

Tim Rappold

[tim.rappold@gmail.com](mailto:tim.rappold@gmail.com)

FOX NEWS VS MSNBC

---

**POLITICAL POLARIZATION IN  
CURRENT AFFAIRS SHOWS**

PREMISE:

# CHOICE OF TOPICS AS A DETERMINANT OF BIAS

---

# DATA

## ▶ Shows:

**MSNBC** ▶ Rachel Maddow Show

**MSNBC** ▶ The Last Word with Lawrence O'Donnell

**FOX** ▶ Hannity

**FOX** ▶ The Ingraham Angle

- ▶ 120 transcripts for each show scraped from [msnbc.com](https://www.msnbc.com) and [foxnews.com](https://www.foxnews.com)

## A HANNITY MONOLOGUE...

HANNITY: All right. Welcome to "Hannity". Stay with us for the hour.

Deep state is cracking right before your very eyes. According to new reports tonight, the fired FBI Deputy Director Andrew McCabe supplied Robert Mueller with a confidential personal memo detailing his so-called concerns about President Trump, the firing of James Comey. We will totally blow this out of the water and tell you how absurd this all is.

Plus, we're going to show you how this uncovered document reveals a stunning new revelation about the Deputy Attorney General Rod Rosenstein.

Also tonight, major developments surrounding the president and spygate. John Solomon is here. He will be breaking a brand new report detailing what are extreme measures the deep state took in order to secretly monitor the Trump campaign no other network has us.

