

# Computational Analysis of Religious and Ideological Linguistic Behavior

---

University of Virginia, MS in Data Science Capstone

Samantha Garofalo, Seth Green, Leigh Harton, Megan Stiles



Advisor: Don Brown

Client: Peter Ochs

Sponsor: ARL

# Problem: How to work with Value-Based Groups?

What is a Value-Based Group?

- *Any group motivated by a set of values*

Who needs to work with Value-Based groups?

- *US Government*
- *Humanitarian Groups*
- *Business Leaders*

Large-Scale Goals:

- *Real-time recommendations to organizations on the ground*

Problem

Data

Analysis

Preliminary Results

# About Our Data

## Input

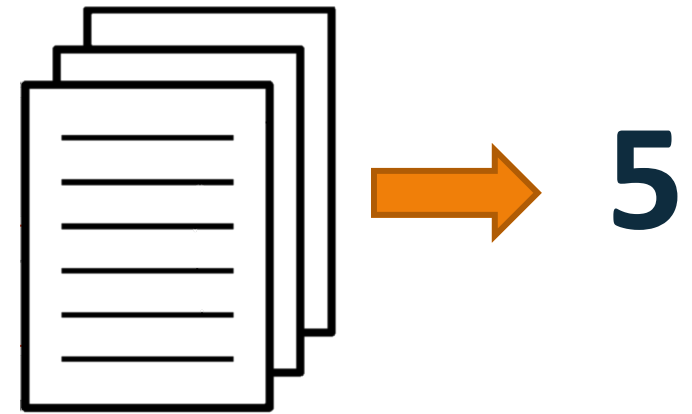
- *Raw ideological text*

## Output

- *Number on Linguistic Rigidity Scale*

## Sources:

- *Last year's corpus*
- *New Groups*
  - Web Scraping
  - OCR for PDFs



Problem

Data

Analysis

Preliminary Results

# Data Wrangling

- *Stemming*
- *Part of Speech Tagger*
- *Appropriate Article Length*

```
def cleanTokens(tokenList):  
    #Convert all text to lower case  
    textList=[word.lower() for word in tokenList]  
  
    #Remove punctuation  
    textList=[word for word in textList if word not in punctuation]  
    textList=["".join(c for c in word if c not in punctuation) for word in textList ]  
  
    #convert digits into NUM  
    textList=[re.sub("\d+", "NUM", word) for word in textList]  
  
    #Stem words  
    textList=[stemmer.stem(word) for word in textList]  
  
    #Remove blanks  
    textList=[word for word in textList if word!= ' ']  
    textList=[word for word in textList if word!= '']  
  
    #Extract tokens  
    return textList
```

Problem

Data

Analysis

Preliminary Results

## TFIDF Matrix

	freq	idf	logidf	tfidf	logtfidf
thee	18	139.000000	4.934474	2502.000000	88.820531
prophet	12	139.000000	4.934474	1668.000000	59.213687
me	12	139.000000	4.934474	1668.000000	59.213687
we	21	75.818182	4.328338	1592.181818	90.895101
uniti	9	139.000000	4.934474	1251.000000	44.410265
fate	8	139.000000	4.934474	1112.000000	39.475791
muhammad	8	139.000000	4.934474	1112.000000	39.475791
decre	8	139.000000	4.934474	1112.000000	39.475791
christ	10	104.250000	4.646792	1042.500000	46.467919

Problem


Data

Analysis

Preliminary Results

# Value Words

Previous Method



A word cloud of terms related to the 2008 US Presidential election. The words are arranged in a circular pattern. The most prominent words are 'worker' and 'love', both in large, bold, blue font. Other words include 'peter', 'room', 'am tri', 'want', 'ask', 'too', 'go', 'talk', 'poor', 'say', 'fr', 'done', 'christ', 'pray', and 'picket'. The words are in various colors (blue, orange, brown) and sizes, creating a dynamic and visually engaging composition.



# Judgment Classification

## Previous Method:

- *Judgment = (noun + “to be” verb + adj/adv)*

## New Method:

- *Judgment = pronoun + key word*

The first priority is to perform hijrah from wherever you are to the Islamic State, from dārul-kufr to dārul-Islām. Rush to perform it as Mūsā (‘alayhis-salām) rushed to his Lord, saying, {And I hastened to You, my Lord, that You be pleased} [Tāhā: 84]. Rush to the shade of the Islamic State with your parents, siblings, spouses, and children. There are homes here for you and your families. You can be a major contributor towards the liberation of Makkah, Madīnah, and al-Quds. Would you not like to reach Judgment Day with these grand deeds in your scales.

Second, if you cannot perform hijrah for whatever extraordinary reason, then try in your location to organize bay’āt (pledges of allegiance) to the Khalīfah Ibrāhīm. Publicize them as much as possible. Gather people in the masājid, Islamic centers, and Islamic organizations, for example, and make public announcements of bay’ah. Try to record these bay’āt and then distribute them through all forms of media including the Internet. It is necessary that bay’ah becomes so common to the average Muslim that he considers those holding back as grossly abnormal. This effort, inshā’allah, will encourage Islamic groups to abandon their partisanship and also announce their bay’ah to the Khalīfah Ibrāhīm.

