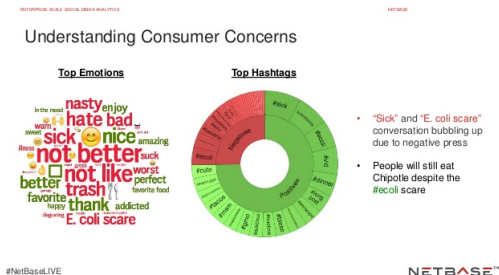- Tension: prediction or interpretation

Qualitative Analysis (Grounded Theory)

Grounded Theory

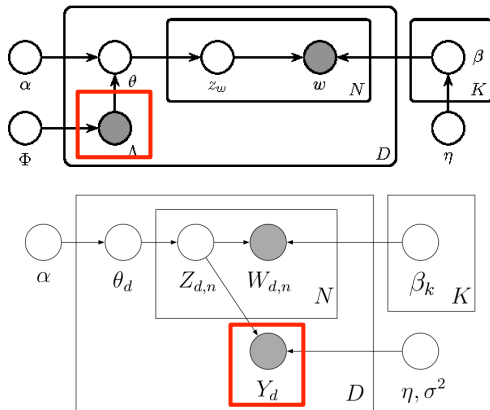Inductive coding of qualitative data to reach understanding

- low-quality models/theories and refining
- topics are "grounded" in specific word tokens
- Baumer et al. [2017]: compare themes
- model gives topic-specific "reading list"
- efficient, insulation from human biases.

Sentiment Analysis

- Given a piece of text, does it have positive or negative sentiment?
- Data: product reviews, social media
- Stanford Sentiment Treebank: syntactic annotations

Upstream vs. Downstream Model

- Is your prediction conditioned on text or vice-versa?
- Upstream: More interpretable, easier to train
- Downstream: Better prediction

Stance Prediction

- Predict how people feel about issue
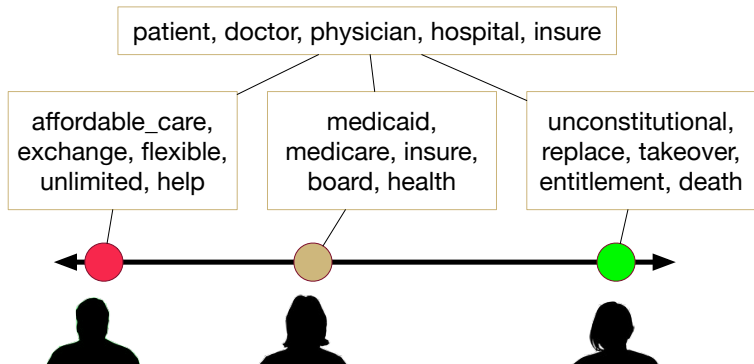- Traditionally: from votes

Stance Prediction

- Predict how people feel about issue
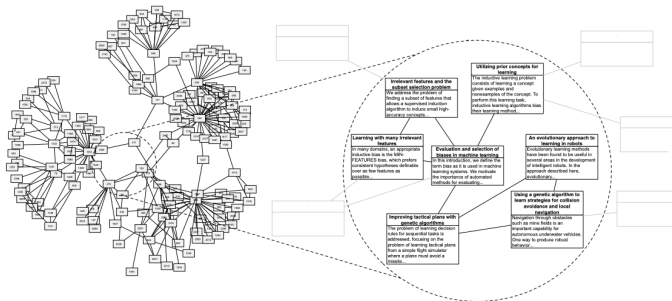- Traditionally: from votes . . . from text?

Framing

Issue: Healthcare

patient, doctor, physician, hospital, insure

affordable_care, exchange, flexible, unlimited, help

medicaid, medicare, insure, board, health

unconstitutional, replace, takeover, entitlement, death

Jointly learn topics and predictions.

Social Networks

- Predict links
- Predict when friendships end (and why)

More importantly . . .

- Suicide risk
- Unhealthy habits
- Bullying

More importantly . . .

- Suicide risk
- Unhealthy habits
- Bullying

Computational social science emphasizes interpretability, broader impacts, and captures what about language is unique to humans