MSNBC

FOX News

9pm

Weeknights



10pm

Weeknights



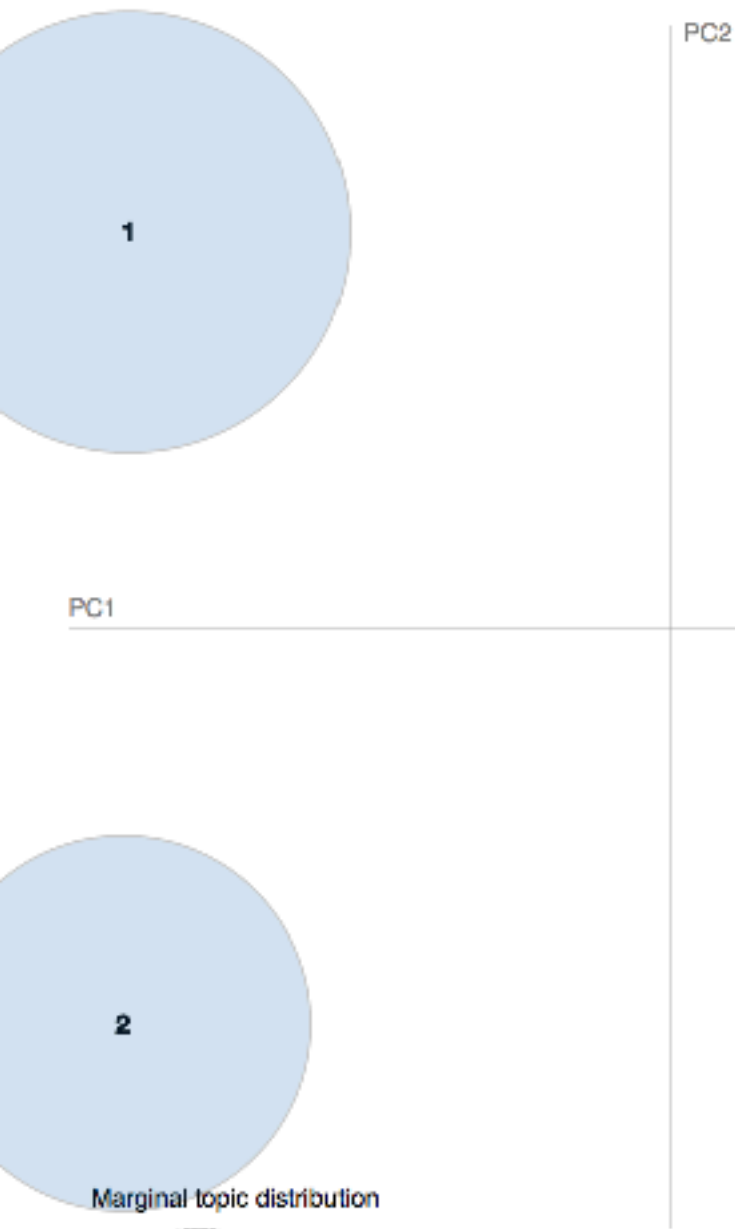
# TOPIC MODELING & TOPIC SEARCH

- ▶ Develop topic models for MSNBC and FOX shows
- ▶ A search tool that maps query to network