If you live in a police state that will arrest you over such bay’āt, then use means of anonymity to con-

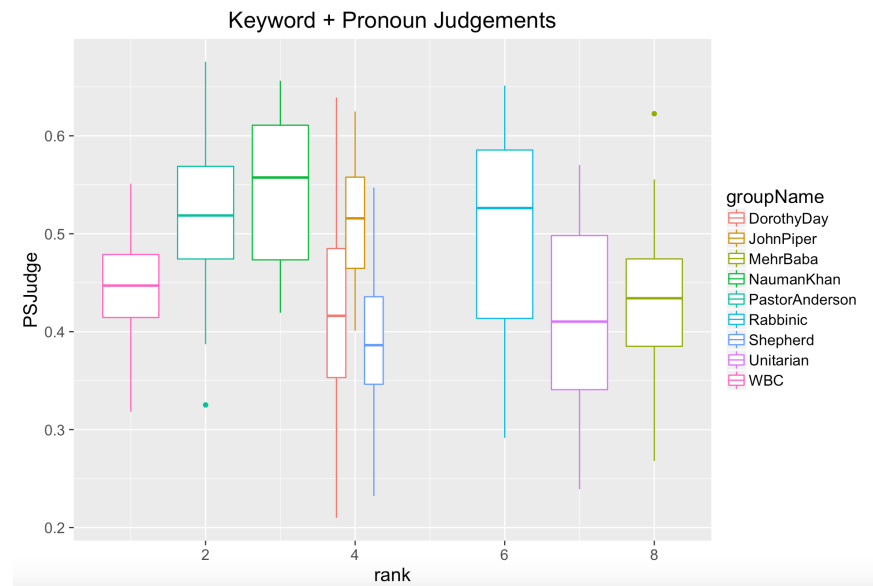
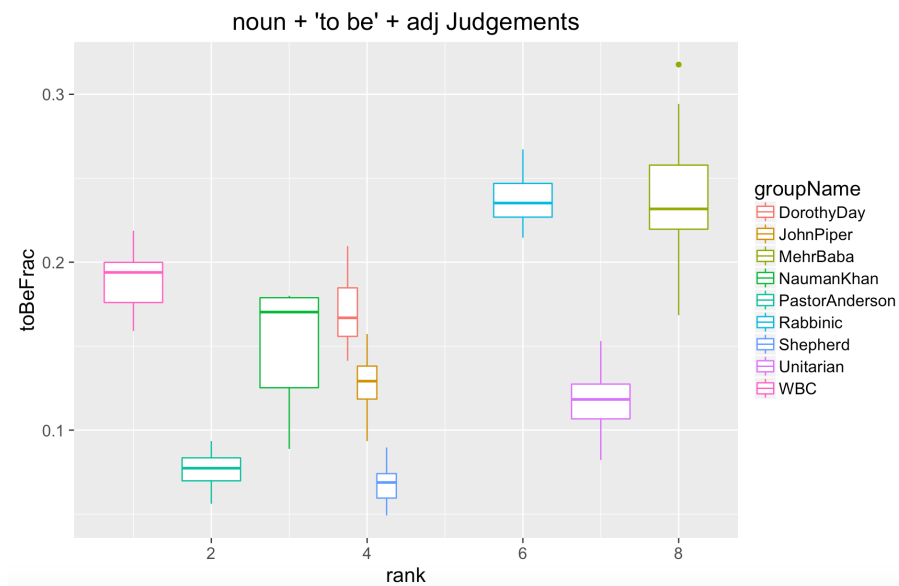
Problem