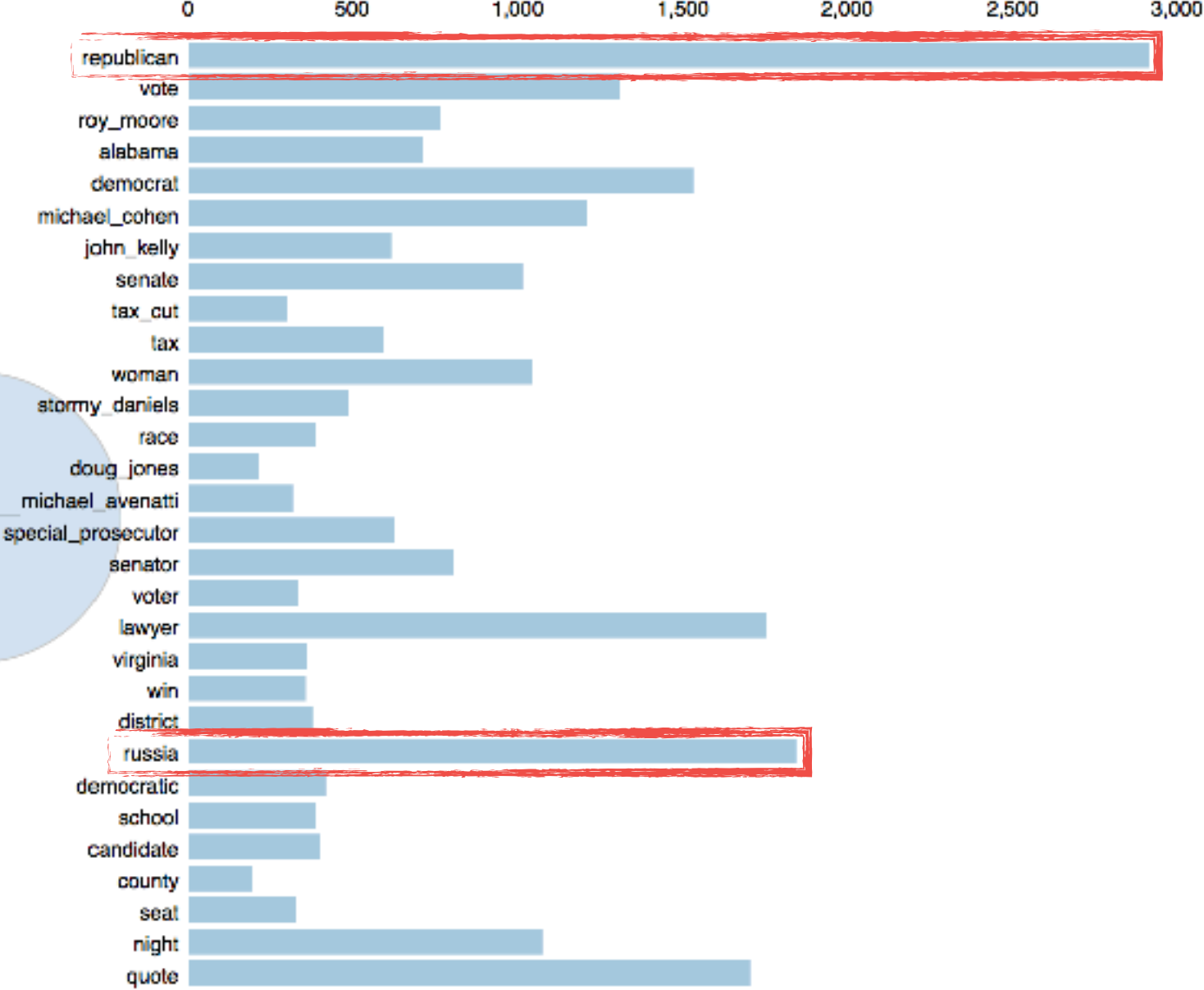
# OVERALL TERM POPULARITY

MSNBC

Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Salient Terms<sup>1</sup>



Overall term frequency

Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) \* [sum\_t p(t | w) \* log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)  
2. relevance(term w | topic t) = λ \* p(w | t) + (1 - λ) \* p(w | t)/p(w); see Sievert & Shirley (2014)



# OVERALL TERM POPULARITY

FOX

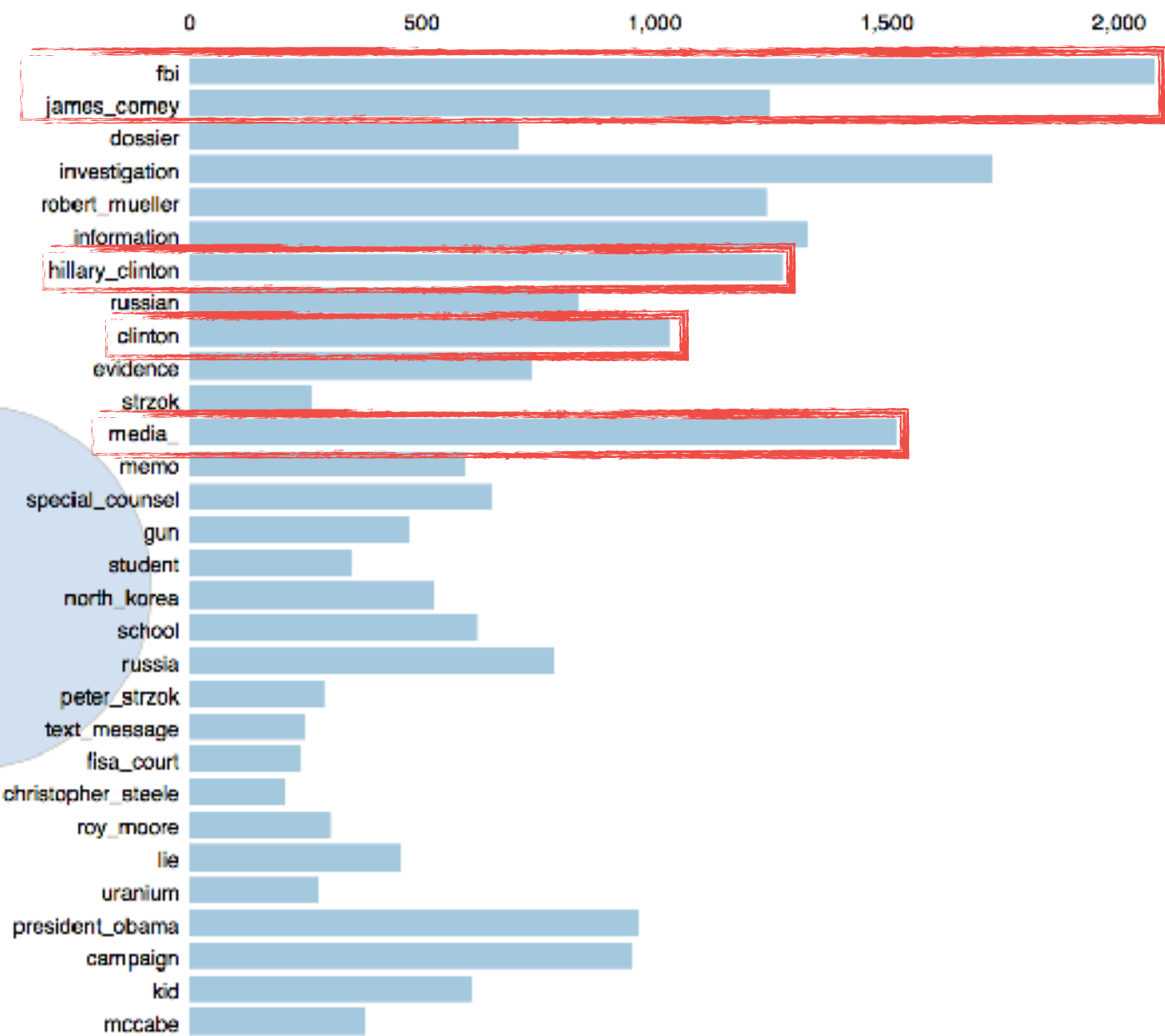
Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribution



Top-30 Most Salient Terms <sup>1</sup>



Overall term frequency

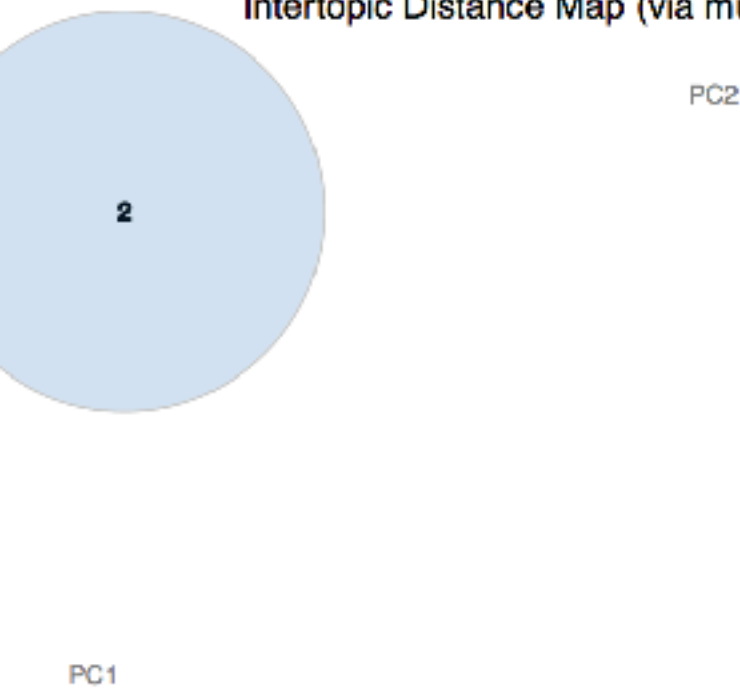
Estimated term frequency within the selected topic

1.  $saliency(terms, w) = frequency(w) * [\sum_t p(t | w) * \log(p(t | w) / p(t))]$  for topics  $t$ ; see Chuang et. al (2012)  
2.  $relevance(terms, w | topic, t) = \lambda * p(w | t) + (1 - \lambda) * p(w | t) / p(w)$ ; see Sievert & Shirley (2014)