Data

Analysis

Preliminary Results

# Judgment Classification



Problem

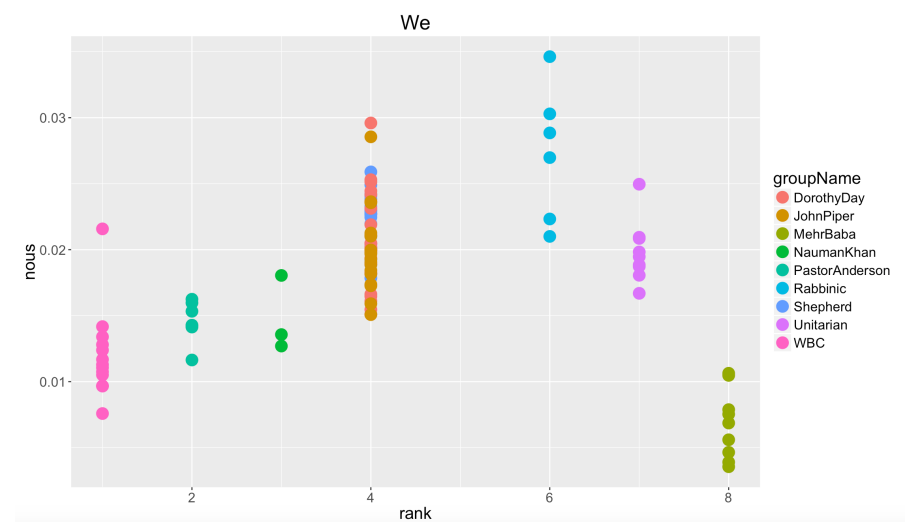
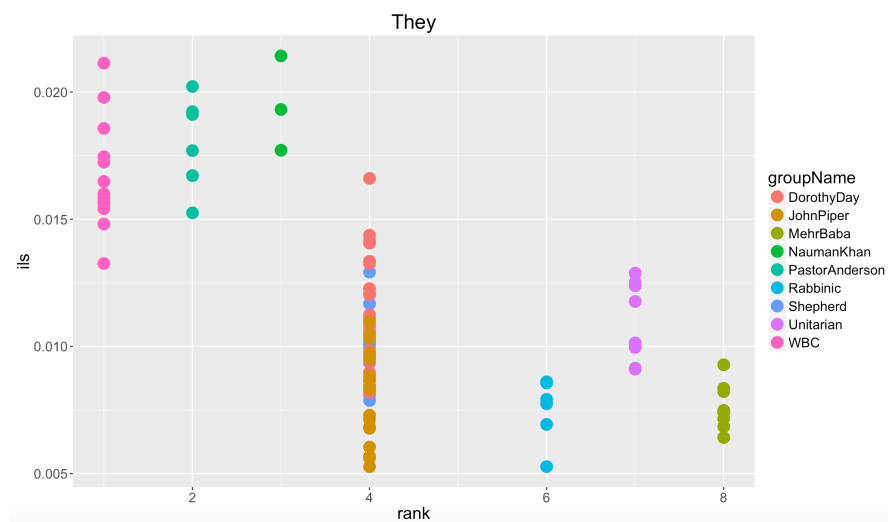
Data

Analysis

Preliminary Results



# Pronoun Analysis



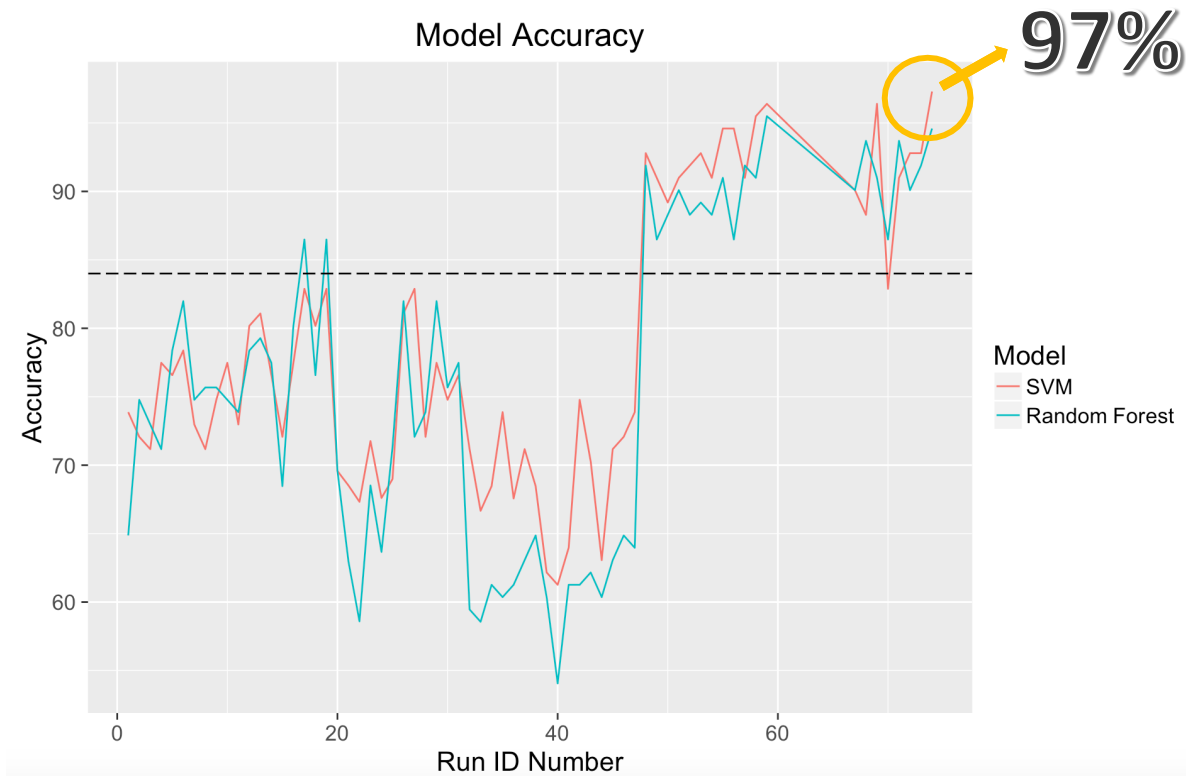
Problem

Data

Analysis

Preliminary Results

# Preliminary Results



Problem

Data

Analysis

Preliminary Results

Questions?