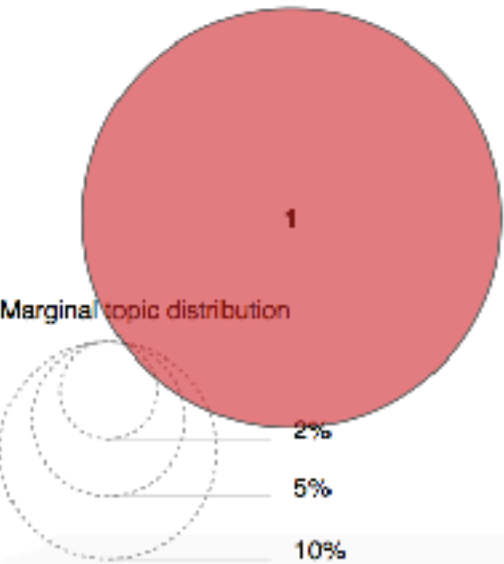


MSNBC

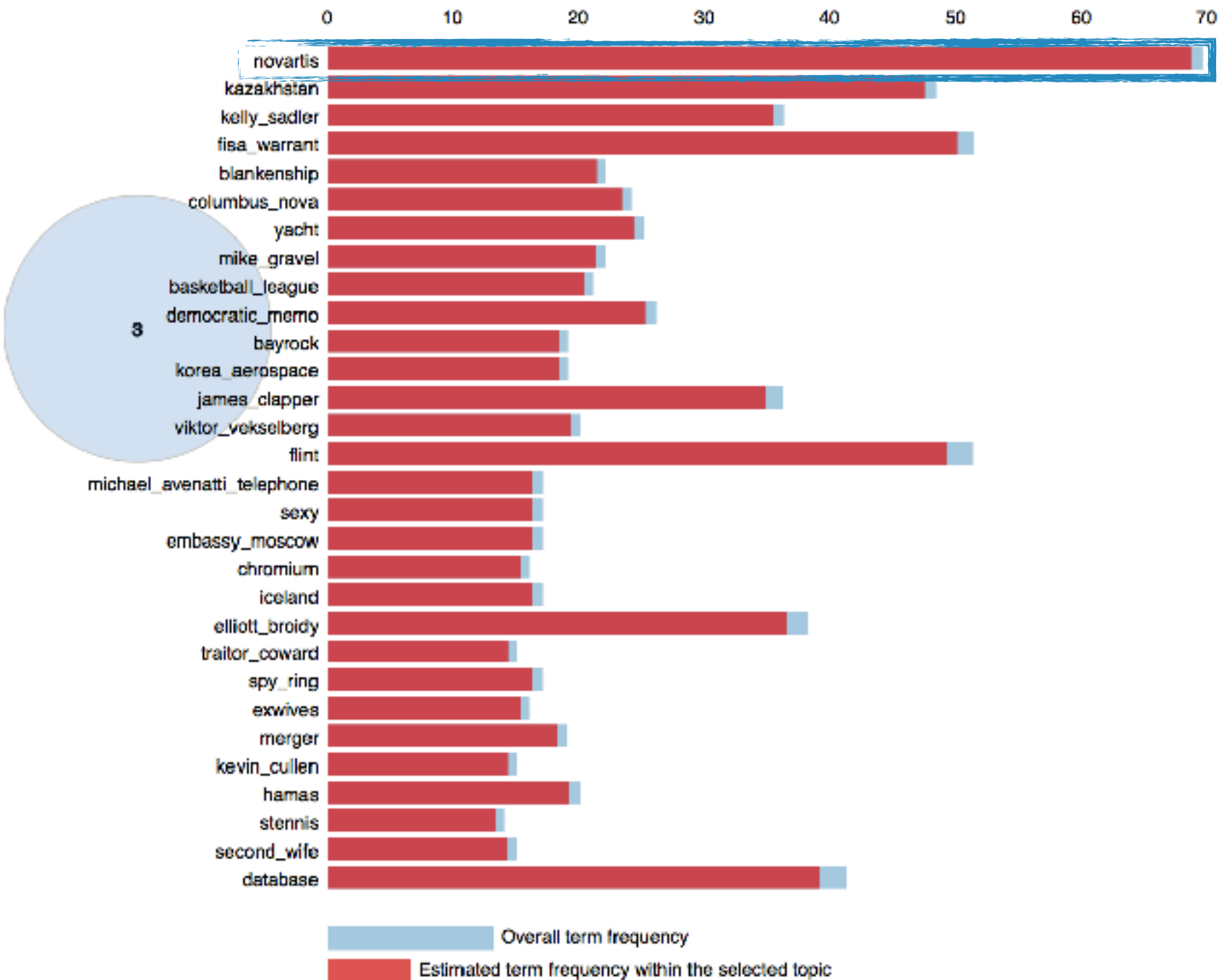
Intertopic Distance Map (via multidimensional scaling)



“CONSERVATIVE  
CORRUPTION”



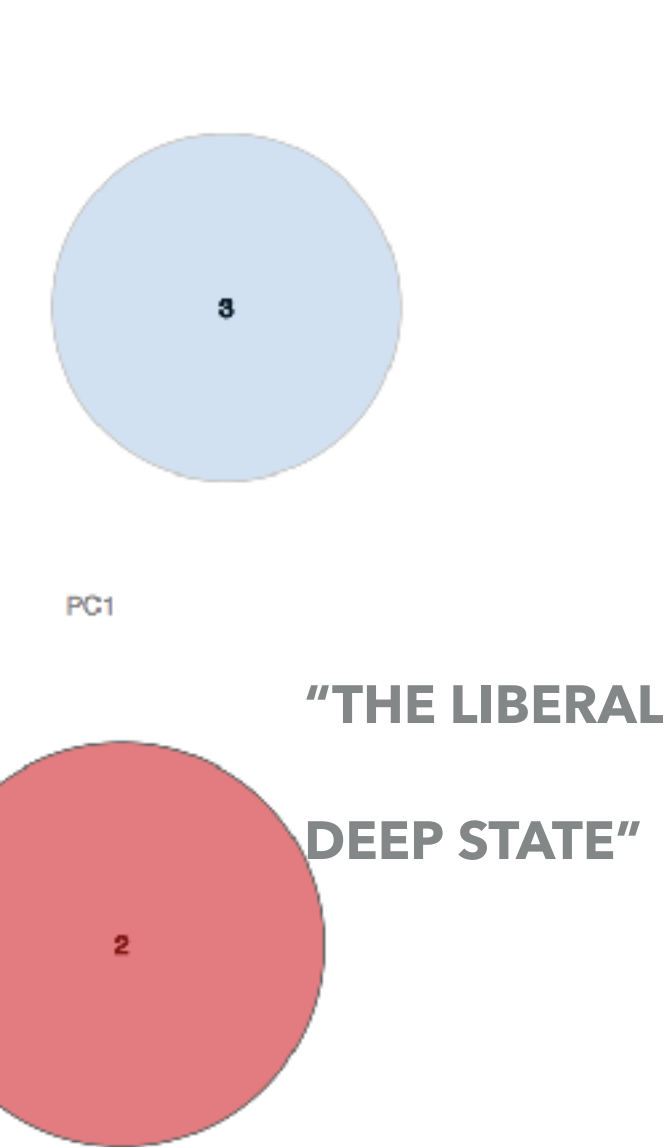
Top-30 Most Relevant Terms for Topic 1 (37.2% of tokens)



1. saliency(term w) = frequency(w) \* [sum\_t p(t | w) \* log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)  
2. relevance(term w | topic t) = λ \* p(w | t) + (1 - λ) \* p(w | t)/p(w); see Sievert & Shirley (2014)

FOX

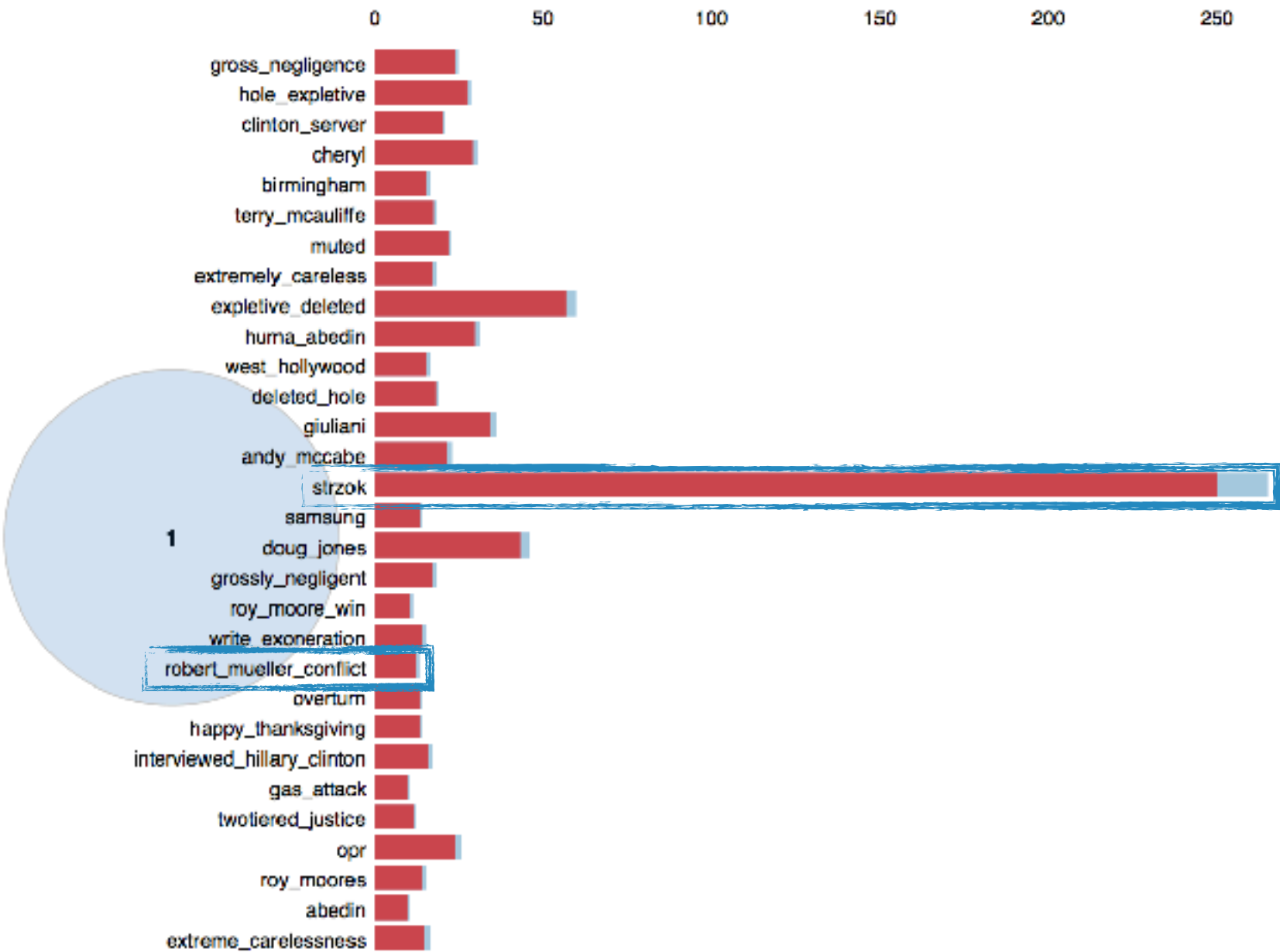
Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribution



Top-30 Most Relevant Terms for Topic 2 (34.7% of tokens)



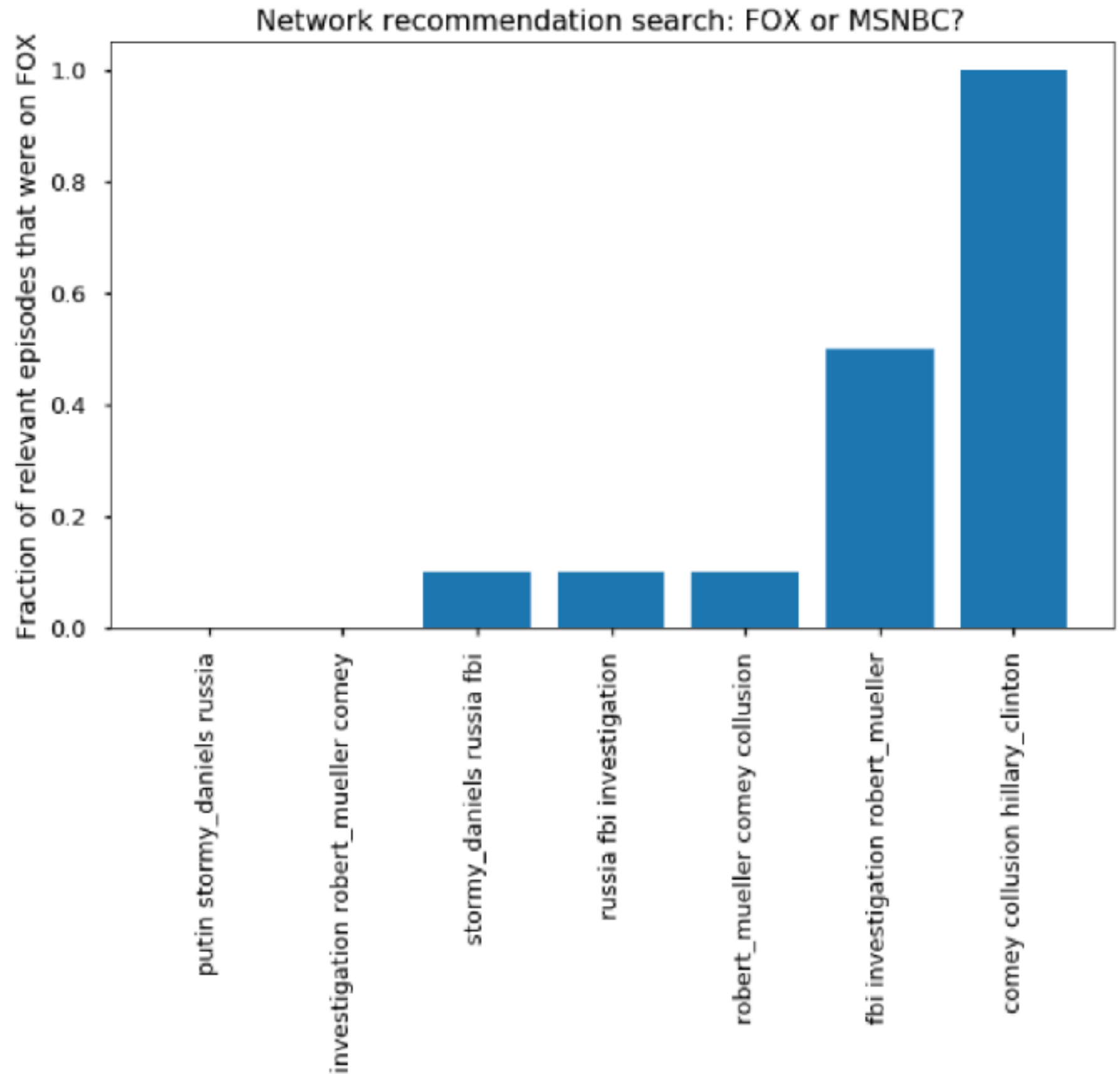
Overall term frequency

Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) \* [sum\_t p(t | w) \* log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)

2. relevance(term w | topic t) = lambda \* p(w | t) + (1 - lambda) \* p(w | t)/p(w); see Sievert & Shirley (2014)

WHAT TOPICS  
DO I WATCH  
WHERE?



---

## CONCLUSION & NEXT STEPS

- ▶ Clear political biases can be identified with LDA and LSI
- ▶ Next steps: deeper dive into topic grouping